Characterising longitudinal dynamics of infant neurodevelopment using functional change point detection

Authors & Affiliations

Samuel Beaton[1], Samantha McCann[1], Sarah Lloyd-Fox[2], Clare E. Elwell[3], Ebrima Mbye[4], Anna Blasi Ribera[3,†], Sophie E. Moore[1,4,†]

(1) Department of Women & Children's Health King's College London, London, UK
(2) Department of Psychology, University of Cambridge, Cambridge, UK
(3) Department of Medical Physics and Biomedical Engineering, University College London, UK
(4) Medical Research Council Unit The Gambia at the London School of Hygiene and Tropical Medicine, Fajara, The Gambia

†: Joint senior authors

Corresponding Author:
Samuel Beaton
Email: samuel.beaton@kcl.ac.uk
Address: Department of Women and Children's Health
School of Life Course and Population Sciences
Addison House
Guy's Campus
King's College London
Newcommen Street
London
SE1 1UL (UK)

Abstract

Habituation is an early-developing cognitive process linked to early learning whereby irrelevant stimuli are filtered out; it can be assessed using neuroimaging tools, including functional near-infrared spectroscopy (fNIRS). However, conventional fNIRS analysis approaches typically assume linear time invariance and analyse summary statistics of the data, potentially limiting insight into the nature of the changing evoked haemodynamic response. Functional data analysis (FDA) is a branch of statistics which treats entire mathematical functions as the unit of analysis rather than discrete points; within this framework, functional change point (FCPt) detection can identify significant changes in data structure, such as the mean haemodynamic response function. Using infant fNIRS data from a rural Gambian population (n= 204) recorded using a habituation and novelty detection paradigm, the use of functional change point detection with wild binary segmentation was implemented to identify structural changes in the evoked haemodynamic response. Group-level analysis from participants assessed at 5-, 8- and 12mo of age identified structural change corresponding to the evoked haemodynamic response in locations consistent with previous analyses. Further analyses indicated consistent habituation at 8- and 12mo and an age-related shift in the timing of decreasing haemodynamic responses, suggesting older infants in the first year of life habituate more rapidly to auditory stimuli. These results provide novel evidence to support the use of FDA and FCPt detection in the fNIRS field more widely, as well as improving understanding of the development of habituation during early infancy.

Keywords functional near-infrared spectroscopy, functional data analysis, functional change point detection, habituation, neurodevelopment, infancy

1 Introduction

1.1 Rationale

Habituation is an early-developing, low-level process of reducing attention to repeated, neutral stimuli (Gonzalez-Gonzalez et al., 2009; Nakano et al., 2009), conserving cognitive resources and potentially aiding survival as a consequence. It is linked to novelty response ('dishabituation'), an increased response to a novel stimulus, which may be unpleasant or hazardous (Sicard-Cras et al., 2022). Habituation is thought to direct attention to support early learning, perhaps through the construction of internal representations of repeated stimuli (Katus et al., 2023; Nakano et al., 2009).

Habituation paradigms are suitable for infant neuroimaging research since they do not rely on overt behaviours and the outcome of interest is cognitive rather than behavioural change (Katus et al., 2023); see Nordt and colleagues (Nordt et al., 2016) for a detailed discussion. Indeed, past infant functional near-infrared spectroscopy (fNIRS) research shows that the evoked infant cortical haemodynamic response can gradually decrease when stimuli are repeated (Bouchon et al., 2015; Lloyd-Fox et al., 2009; Lloyd-Fox et al., 2019; Nakano et al., 2009). However, fNIRS data analyses typically rely on averaging the signal across trials – time blocks concurrent with a task or stimulus – to reduce noise (Yücel et al., 2021). This is based on the linear time-invariance (LTI) assumption (Lindquist et al., 2009) that haemodynamic responses remain constant across trials and sum linearly with other physiological phenomena to give the recorded signal (Hemmati Berivanlou et al., 2014).

As an example, habituation in haemodynamic responses have previously been identified during an auditory fNIRS habituation task in infants in the first year of postpartum life (Blasi Ribera et al., 2025; Katus et al., 2023; Lloyd-Fox et al., 2019). The observed habituation of Gambian infants persisted even when introducing a novel stimulus intended to elicit a heightened response, indicating continued habituation. These previous results were generated using block averaging methods (Luke et al., 2021) to determine not only channels of activation but also condition contrasts between groups ('epochs') of 5 trials – see Lloyd-Fox and colleagues (2019) for more information. Whilst informative, this approach loses trial-specific and temporal information.

In contrast, functional data analysis (FDA), a branch of statistics concerned with the analysis of mathematical functions as the core objects (Wang et al., 2016), may enable trial-by-trial analysis utilising latent, temporal information from the entire response. This may enable improved differentiation between the habituation in haemodynamic response of individuals or groups, particularly useful given the absence of a clear novelty-response in this population.

Within FDA, discrete data can represent a regular sampling of a continuous, smooth underlying process allowing the function itself to be treated as the object of interest (Kokoszka & Reimherr, 2017a; Ramsay & Silverman, 2005). This representation reduces concerns about correlation between repeated measures and accounts for temporal structure (Gertheiss et al., 2024; Ullah & Finch, 2013). An important concept in FDA is the operator: an integral across the entirety of the domain (e.g. time or space) which maps one function to another in the same domain, usually altering or transforming it in some way (Hsing & Eubank, 2015). Operators in FDA are

analogous to mathematical functions in multivariate statistics, which act on matrices to alter their properties. Full overviews of FDA can be found in works by Kokoszka and Reimherr (2017), Ramsay and Silverman (2005) and Müller (2009), but relevant concepts are outlined briefly here for clarity, with additional detail found in Supplementary Materials 1.

Further, observed or measured curves, X_n, can be seen as realisations of a random function, X, leading to the concepts of mean and covariance functions (Muller, 2009). The group of approaches which can be used to determine points in functional series where these structural properties change (see Figure 1) is known as functional change point (FCPt) detection (Berkes et al., 2009)

During FCPt detection, a change point occurs when there is a significant difference in the selected structural property (e.g. the mean function) between data segments before and after the point; changes can be global (i.e. across the whole mean function) or more localised, and indicate trends in the data or a shift in the mean level of the function. During habituation, changes in the mean of the haemodynamic response are expected, alongside consistency in variance and shape (i.e. the form will still resemble a canonical haemodynamic response). FCPt detection was therefore applied in this work to the functional mean haemodynamic response, to examine habituation characteristics in infant fNIRS data.

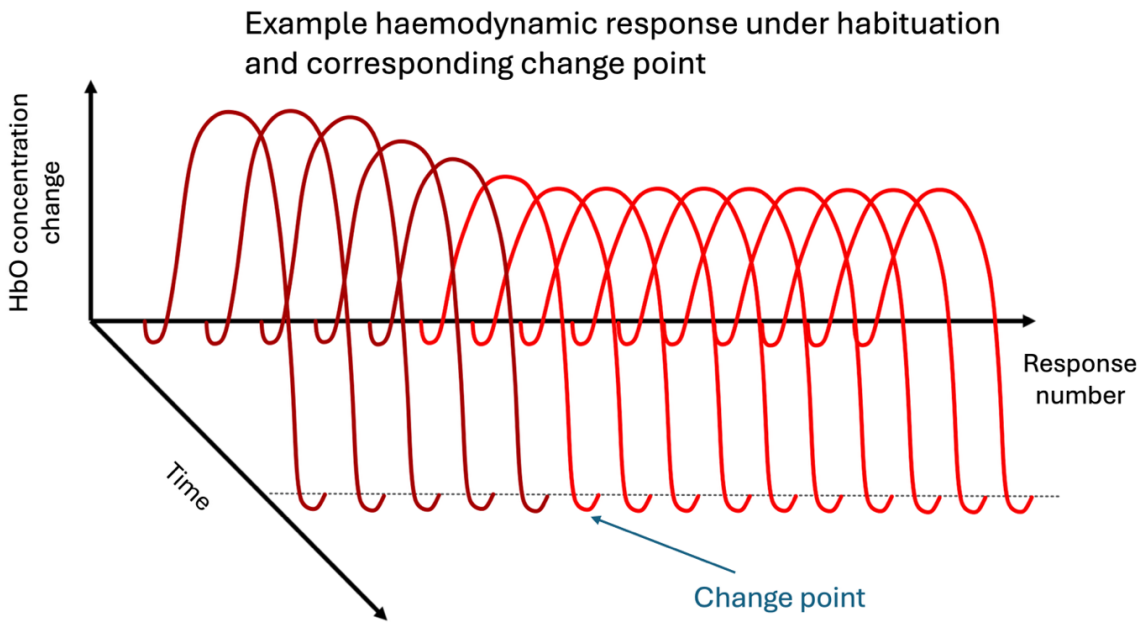

Figure 1: Hypothesized habituation of an evoked haemodynamic response and a corresponding example change point (FCPt).

Hypothesized FCPts occur when significant structural change is present in the functional representation of the haemodynamic response; they are represented at the points by a change in colour of the haemodynamic response functions. The set of consecutive evoked haemodynamic responses form a functional time series within FDA i.e. a series of functional objects.

Previous applications of FCPt detection in neuroimaging data to identify changes in the mean function have examined epidemic changes – where the mean function returns to its original, pre-change state – in fMRI data (Aston & Kirch, 2011,

2012, 2012). However, no work has utilised FCPt detection on either infant neuroimaging or fNIRS data, making this the first such example of either case.

A common approach to FCPt detection is to first perform dimension reduction, using functional principal component analysis (FPCA), a decomposition of the data analogous to ordinary principal components analysis (Kokoszka & Reimherr, 2017b, p. 41). However, structural change in modes orthogonal to the FPCs may be undetected using this approach (Aue et al., 2018). In contrast, fully functional (FF) change point detection methods, which utilise the data in their entirety, may be more effective, though they require more computational resources. Both FPCA-based and FF approaches were applied during analysis here, however, the FF approach yielded results more spatially aligned with previous analyses in comparison to the FPCA approach, with little difference in time taken to generate results. Accordingly, the FF approach and subsequent findings yielded by its implementation are outlined in the main body of this work; methods for FPCA-based FCPt detection and full results can be found in the Supplementary Materials.

1.2 Objectives

This work has three objectives:

(a) To compare the functionality of FCPt detection against existing analytical methods for infant fNIRS data

(b) To explore the ability of FCPt detection to gain more insight into the habituating infant haemodynamic response than has been previously possible using established approaches

(c) Use FCPt detection to infer longitudinal change in evoked infant habituation responses

With respect to these objectives, the following hypotheses were formulated:

(1) FCPt detection will identify channels similar to those which are identified by existing univariate methods

(2) FCPt detection will provide information on the timing of changes in structure related to habituation in a way which univariate approaches are unable to

(3) If age-associated differences exist in the timing of habituation, older participants will exhibit earlier reductions in haemodynamic response and/or completion of the habituation process

2 Methods

2.1 Participants

Data described herein were collected as part of the Brain Imaging for Global HealTh (BRIGHT) Project, which aimed to expand diversity in neurodevelopmental studies, establish brain function-for-age curves in infants from high- (UK) and low-resource (The Gambia) settings and explore the effects that contextual factors related to living in a low-resource context may have on infant development (Lloyd-Fox et al., 2023).

204 infants were enrolled into the BRIGHT Project in The Gambia. Families were identified using the West Kiang Demographic Surveillance System (Hennig et al., 2015), primarily from the village of Keneba and neighbouring villages in the rural

West Kiang District. In West Kiang, the predominant (79.9%) ethnicity is Mandinka (Hennig et al., 2015) therefore recruitment was limited to this group to avoid confounds related to study protocol translation. Ethical approval was given by the joint Gambia Government-MRCG Ethics Committee (SCC 1451v2). Prior to participation, informed consent was obtained from all parents/carers in writing or via thumbprint if they were unable to write. Mothers were initially seen at a 34-36 weeks antenatal visit and assessment study visits took place at 7-14 days, 1-, 5-, 8-, 12-, 18- and 24mo of age thereafter. For the current analysis, data collected at 5-, 8- and 12mo of age (hereafter referred to as $x$mo for x-months of age) is used.

Post-natal exclusion in BRIGHT included infants who were born preterm, however a minimum birth weight was not specified for Gambian infants. Anthropometric measures (head circumference, weight, and length/height) were collected and used to calculate z-scores. These included head circumference (HCZ), weight-for-height (WHZ), weight-for-age (WAZ) and height-for-age (HAZ). Low WHZ, WAZ and HAZ scores (< 2 standard deviations below the WHO reference) indicate that a child is wasted, underweight, or stunted, respectively.

2.2 Experimental procedures

2.2.1 Data collection

Testing took place between June 2016 and July 2019 at the Medical Research Council Unit The Gambia at the London School of Hygiene and Tropical Medicine (MRCGLSHTM; www.mrc.gm).

Custom-made headgear was fitted after head measurements (head circumference, and ear-to-ear both around the forehead and over the top of the head) had been taken to aid with the alignment of fNIRS headgear with the 10/20 system anatomical landmarks. The headgear was designed to record responses bilaterally from auditory associative brain regions (see Figure 2), including the inferior frontal gyrus (IFG), middle and superior temporal regions, and the tempero-parietal junction (Collins-Jones et al., 2021; Lloyd-Fox et al., 2014; Lloyd-Fox et al., 2019). The fNIRS array comprised 17 channels in each hemisphere with a source-detector (S-D) distance of 2cm, corresponding to a penetration depth of ~1cm from the scalp surface and thus permitting measurement of the gyri and superficial sulci (Lloyd-Fox et al., 2019). fNIRS data were collected using this array design with the NTS optical topography system (Gowerlabs Ltd., UK, source wavelengths of 850 and 780nm).

During data acquisition, infants sat on a carer's lap. Carers were discouraged from interacting with the infant to minimise risk of confounding stimuli. However, infants' attention was engaged by (non-social, non-auditory) bubble-blowing and silent demonstration of soft toys, which helped minimise infant head movement. The Habituation and Novelty Detection (HaND) task was part of a larger battery of fNIRS assessments with a total recording time of at least ~21min 30sec (6min social task; 4min functional connectivity data acquisition; 7min 30sec HaND; 4min further functional connectivity) (Benerradi et al., 2025; Bulgarelli et al., 2024; Lloyd-Fox et al., 2019). The protocol for 8- and 12mo infants was even longer, as it also included a working memory and deferred imitation task. Where possible, tasks were completed uninterrupted; sessions were paused and subsequently resumed in the event of infant discomfort.

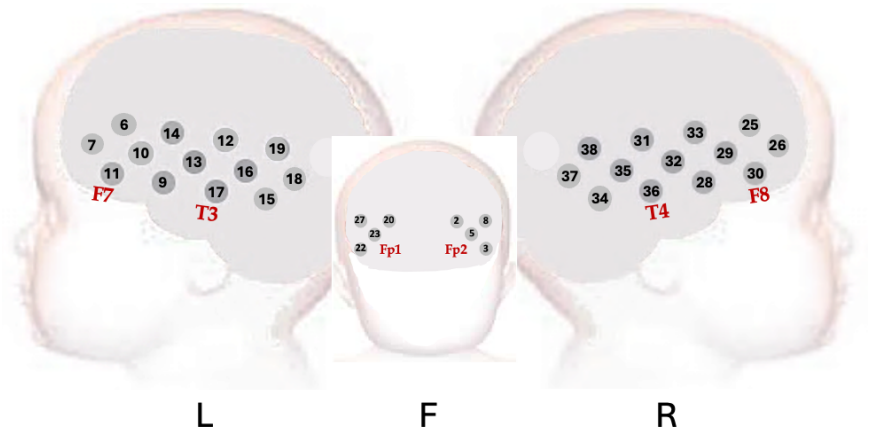

Figure 2: Array layout during data collection.

2.2.2 The HaND paradigm

The aim of this paradigm is to enable measurement of habituating hemodynamic responses related to a familiar stimulus and subsequent increase in response related to a novel stimulus Lloyd-Fox et al., (2019).

The experimental paradigm consisted of 25 trials, each consisting of the same spoken 8sec sentence in Mandinka followed by 10sec of silence - traditionally considered the minimum amount of time necessary for the hemodynamic response to return to baseline (see Figure 3). Working with a battery of multiple tasks with infant participants meant time allocation for tasks was constrained, so the inter-stimulus interval (ISI) was chosen to be the bare minimum possible whilst still theoretically enabling a return to baseline after the hemodynamic response.

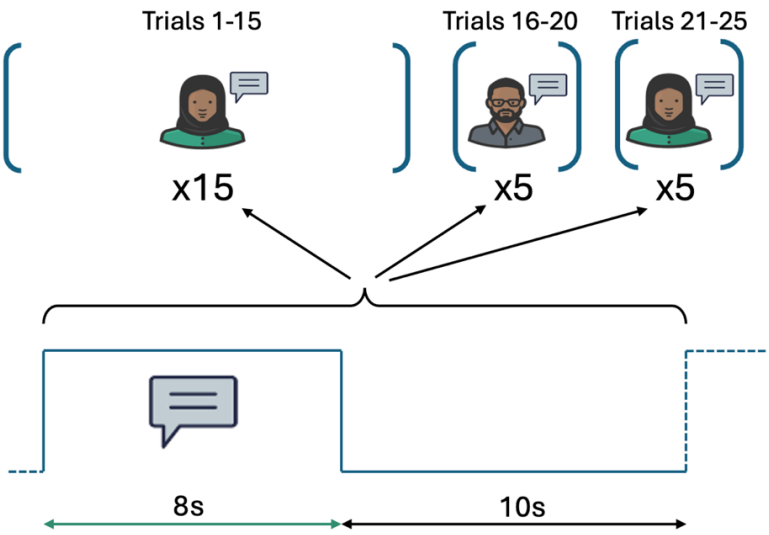

Figure 3: The HaND paradigm in the BRIGHT Project, with 25 repetitions of an identical spoken 8s sentence, followed by a 10s ISI.

Trial 1-15 and 21-25 have the same speaker; in trials 16-20 the identity and sex of the speaker is changed.

The same recording, with a female voice, was used for trials 1-15 (the Habituation trials); a different, male voice was used for trials 16-20 (the Novelty trials); finally, the original, female recording was again used for trials 21-25 (the Post-Test trials). The sentence in Mandinka: *"Denano a be nyadii. I be kongtan-rin? Abaraka bake ela naa kanan njibee bee, n kontanta bake le ke jeh"* roughly translates to: *"Hi baby! How are you? Are you having fun? Thank you for coming to see us today. We're very happy to see you".* Technical detail on the recording, processing and playback of the auditory stimulus can be found in previous work by Lloyd-Fox and colleagues (2019) and Katus and colleagues (2023)

2.3 Data processing

First, channels with poor optode coupling were pruned. Initially channels were removed if the signal was weak or saturated, indicated by a mean signal amplitude less than 3×10^{-4} or greater than 1×10^{7}, respectively based on the NTS system specifications. Subsequently, the coefficient of variation was calculated on a pre-selected motion-free section of the data for remaining channels, which were pruned if this value was below 8% (Lloyd-Fox et al., 2009). Infants were excluded if more than 40% of channels were pruned (Lloyd-Fox et al., 2019). Channel signals were then converted from light attenuation to optical density, enabling motion artefact correction using the Homer2 software package (Huppert et al., 2009).

Motion artefact correction was conducted using the spline interpolation approach (Scholkmann et al., 2010) followed by wavelet denoising (Molavi & Dumont, 2012). Spline interpolation corrects pre-identified motion artefacts channel-by-channel by estimating the true signal based on neighbouring data, via interpolation. The motion-free estimation is subtracted from the recorded signal; signal continuity is maintained by aligning the signal before and after the artefact, which also corrects baseline shifts (Cooper et al., 2012). Motion artefacts were identified using *hmrMotionArtifactByChannel* with parameters *tMotion = 1, tMask = 1, STDEVthresh = 15* and *AMPthresh = 0.4.* Following previous findings and the developers' recommendations, a parameter of $p = 0.99$ was used for spline interpolation (Di Lorenzo et al., 2019; Scholkmann et al., 2010). Wavelet denoising decomposes the signal into constituent wavelets, removing those with characteristics corresponding to abrupt changes caused by motion, before subsequent reconstruction. An interquartile range factor of 0.8 was implemented for the function *hmrMotionCorrectWavelet* (Di Lorenzo et al., 2019). Persistent motion artefacts were subsequently identified using *hmrMotionArtifactByChannel* as above; trials within 4 seconds after or 12 seconds before detected motion artefacts were removed. Channels with fewer than 60% (3/5) valid trials in either the first (trials 1-5), second (trials 6-10) or third (trials 11-15) blocks were removed. Participant exclusion was conducted again, with infants excluded if fewer than 60% channels remained after channel pruning and motion artefact removal.

Consistent with previous work with this data, a low-pass filter was applied (0.8Hz frequency cutoff) and a first-degree polynomial was fitted and subtracted from the data for each trial, in each channel, for each participant, to remove linear trends (Blasi Ribera et al., 2025; Lloyd-Fox et al., 2019). The low-pass filter removes high-frequency oscillations associated with physiological confounds such as the participant heart rate. Instead of high-pass filtering on optical density data (Blasi Ribera et al., 2025) linear detrending of the chromophore data was used, since high-pass filtering

retained low-frequency drift evident during exploratory analysis using FPCA (see Supplementary Materials 2). Optical density data was converted to oxy- (HbO) and deoxyhaemoglobin (HbR) data using the modified Beer-Lambert law and age-appropriate differential pathlength factors to account for age-related tissue composition changes (Delpy et al., 1988; Scholkmann & Wolf, 2013). Linear detrending was applied to concentration change data on a trial-by-trial basis, with each trial consisting of 2 seconds of recorded data pre-stimulus onset (baseline) and 18 seconds post-stimulus onset (Lloyd-Fox et al., 2019).

2.4 Functional change point detection

Unless otherwise stated, change point detection was implemented in Python 3.9.16 (Python Software Foundation, 2021) using in-house code.

The aim of FCPt detection is to identify structural breaks in a functional time series (see Figure 1) which are subsequently tested for statistical significance. This approach was used to look for fundamental haemodynamic response changes in the HaND data. The FCPt detection approach itself is described here, with full mathematical detail in Appendix A, and was applied to cross-infant mean-averaged data per channel.

The fNIRS time series data represents a discretely sampled measurement of the underlying functional (haemodynamic response) process across some time range T. For each HaND paradigm trial ($n = 1, \dots, N$, where $N = 25$), functional data, $Y_n(t)$ with sampled data at $t = \{T_1, \dots, T_2\}$, is expressed in the FDA framework as:

$$Y_n(t) = \mu(t) + \delta(t) \cdot \mathbf{1}_{\{n>n_c\}} + \epsilon_n(t)$$

Equation 1

where, $\mu(t)$ is the mean function, $\epsilon_i(t)$ is the measurement or systemic error and $\delta(t)$ indicates a structural break at trial $n = n_c$. The null hypothesis:

$$H_0\colon \delta(t) = 0 \text{ for all } n$$

Equation 2

is tested. If there are no structural breaks, H_0 holds. Under the alternate hypothesis, H_A, $\delta(t) \neq 0$ for all $n > n_c$ if there is a change point at some trial $n = n_c$.

FF FCPt detection uses a test statistic, $T_N^F(n)$, calculated for each trial number, n. In intuitive terms, $T_N^F(n)$ measures how much the average haemodynamic response up to each trial differs from the overall average across all trials. These cumulative differences fluctuate randomly around zero in the case of a stable underlying haemodynamic response or when no consistent pattern is present in the data. Under H_A, when the shape or magnitude of the mean response changes – such as when the haemodynamic response declines with habituation – the cumulative difference grows sharply due to deviation within each trial caused by $\delta(t)$ in Equation 1. The trial number, n_c, at which this divergence reaches its largest value marks the most likely location of a structural change in the data. In practice, this "cumulative-sum" (CUSUM) approach captures the trial at which the signal's average behaviour shifts most strongly over time (full mathematical formulation in Appendix A, Equation 7).

Each channel/chromophore combination contains a functional time series of 25 functions corresponding to the 25 trials in the HaND paradigm. For each of these

channel/chromophore combinations, candidate change points, n^*, for each functional time series were identified, corresponding to trial numbers where the largest structural change occurred. Candidate FCPts were identified as the smallest n^* such that $n^* = \min\left(\arg\max_{1 \le n \le N} T_N^F(n)\right)$: the smallest (earliest occurring) trial number which produces the largest cumulative difference from the overall functional mean in each channel/chromophore combination. Each n^* was then subsequently tested for statistical significance.

Following the recommendations of the developers of the FF method, standard FDA approaches were adopted and the fNIRS data functional object in the trial window was estimated using a set of 21 basis spline functions (Aue et al., 2018; Muller, 2009; Ramsay & Silverman, 2005, pp. 46–47).

2.5 Statistical analysis

Statistical significance of candidate change points was assessed via Monte Carlo simulation, comparing observed test statistics to those derived from Brownian Bridge processes. Multiple comparison correction was applied at each hierarchical level to control the family-wise error rate and ensure robustness of the findings.

2.5.1 Wild binary segmentation

The FF-based change point detection method detects a single change point; binary segmentation can be used to detect multiple FCPts using this same detection method (Rice & Zhang, 2022). During binary segmentation, if a statistically significant FCPt is detected in a functional time series, the time series is split into two subintervals either side of this change point and the process repeats recursively until no significant FCPts remain (see Figure 4). Binary segmentation is efficient, however it can overlook closely spaced change points (Fryzlewicz, 2014; Killick et al., 2012).

Wild binary segmentation (Fryzlewicz, 2014) addresses this by generating multiple subintervals of different lengths in each data segment; detecting candidate change points in each; and recording the largest test statistic T_N^F. The candidate change point n^* for the entire current data interval corresponds to the location of the largest test statistic across all subintervals.

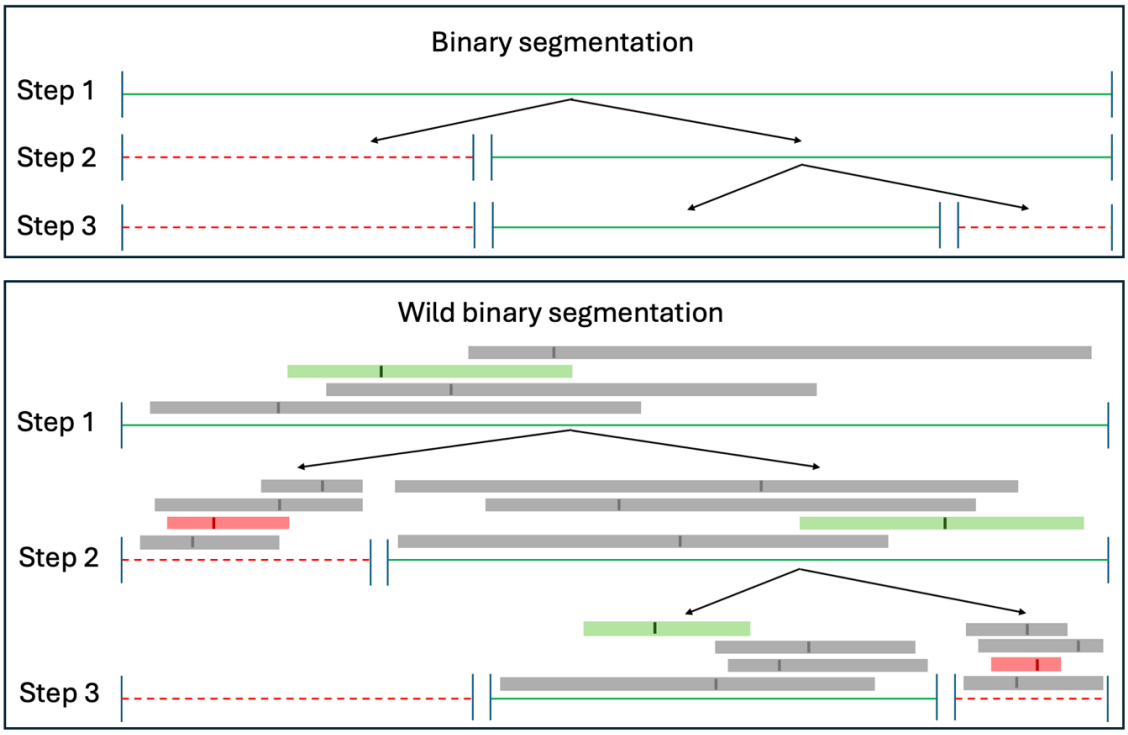

Figure 4 Binary segmentation and wild binary segmentation.

Green lines at each step represent segments with significant changepoints; red dashed lines represent segments with non-significant change points. For wild binary segmentation, grey bars represent sampled sub-segments, with candidate change points marked. Green bars represent the sub-segment of data containing the point with the highest CUSUM statistic value, marked by a vertical line. Red bars represent the sub-section with the highest CUSUM statistic value (marked) which did not meet statistical significance.

2.5.2 Monte Carlo simulation

The test statistic, $T_N^F(n)$, asymptotically converges to a weighted Brownian Bridge distribution under H_0, (Berkes et al., 2009; Kiefer, 1959); for full Mathematical detail see Appendix B. Put plainly: as N increases, the test statistic under H_0 is capped by the largest value in a weighted sum of random fluctuations (Brownian Bridges) which are tethered to zero at each end (see Figure 5), with the weights corresponding to different components of data variability.

The convergence in distribution provides a possible method for significance testing using empirical critical values of the distribution itself (Horváth & Kokoszka, 2012b). In this work, to permit multiple comparison corrections, it was necessary to calculate exact significance values, so the distribution of T_N^F was approximated using Monte Carlo simulation. Statistical testing of the significance of candidate change points n^* was therefore conducted by comparing their associated test statistics to the maximum values of $M = 10\,000$ generated using summed Brownian Bridges of length N_b, where N_b is the length of the functional time series interval being assessed at each stage of the recursive wild binary segmentation process. The statistical significance of

$T_N^F(n^*)$ was then given by the proportion of the M randomly generated values larger than $T_N^F(n^*)$.

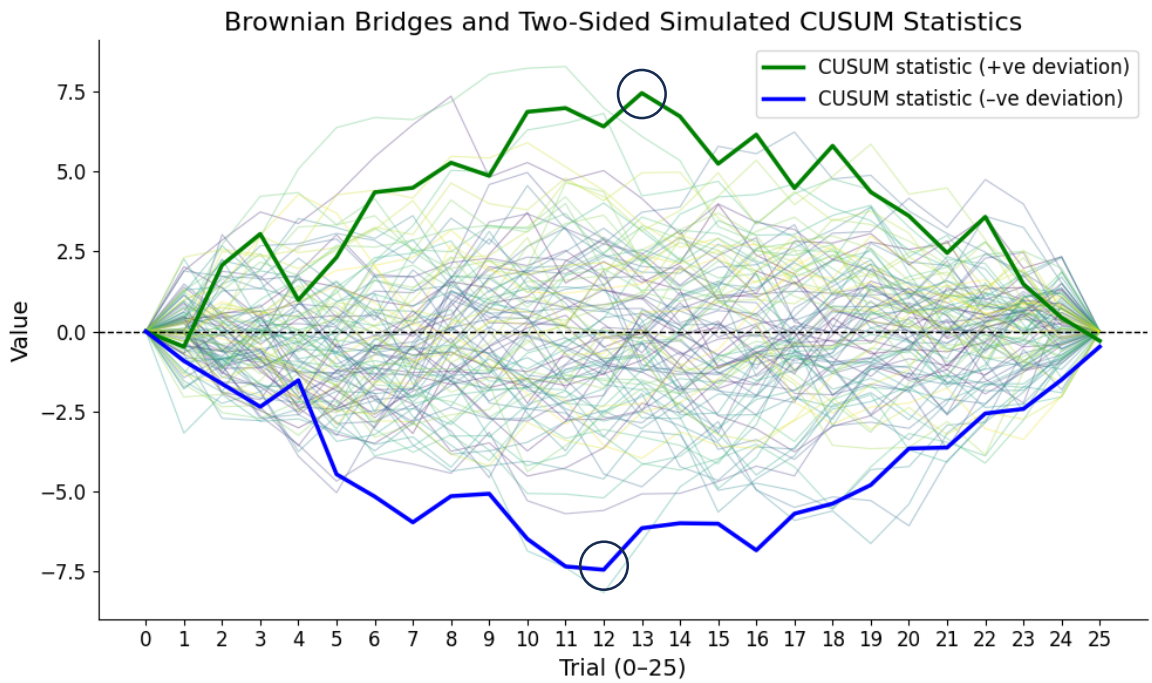

Figure 5: 100 simulated weighted sums of Brownian Bridges, anchored at the start and end points, with 25 steps.

The maximal value of each generated summed bridge were recorded to create a 'null' distribution to compare against the maximal partial sum values across the 25 HaND trials. In this instance, the deviations of the positive (green) and negative (blue) CUSUM statistics would likely be significant, given the absolute magnitude of their largest values (circled).

Each comparison with Brownian Bridge distributions during wild binary segmentation amounts to a separate statistical test and increases the family-wise error rate (FWER). To correct for this, within-channel multiple comparison correction using the Bonferroni-Holm method (Holm, 1979) was conducted, which required the generation of significance values. To address the hierarchical nature of the binary segmentation process, in which the importance of change points at higher levels is predicated on the statistical significance of those which occur beforehand (see Figure 4) a step-down multiple comparison procedure was conducted within each channel at every binary search level. Each hierarchical level, l, of binary segmentation has a maximum of 2^{l-1} candidate change points; correction for a maximum of $\sum_l 2^{l-1}$ tests at each level was therefore conducted, corresponding to the change points at the current and all preceding levels. If a test at any level failed to meet the significance threshold post-correction, all descendant tests were discarded. The Bonferroni-Holm method was applied at each level using the *statsmodels* Python package (Holm, 1979; Seabold & Perktold, 2010). An alpha value of 0.01 was used to determine the statistical significance of change points.

2.5.3 Assessing channel overlap with previous analysis

For this specific dataset, previous work has been published which used t-test contrasts (Lloyd-Fox et al., 2019) and spatio-temporal clustering (Blasi Ribera et al., 2025) to locate channels co-located with activation. FCPt detection is intended to complement analytical approaches to determine the location of evoked haemodynamic responses, as FDA methods already exist for one- and two-sample problems (Horváth & Kokoszka, 2012a; Pini & Vantini, 2017). Nevertheless, consistent structure across a set of functions is required for large test statistic values, which in this case would indicate a temporal structure time-locked to the HaND paradigm's stimulus presentation. Correspondence between 'change point' channels using the approach described and activation channels from previous analyses would indicate FCPt detection can identify underlying physiological changes corresponding to the haemodynamic response.

Accordingly, agreement was assessed between channels with detected change points and those channels in which a haemodynamic response was identified across the first 5 trials, derived from previous work using the spatiotemporal clustering method Threshold-Free Cluster Enhancement (TFCE) (Blasi Ribera et al., 2025; Smith & Nichols, 2009). Using the survival function of the hypergeometric distribution, implemented with the Python class *scipy.stats.hypergeom* (Virtanen et al., 2020) the probability of observed overlap between the set of FCPt channels and the set of channels with a haemodynamic response according to TFCE was assessed. The channels identified for (i) HbO, (ii) HbR, and (iii) both chromophores were compared.

This analysis seeks to check inherent sensitivity to the data structure and detect meaningful overlap with prior work, so controlling the false discovery rate was prioritised rather than strict control of the FWER. Consequently, Benjamini-Hochberg's method for multiple comparison correction was conducted across the three age groups using the *statsmodels* Python package (Benjamini & Hochberg, 1995; Seabold & Perktold, 2010).

2.5.4 Assessing characteristics of change points

The motivation for using FCPt detection is to better understand changes in the haemodynamic response during the course of the experimental session. Consequently, timing of change points alone provides incomplete information, so the nature of structural change in the mean haemodynamic response (i.e. whether the change corresponds to an alteration in response magnitude) was also considered.

Accordingly, to determine whether the mean haemodynamic response increased or decreased following change points, the mean haemodynamic response was calculated before and after each statistically significant FCPt, then Simpson's rule was used to estimate the definite integral of the curve's absolute values, using *scipy.integrate.simpson* (Virtanen et al., 2020). The definite integral estimate after each change point was subtracted from that before it, as an estimate of area-under-the-curve (AUC) difference between functions; a positive value indicated a decrease in response. The proportion and statistical significance of the number of decreasing change points – suggesting habituation – was assessed using two-sided binomial tests, comparing the observed proportion of decreasing change points against a null hypothesis of random directionality. Tests were conducted on the full dataset plus

three subgroups defined by age, with Bonferroni-Holm correction for multiple comparisons to control the FWER (Holm, 1979).

2.5.5 Analysing the effect of age

The effect of age group was then examined by comparing the distributions across age groups and chromophores of the trial numbers of:

(i) First change points

(ii) Final change points

(iii) Decreasing change points

(iv) Decreasing change points which were ***also*** the first to occur in a channel ('First And Decreasing')

(v) Decreasing change points which were ***also*** the last to occur in a channel ('Final And Decreasing')

(vi) The earliest occurring decreasing change points in a channel ('Earliest Decreasing')

(vii) The latest occurring decreasing change points in a channel ('Latest Decreasing')

Distributions (vi) and (vii) were generated to account for scenarios in which an increase in the haemodynamic response may be present before or after habituation and which would therefore be discounted from distributions (iv) and (v).

Weighted ordinal logistic regression (OLR) was used to assess inter-age differences between trial numbers for each distribution (i)-(vii), supplemented with post-hoc testing and multiple comparison corrections. OLR (McCullagh, 1980) preserves the ordinal nature of the outcome (FCPt Trial Number); and was conducted using the *OrderedModel* from the *statsmodels* package (Seabold and Perktold, 2010). The predictors 'Age' and 'Change Point Count Per Channel' (the sum of age/chromophore combinations with identified change points, per channel, ranging from 0 to 6) were used to investigate associations with age and control for Channel differences. Weights were applied based on participant contributions to each functional observation in question, to account for varying participation contributions across observations.

A post-hoc Wald test (Wald, 1943) confirmed any significant impacts of age on change point outcomes. Pairwise differences between age groups driving significant age effects were identified using an ANOVA model followed by Tukey's Honestly Significant Difference (HSD) (Tukey, 1949) to control the FWER.

3 Results

Sample size, data retention, reasons for exclusion or withdrawal, and descriptive statistics for remaining fNIRS data are detailed in Figure 6.

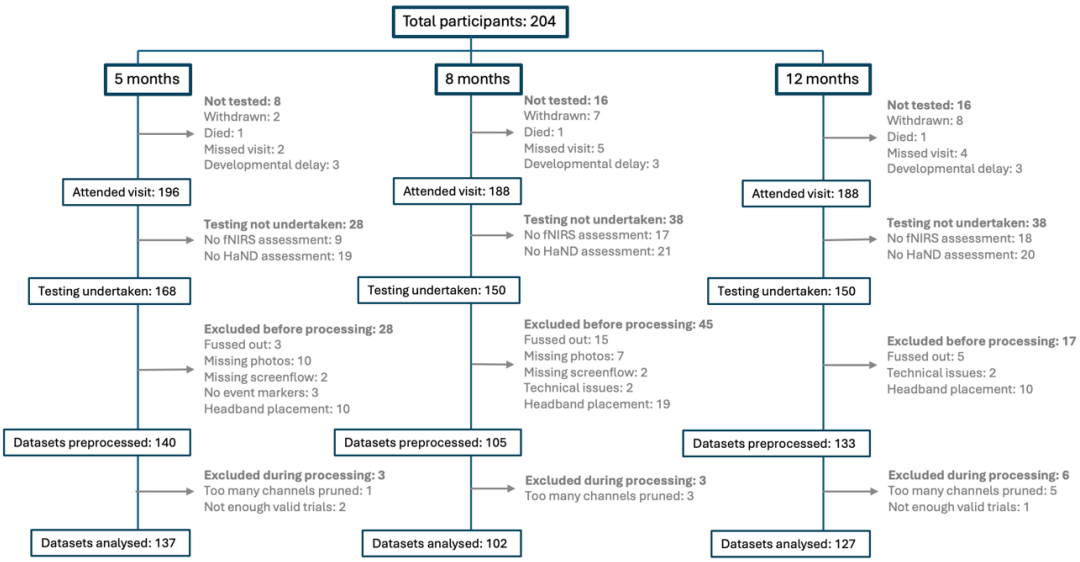

Figure 6: Data retention for each of the 3 ages during HaND fNIRS assessment and processing.

3.1.1 Functional representation of the haemodynamic response

Qualitatively, functional data representations of the evoked haemodynamic responses were found to be faithful to their discrete representations. Further, FCPt detection was able to both capture the structural representation of the evoked responses and identify changes in the mean function, even in channels which had not previously been identified as containing significant haemodynamic responses by the TFCE analysis used in Blasi et al., 2025 (see Figure 7).

At all 3 ages, each of the two Hb chromophores yielded at least 10 channels with statistically significant FCPts amongst the 25 paradigm trials (see Table 1). Statistically significant FCPts were identified in both chromophores in 6 channels at all ages, though these were not the same 6 channels at each age. 2 channels consistently yielded FCPts in both chromophores at all ages, with one in each hemisphere located in the same temporal location bilaterally (see Figure 8).

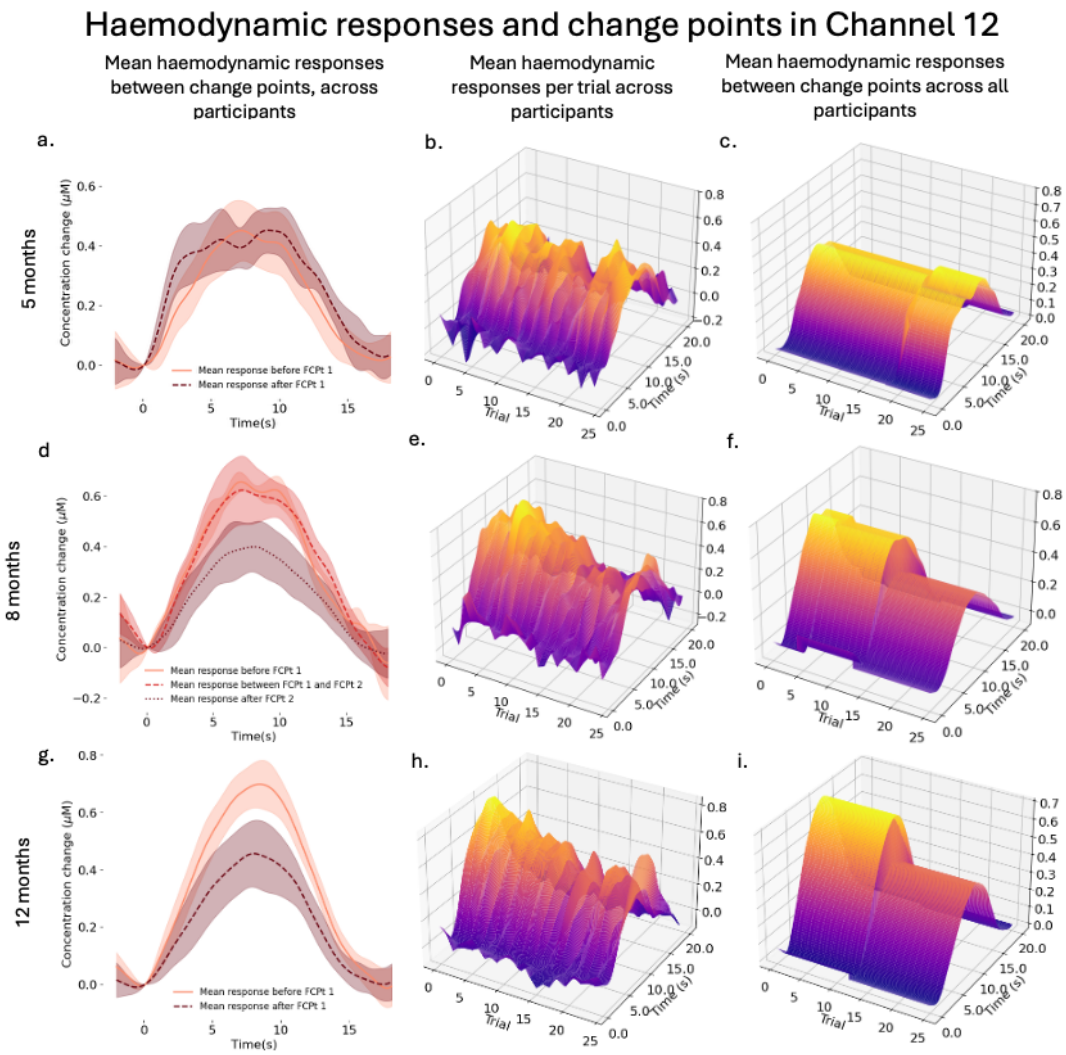

Figure 7: Haemodynamic responses, and their mean representations between FCPts, in a single channel (channel 12)

See Figure 8 for channel location. The channel demonstrates a clear haemodynamic response structure at all three ages (in both chromophores – see Supplementary Materials 9). Figures (c), (f) and (i) display the mean functional haemodynamic response between FCPts for the data in figures (b), (e) and (h), respectively. Figure (c) illustrates the capacity to distinguish between haemodynamic responses based on structural information rather than solely the magnitude of the sampled fNIRS values. Figures (a), (d) and (g) illustrate an increasing difference pre- and post-FCPt with age.

3.1.2 Channel overlap with previous analysis

Age-specific ROI channels

Overlap was significant at the 0.05 threshold for channels identified in both chromophores using the FF method at 8- ($p = 0.043$) and 12mo ($p = 0.013$), as shown in Table 1 and Figure 8. Statistical significance at this threshold was also exhibited at 12mo for HbO data ($p = 0.030$). Beyond this, channels identified for both

chromophores overlapped with a p-value just above the 0.05 threshold at 5mo ($p = 0.053$).

Table 1: Overlap between FF FCPt detection and ROI channels outlined using TFCE.

			Overlap					
			Age specific TFCE channels			Overall TFCE channels		
Age	Chromophore	n	K	k	p-value (adj.)	K	k	p-value (adj.)
5mo	HbO	15	*11*	*7*	*0.112*	*10*	*6*	*0.307*
	HbR	10	*8*	*4*	*0.154*	**8**	**4**	**0.047***
	Both	6	*6*	*3*	*0.053*	**6**	**4**	**0.004****
8mo	HbO	10	*9*	*5*	*0.090*	*10*	*4*	*0.316*
	HbR	14	*6*	*5*	*0.078*	**8**	**7**	**0.012***
	Both	6	**5**	**3**	**0.043***	**6**	**4**	**0.004****
12mo	HbO	12	**10**	**7**	**0.030***	**10**	**7**	**0.031***
	HbR	11	*8*	*5*	*0.078*	*8*	*5*	*0.052*
	Both	6	**6**	**4**	**0.013***	**6**	**4**	**0.004****

Key: K = number of age-specific channels identified by TFCE; n = number of channels identified by FCPt detection; k = number of channels which overlap between the two methods. Significance in overlap rounded to 3 significant figures (p<0.05; p**<0.01).*

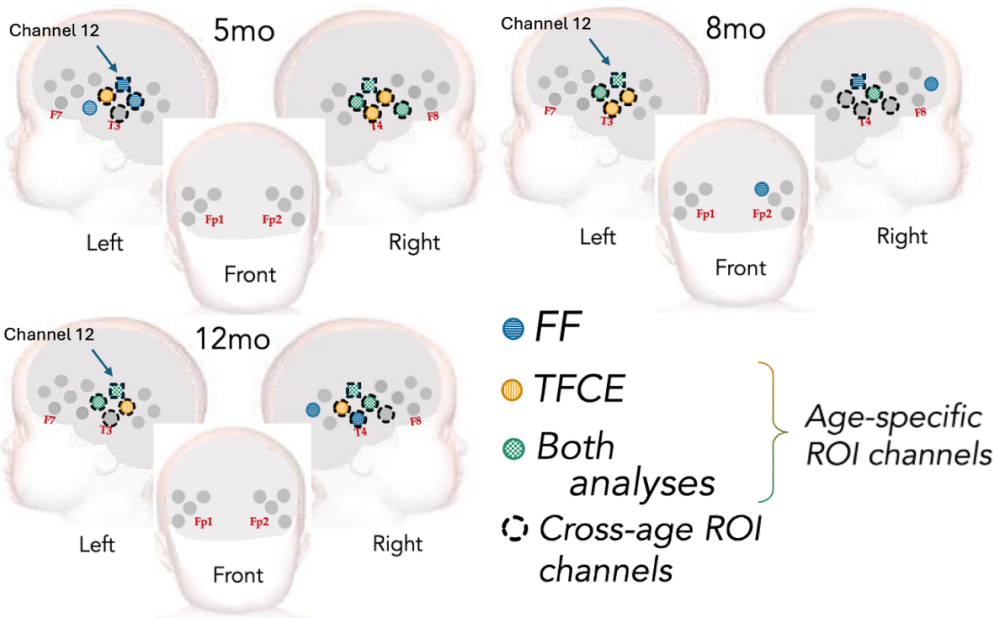

Figure 8: Channel overlap between FCPt and TFCE methods.

Blue, horizontal stripes are channels with identified FCPts. Orange, vertical lines represent age-specific channels with significant haemodynamic responses for both chromophores identified in the first 5 trials of the HaND paradigm, using the TFCE approach. Green, checkerboard represents channels which overlap between the two analyses. Black, dashed-circled channels are ROI channels for all 3 ages. Black, dashed-line square icons represent channels with identified FCPts in both chromophores for all 3 ages. Channel 12 is identified as it's FCPts are included in Figure 7.

Cross-age, chromophore-specific ROI channels

Channels containing significant change points in both chromophores were primarily sensitive to the superior- and middle temporal gyri (Figure 8), as identified using channel registration in previous work (Collins-Jones et al., 2021), and exhibited typical haemodynamic response structures (see Figure 7 and Figure 9), even in channels in which a haemodynamic response had not been previously identified for both chromophores. Outside of these regions, change points were observed in one channel at 12mo in the posterior right hemisphere and two channels at 8mo in the left inferior frontal gyrus. Channels within the superior- and middle temporal gyri deemed to contain FCPts often corresponded with those where a significant haemodynamic response had previously been identified using TFCE analysis for one of the three ages, even if the channel identified was not matched for age across both analyses. This raised the possibility that each of the methods may be identifying relevant information about the haemodynamic response overlooked by the other.

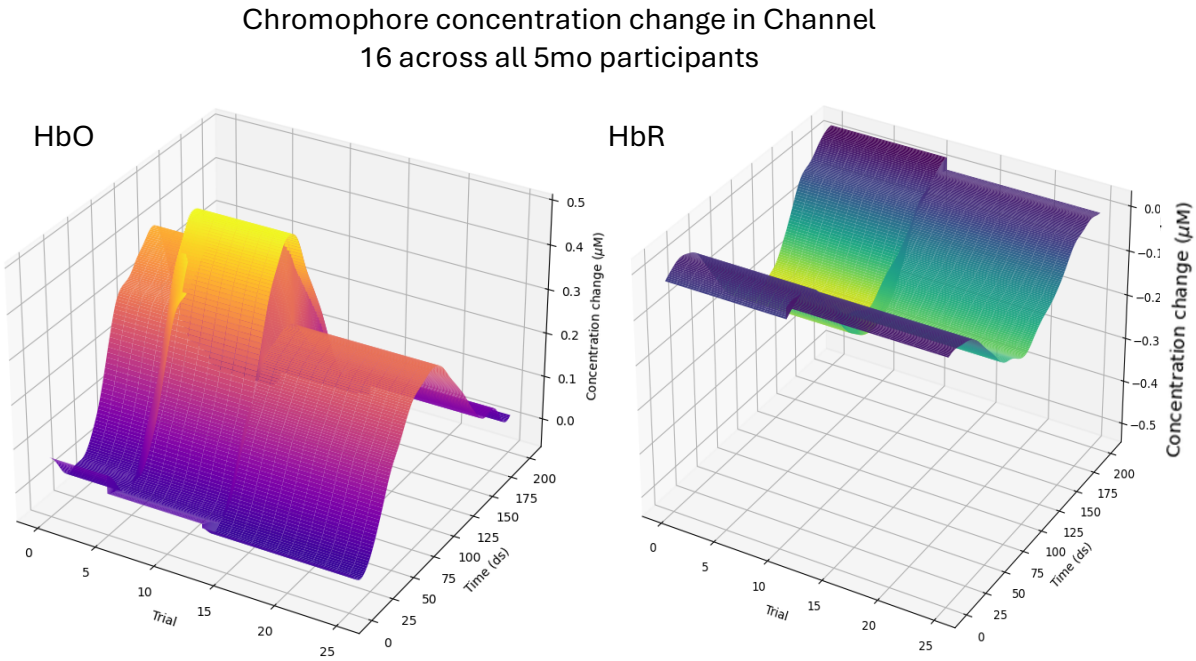

Figure 9: Change points and mean surface between them for a cross-age ROI channel using 5mo data.

The channel is not an age-specific channel as prior analysis did not identify a significant response in HbR data for this age, yet it still shows clear haemodynamic response structure and habituation profile in both chromophores. It also exhibits an early apparent increase in the HbO response.

Consequently, an overall, age-agnostic set of ROI channels was created for (i) HbO, (ii) HbR, and (iii) both chromophore signals, by combining the sets of channels identified by TFCE as exhibiting a haemodynamic response in the first 5 trials at any of the three ages. The survival function of the hypergeometric function was used to assess statistical significance of the overlap between age- and chromophore-specific change point channels and these three sets of chromophore-specific, cross-age ROI channels (see Table 1; full results in Supplementary Materials Table S2).

Significant overlap was found between channels in which haemodynamic responses were exhibited in the initial trials, and those where change points were identified, in both chromophores for all three ages ($p = 0.004$). There was also additional statistical overlap between the HbR channels at 5- ($p = 0.047$) and 8mo ($p = 0.012$) in addition to the HbO channels at 12mo identified during comparison with channels ($p = 0.031$, unchanged).

3.1.3 Characteristics of change points

Analyses for both the characteristics and age effects of detected FCPts are reported for the age-specific and the cross-age ROI channels identified for both chromophores, to avoid biasing analyses with changes unrelated to the evoked haemodynamic response in channels spatially distinct from the ROI channels (see Supplementary Materials 9).

Age- and chromophore-specific ROI channels

Change points in age-specific ROI channels identified by TFCE analysis were first examined. The results are summarised in Table 2. For the set of combined change points across all ages, the proportion of all detected change points, first change points, and last change points which decreased were all significantly higher than expected by chance (p-values < 0.001, $p = 0.006$, and $p = 0.017$, respectively). At 5mo, no statistically significant differences were observed in the proportion of decreasing change points across all categories of change points (p-values > 0.05). At 8- and 12mo, there were significant proportions of decreasing change points when considering all detected change points (p-value $= 0.006$ for both ages) however no statistical significance was observed for the sets of first or last change points in each channel, for either age (p-values > 0.05).

Cross-age, chromophore-specific ROI channels

Expanding to a set of cross-age TFCE ROI channels (Figure 8), similar significance testing was conducted. For the set of combined change points across all ages, the proportion of all detected FCPts, first FCPts, and last FCPts which decreased were all significantly higher than expected by chance (p-values < 0.001, $p = 0.005$, and $= 0.013$, respectively), similarly to the age-specific ROI channels. For infants at 5mo, again, no statistical significance was observed in the proportion of decreasing FCPts across all categories of change points (p-values > 0.05). At 8mo, all change point categories (all detected, first, last) contained a significant proportion of decreasing change points ($p < 0.001$, $p = 0.013$, $p = 0.013$, respectively). At 12mo, the last detected change points in each channel now contained a significant proportion which decreased ($p = 0.023$) alongside the significant proportion of all detected change points ($p = 0.005$), however decreasing change points within the set of first change points still did not meet significance ($p = 0.167$).

Table 2: Proportion of decreasing FCPts which occurred first, last, or anywhere during the channels' time functional series

		Age-specific TFCE channels				Cross-age TFCE channels			
Age	Change points	Total	Decreasing	Proportion (%)	p-value (adj.)	Total	Decreasing	Proportion (%)	p-value (adj.)
All	ALL	**43**	**37**	**86.0**	**0.001*** **	**43**	**37**	**84.1**	**0.001*** **
	First	**21**	**18**	**85.7**	**0.006** **	**21**	**18**	**85.7**	**0.005** **
	Last	**21**	**17**	**80.1**	**0.017* **	**21**	**17**	**81.0**	**0.013* **
5mo	ALL	*8*	*4*	*50*	*1*	*13*	*8*	*61.5*	*0.688*
	First	*4*	*3*	*75*	*0.682*	*6*	*4*	*66.7*	*0.688*
	Last	*4*	*0*	*0*	*0.150*	*6*	*2*	*33.3*	*0.688*
8mo	ALL	**10**	**10**	**100**	***0.006*** **	**16**	**16**	**100**	**0.001*** **
	First	*5*	*5*	*100*	*0.083*	**8**	**8**	**100**	**0.013* **
	Last	*5*	*5*	*100*	*0.083*	**8**	**8**	**100**	**0.013* **
12mo	ALL	**10**	**10**	**100**	**0.006** **	**14**	**13**	**92.3**	**0.005** **
	First	*5*	*5*	*100*	*0.083*	*7*	*6*	*85.7*	*0.167*
	Last	*5*	*5*	*100*	*0.083*	**7**	**7**	**100**	**0.023* **

Significance reported after Bonferroni-Holm correction (p<0.05; p**<0.01, p***<0.001). For the cross-age ROI channels, all types except for the first occurring at 12mo were significant. Proportions for the 5mo data were more varied, with no FCPt category containing a significant number of decreasing changes either in age-specific or cross- age ROI channels.*

3.1.4 Effect of Age

Weighted ordinal logistic regression was used to examine the effect of infant age on the timing of various types of change points. Change Point Count Per Channel was used as a covariate and participant counts for each functional observation as weights. The model was fit using the logit distribution and optimised using the Broyden–Fletcher–Goldfarb–Shannon algorithm.

Age- and chromophore-specific ROI channels

Results are reported fully in Table 3 and Table 4. For age-specific ROI channels, the age coefficient was significant for the Final ($\beta = -0.626, SE = 0.209, z = -3.00,\ p = 0.003$), Decreasing ($\beta = -0.272,\ SE = 0.117, z = -2.33,\ p = 0.020$), Final And Decreasing ($\beta = -0.495,\ SE = 0.218, z = -2.27,\ p = 0.023$), and Latest Decreasing change points ($\beta = -0.424,\ SE = 0.176, z = -2.41,\ p = 0.016$). Wald tests confirmed the significance of these differences (see Table 4).

Table 3: Weighted ordinal logistic regression results for FCPts.

	Age-specific				Cross-age			
Change points	β	S.E.	z	p (adj.)	β	S.E.	z	p (adj.)
First	*-0.228*	*0.147*	*-1.54*	*0.122*	*-0.228*	*0.147*	*-1.54*	*0.122*
Final	**-0.626**	**0.209**	**-3.00**	**0.003****	**-0.626**	**0.209**	**-3.00**	**0.003****
Decreasing	**-0.272**	**0.117**	**-2.33**	**0.020***	**-0.247**	**0.114**	**-2.17**	**0.030***
First And Decreasing	*-0.216*	*0.165*	*-1.30*	*0.192*	*-0.216*	*0.165*	*-1.30*	*0.192*
Final And decreasing	**-0.495**	**0.218**	**-2.27**	**0.023***	**-0.497**	**0.218**	**-2.28**	**0.022***
Earliest Decreasing	*-0.195*	*0.145*	*-1.34*	*0.180*	*-0.195*	*0.145*	*-1.34*	*0.180*
Latest Decreasing	**-0.424**	**0.176**	**-2.41**	***0.016****	**-0.359**	**0.175**	**-2.06**	***0.040****

Detected across both chromophores, after Bonferroni-Holm correction. Significant effects for age were found in the change points which decreased, particularly those occurring later in the series of 25 trials. Significance: p<0.05; p**<0.01, p***<0.001.*

Table 4: Post-hoc testing of change point categories for which significant channel effects were identified using weighted ordinal logistic regression.

		Wald test		5 & 8 pairwise		5 & 12 pairwise		8 & 12 pairwise	
ROI Channels	Change points	$\chi^2(1)$	p	M	p	M	p	M	p
Age-specific	Final	**9.00**	**0.003****	**-5.21**	**<0.001*****	**-5.48**	**<0.001*****	*-0.268*	*0.951*
	Decreasing	**6.00**	**0.014***	**-3.58**	**0.026***	**-3.66**	**0.029***	*-0.087*	*0.997*
	Final And Decreasing	**5.15**	**0.023***	**-4.54**	**0.006****	**-4.81**	**0.005****	*-0.268*	*0.957*
	Latest Decreasing	**5.80**	**0.016***	**-2.88**	**0.046***	**-3.14**	**0.033***	*-0.268*	*0.966*
Cross-age	Final	**9.00**	**0.003****	**-5.21**	**<0.001*****	**-5.48**	**<0.001*****	*-0.268*	*0.951*
	Decreasing	**4.70**	**0.030***	*-3.15*	*0.062*	*-3.24*	*0.066*	*-0.087*	*0.997*
	Final And Decreasing	**5.21**	**0.024***	**-4.54**	**0.006****	**-4.81**	**0.005****	*-0.268*	*0.957*
	Latest Decreasing	**4.23**	**0.040***	**-3.08**	**0.049***	**-3.34**	**0.036***	*-0.268*	*0.967*

Wald testing confirmed the significance of Age effects in all cases. Significant pairwise differences were found between 5- and both 8- and 12mo data for all categories tested, aside from the decreasing FCPts identified in the cross-age ROI channels, suggesting the habituation process continues for longer at younger ages. Significance: p<0.05; p**<0.01, p***<0.001.*

In all four cases, Tukey's HSD results showed that differences between age groups were driven by later change point occurrences at 5mo than the two later time points. Final change points at 5mo occurred a mean of 5.21 trials later than at 8mo ($p < 0.001$) and 5.48 trials later than at 12mo ($p < 0.001$). Decreasing change points occurred a slightly smaller 3.58 trials later than at 8mo ($p = 0.026$) and 3.66 trials later than at 12mo ($p = 0.029$). More specific analysis of late-occurring, decreasing change points showed that change points which were both Final And Decreasing had mean trial occurrences 4.54 later at 5mo when compared to 8mo ($p = 0.006$) and 4.81 trials later when compared to 12mo ($p = 0.005$). Similarly, the Latest Decreasing change points occurred an average of 2.88 trials later at 5mo than at 8mo ($p = 0.046$) and 3.14 trials later than at 12mo ($p = 0.033$). Interestingly, the earliest occurring FCPts were relatively consistent across age groups, with no other significant results found.

These differences across ages in the later change point occurrences are best exemplified by the differences between the Latest Decreasing change points (see Figure 10; similar figures for other change point types can be found in Supplementary Materials 6), which shows both the smaller mean trial number for the two older age groups and the greater distribution breadth of responses found in younger infants. No significant effects of Change Point Count Per Channel were found for any of the change point types.

Cross-age, chromophore-specific ROI channels

Similar results were found when expanding the set of ROI channels so that they represented a cross-age set of ROI channels for both chromophores. Weighted ordinal logistic regression results for age-related differences in the timing of Final change points was identical (see Table 3) – as were the post-test findings (see Table 4) – whereas both the effect size and statistical significance of age on Decreasing ($\beta = -0.247$, $SE = 0.114, z = -2.17$, $p = 0.030$), Final And Decreasing ($\beta = -0.497$, $SE = 0.218, z = -2.28$, $p = 0.022$), and Latest Decreasing ($\beta = -0.359$, $SE = 0.175, z = -2.06$, $p = 0.040$) change points all decreased slightly. Wald tests confirmed the significance of these effects; pairwise differences were again driven by later occurrences at 5mo than the two later Ages for Final And Decreasing (5- & 8mo: $M = -4.54, p = 0.006$; 5- & 12mo: $M = -4.81, p < 0.001$) and Latest Decreasing (5- & 8mo: $M = -3.08, p = 0.049$; 5- & 12mo: $M = -3.34, p = 0.036$). No significant pairwise effects were found for Decreasing change points. No significant effects for the Change Point Count Per Channel covariate were identified for any of the FCPt types.

Trial-wise decline in haemodynamic response

To visualise this trend with age directly, Figure 11 shows the change in mean haemodynamic response AUC across the 25 trials, calculated using Simson's rule for integration. The overall decline, particularly evident at 8- and 12mo, mirrors the pattern of earlier instances of decreasing FCPts identified in the analyses above. It also illustrates greater magnitude of decrease in size for the two older age groups, as well as the trajectory of decrease in which a sharp, step-like drop occurs after a stable initial phase and before stabilising again at a steady lower level after about 15 trials. These phenomena are exhibited to a lesser extent at 5mo, though the decrease is smaller in size and reaches a more unstable plateau towards the end of the series of paradigm trials.

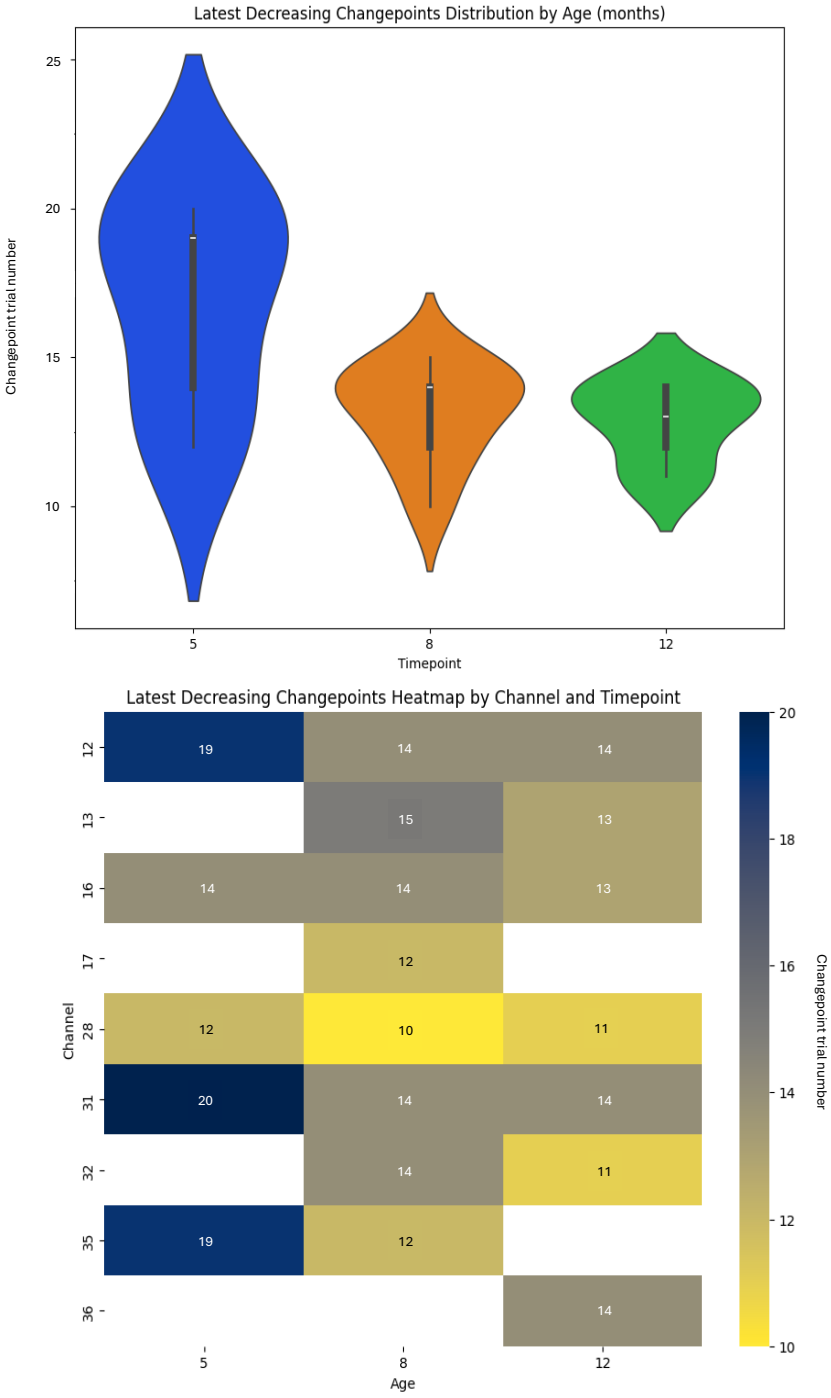

Figure 10: Violin plot and heatmap showing the age effects for latest decreasing change points across both chromophores in age-specific channels.

Distribution differences can be observed between 5- and both 8- and 12mo data, with similar distributions for the older two ages.

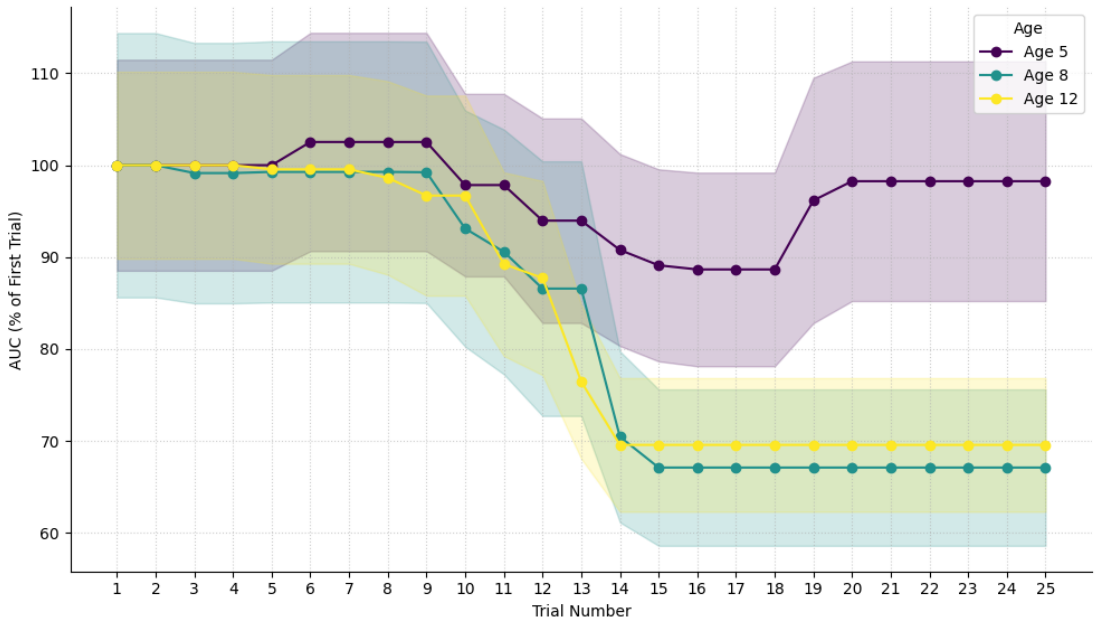

Figure 11: Mean AUC across FCPt channels, expressed as a percentage of the 1st of the 25 HaND trials.

Calculated across all channels in which FCPts were identified for each age. Greater decreases which plateau at an earlier point in the paradigm can be observed at 8- and 12mo when compared to those at 5mo. Shaded regions represent $\pm 1SEM$.

4 Discussion

This work demonstrates the novel application of FCPt detection to characterise the infant haemodynamic response, evoked by repeated stimuli and measured using fNIRS, in greater detail than has been previously possible using conventional approaches. This characterisation of the timing of structural change in the haemodynamic response also enabled assessment of longitudinal changes in evoked infant habituation responses, demonstrating a faster habituation in older infants.

4.1 Functional change point detection

FCPt detection using a fully functional approach appears to accurately capture structure in fNIRS signals corresponding to the evoked haemodynamic response. Changes in ROI channels were qualitatively representative of change in the haemodynamic response (Figure 7, and Supplementary Materials 9) whereas changes elsewhere, including frontal areas, corresponded to changes in other structural information in the processed signal such as a greater prevalence of higher frequency data components or temporally-localised changes in shape rather than magnitude (see Figure 7c, for example). Future work may explore alternative approaches to investigate different aspects of the changing haemodynamic

response, such as multivariate, spatiotemporal functional change point detection (Padilla et al., 2022).

The overlap between the channels containing FCPts detected using a fully functional method and those showing a significant haemodynamic response via TFCE (Blasi Ribera et al., 2025) was compared. Channels and cortical locations identified by fully functional FCPt detection showed consistency with those previously identified as drivers of significant activation in the same data (Blasi Ribera et al., 2025; Lloyd-Fox et al., 2019) when constraining analysis to channels where both chromophore signals reached significance. Analysing both chromophore signals is important and likely to reduce false positive results given the effect of neurally-induced haemodynamic coupling on both oxygenated and deoxygenated haemoglobin, and aligns with growing advocacy for reporting of results for both chromophores in fNIRS studies (Tachtsidis & Scholkmann, 2016; Yücel et al., 2021).

Overlap for single chromophore channels was marginal at all ages, further supporting the analysis of both chromophores to increase specificity and reduce false positives, but also demonstrating a limitation of FCPt detection, particularly when applied to data which is noisier than is typical for this approach (Huppert, 2016; Jiao et al., 2022; Yücel et al., 2021). However, it is important to re-emphasise that FCPt detection is intended to complement statistical approaches used to determine the statistical significance of haemodynamic responses. Differences in results between FCPt detection and the TFCE approach reflect the strengths and limitations of each, with FCPt detection capable of detecting changes in both shape and magnitude (Figure 7c). Supporting this, habituation in the form of decreasing change points was found in two channels in the HbO chromophore at 5mo and one channel each at 8- and 12mo where significant haemodynamic responses were not found using TFCE (see Figure 9). This possibly highlights the incorporation of important structural information during FCPt detection which may otherwise be overlooked using methods which reduce data to simpler scalar summaries (Gertheiss et al., 2024). Alternatively, these differences in identified channels may suggest that using block averages with a range defined by the limits of the structural changes is better able to capture structural information than when using predefined epochs (see Figure 7a, d and g).

4.2 Functional data analysis

Whilst FCPt detection showed significant overlap with TFCE channels, this overlap was imperfect. This emphasises the distinction already outlined between change point detection methods and approaches specifically designed to address one- and two-sample problems, which are necessary to determine spatial locations of haemodynamic responses. Functional data methods exist for this purpose using both fully functional (Olsen et al., 2023; Pini & Vantini, 2017; C. Zhang et al., 2010; J.-T. Zhang et al., 2010) and FPCA-based (Horváth & Kokoszka, 2012a; Pomann et al., 2016) approaches. These methods may offer more informative insight than the commonly-used uni- or multivariate approaches given the inherent holistic consideration of the functional observation (Gertheiss et al., 2024). This holistic consideration may prompt the application of other FDA analysis approaches such as functional analysis of variance (ANOVA) or functional linear regression to fNIRS

data for other aspects of statistical analysis during fNIRS research (Cuevas, 2014; Gertheiss et al., 2024; Horváth & Kokoszka, 2012d).

4.3 Habituation in auditory haemodynamic responses

FCPt detection was used to extract the timing and nature of significant structural changes in the evoked haemodynamic response in each channel before comparing timings across ages.

The high proportion of decreasing change points suggests a general trend of habituation, particularly when considering ROI channels from all ages taken from the previous TFCE analysis. This trend is especially notable at the two older age points (8- and 12mo) when considering the cross-age ROI channels. In the cases where there is a failure of the first- and last-detected change points in the age-specific ROI channels to reach significance, this is likely attributable to the limited number of FCPts overall, rather than to their relative distribution (see Table 2). These results, driven by the results at older ages, are consistent with the previously reported continued habituation of haemodynamic response in this Gambian cohort, even following exposure to the novel stimulus (Blasi Ribera et al., 2025; Lloyd-Fox et al., 2019). Further, this is consistent with other studies which have previously reported repetition suppression in response to repeated stimuli in developmental populations (Emberson et al., 2017; Nakano et al., 2009).

The response in 5mo infants was more ambiguous. No statistical significance was found for the proportion of decreasing changes when considering the first, last, or overall set of detected change points; this is also consistent with the univariate analysis. In part, this may be driven by some channels which exhibited early increases in haemodynamic response (see Figure 9), possibly reflecting an initial increase in responsiveness ('sensitisation') before habituation (Rankin et al., 2009). Conversely, some channels contained increasing change points towards the end of the experimental paradigm. Given that the peak of mean haemodynamic responses before and after structural changes in several of these channels is similar in magnitude, it is possible that the higher percentage of increasing FCPts reflects detection of structural change which is overlooked by analysis using summary scalar values (see Figure 7c). However, this doesn't explain why these types of detected changes seem to be unique to the 5mo data beyond the possibility that the data collected at younger ages contains more noise due to motion and/or poor scalp-optode coupling (Beaton et al., 2025). Alternatively, it is possible that the variation in the nature of structural changes at this age corresponds to the ongoing development of important auditory and speech perception processes (Cabrera & Lau, 2022; Puertollano et al., 2024); anomalous results caused by variation due to a smaller sample size; or possibly even a true novelty response, or a combination of these factors.

This ongoing development and variation due to sample size is also evident in the assessment of age effects on the timing of various types of change points: the latest decreasing structural changes in the haemodynamic response typically occur earlier at 8- and 12mo than at 5mo, where the FCPt timing distribution is more varied. Overall age effects and pairwise comparisons suggest that infants' responses and associated haemodynamic responses may attenuate more quickly to repeated spoken stimuli at older ages, driven by differences between both 8-

and 12mo data and the earlier data collected at 5mo, despite the absence of a consistent novelty detection response (see Figure 7 and Figure 10). It is also noteworthy that there appeared to be no significant difference in earlier FCPts across ages. This is interesting in and of itself, given the age-related differences in time taken to 'finish' habituating, as it suggests habituation started at similar times across all three ages. It is also informative for future researchers designing habituation studies as it provides an indication of how many trial repetitions may be required in the experimental paradigm. Most importantly, it also outlines the key advantage of using FCPt detection as a form of analysis for this paradigm and similar studies: the potential to distinguish between cohorts based on the rate of habituation.

This is evident when examining the trajectories of the decrease in responses for each age, as illustrated in Figure 11, with the decrease in AUC beginning earlier and showing a larger overall reduction at both 8- and 12mo than at 5mo. The form of the decrease in response is also notable: an initial plateau is followed by a sharp decrease, before a later, stable plateau at lower AUC values, reminiscent of sigmoidal decay. Taken to correspond to habituation, the form of this decrease contrasts other theories and findings which suggest a linear or exponential form of habituation (Thompson and Spencer, 1966; Sicard-Cras et al., 2022; Rankin et al., 2009).

4.4 Strengths and limitations

The work presented here is the first application of FCPt detection to NIR imaging data, demonstrating the utility of the novel application of an analytical approach beyond those conventionally used in the field. Beyond this, FCPt detection was sensitive to trial-wise structural changes in the haemodynamic response, capturing nuanced variation in both magnitude and shape which univariate methods may overlook. This enabled trial-by-trial analysis of the data using measures beyond summary statistics in predefined time periods (e.g. mean of values in the peak response window), which uses more of the data while preserving the temporal structure of the haemodynamic response. For example, previous analysis using TFCE was also able to identify continued habituation for a cluster including channel 16, for which FCPt findings are displayed in Figure 7; however, FCPt detection was additionally able to provide information about the timing of this habituation. FCPt detection serves as a complementary approach to conventional, univariate methods, with results yielded by each method providing different insights into the data.

A limitation of this work is the use of CUSUM-based statistics, which are biased towards detection in the middle of the observation series. This bias stems from the fact they are designed to accumulate deviations from a reference value over time; as such, near the start/end of the series, there are limited observations with which to establish a baseline or confirm a persistent shift (Otto & Breitung, 2023). This bias persists even with wild binary segmentation and is particularly important to consider when interpreting the apparent shape of the habituation trajectory. Alternative approaches altogether may improve FCPt detection at series boundaries and mitigate this bias (Horváth et al., 2022).

Persistent physiological signal components in individual data did not permit meaningful change point detection or subsequent analysis at this level. FCPt detection was therefore conducted on group data, using cross-participant mean data for each channel and trial, as this mitigated the persistent confounds to a greater extent by performing a similar function to that normally provided by averaging within conditions. Consequently, small (group-wise) sample sizes for change points at each age increased volatility in results, making them sensitive to minor variations in ROI channels, meaning caution is warranted when interpreting results concerning age-related differences. Future work may improve individual-level robustness by exploiting the greater spatial resolution afforded by HD systems, either to analyse spatially averaged individual data e.g. using parcellations (Uchitel et al., 2022) or by incorporating larger channel numbers in group analysis. It is also important to note that sample size issues may be prevalent when using smaller numbers of participants, if analysing cross-participant mean responses: the ability to mitigate effectively against physiological confounds relies on averaging across a sufficient set of cross participant data per trial, which is not guaranteed when using a smaller *N.*

A further limitation is that the approach used here assumes that the residual functions $\epsilon_n(t)$ are independent, though temporal dependence between observations likely exists due to the design of the HaND paradigm. The long-run (LR) covariance function can account for the cumulative effect of temporal dependencies by summing the autocovariances over time (Kokoszka, 2012) and can be used during FCPt detection (Horváth & Kokoszka, 2012c). However, it was found that the length ($N = 25$) of the functional datasets was not sufficient to generate reliable LR covariance estimates (see Supplementary Materials 12), limiting detection accuracy. Using the ordinary covariance function as in this work introduces potential standard error underestimation and consequent false positive inflation, though a stable yet biased estimate was preferred to a variable LR estimate. Moving forward, it may be interesting to look at the relationship between the number of trials and the stability of the LR covariance function (possibly using simulated data), since the extent to which the error introduced by using the ordinary covariance function is currently unknown.

5 Conclusion

In this work, a novel approach to identify structural changes in the evoked haemodynamic response in fNIRS data has been explored. Using infant data from a rural Gambian population recorded using a HaND paradigm, the use of functional change point detection with wild binary segmentation has been tested and demonstrated as robust. Group-level analysis from participants assessed at 5-, 8- and 12mo identified structural change corresponding to the evoked haemodynamic response in locations consistent with previous analyses of the same dataset. Further analyses indicated consistent habituation at 8- and 12mo and an age-related shift in the timing of decreasing haemodynamic responses, suggesting older infants in the first year of life habituate to auditory stimuli sooner. These findings demonstrate the utility of analysing fNIRS data using FCPt detection, with potential future research avenues including implementation of alternative change point detection methods or use with high-density imaging setups.

Data & Code Availability statement

The code used to conduct the analyses and generate the figures presented in this paper are available at https://github.com/sam-beaton/fcpt and rely on the following open source Python packages: *h5py* (Collette, 2013), *numpy* (Harris et al., 2020) , *matplotlib* (Hunter, 2007), *os* (Python Software Foundation, 2025a), *pandas* (McKinney, 2010), *pickle* (Python Software Foundation, 2025b), *scikit-fda* (Ramos-Carreño et al., 2024), *scipy* (Virtanen et al., 2020), *statistics* (Python Software Foundation, 2025c), *sys* (Python Software Foundation, 2025d).

The data used to support this study are stored in the Brain Imaging for Global Health Data Repository. The conditions of ethics approval do not allow public archiving of pseudo-anonymized study data. The data cannot be fully anonymized due to the nature of combined sources of information, such as neuroimaging, sociodemographic, geographic and health measures, making it possible to attribute data to specific individuals. Hence, the release of data would not be compliant with GDPR guidelines unless additional participant consent forms are completed (Blasi Ribera et al., 2025). Access to any data collected during or generated by the BRIGHT project is fully audited and, to ensure data security, is overseen by the data management team in the UK and The Gambia. Relevant data sharing procedures were created in consultation with stakeholders and external consultation (Begum Ali & Holman, 2023). To access the data, interested readers should use the Contact page of the BRIGHT website. Access will be granted to named individuals following ethical procedures governing the reuse of sensitive data. Specifically, requestors must pre-register their proposal and clearly explain the planned analysis to ensure that the purpose and nature of the research is consistent with that to which participating families originally consented. Additionally, requestors must complete and sign a data sharing agreement to ensure data is stored securely. Approved projects would need to adhere to the BRIGHT project's policies on Ethics, Data Sharing, Authorship and Publication.

Author Contributions (CRediT)

SB: conceptualisation, data curation, formal analysis, methodology, software, visualisation, writing – original draft, writing – review and editing; CE: funding acquisition, project administration, resources, writing – review and editing; SLF: funding acquisition, project administration, resources, writing – review and editing; EM: data curation, investigation, project administration; SMc: project administration, supervision, writing – review and editing; AB: data curation, supervision, writing – review and editing; SEM: funding acquisition, project administration, resources, supervision, writing – review and editing.

Funding

Sam Beaton, Ebrima Mbye, Samantha McCann (to July 2024) and Sophie Moore are supported by a Wellcome Trust Senior Research Fellowship (220225/Z/20/Z) held by Sophie Moore. The BRIGHT Study was funded by the Gates Foundation (OPP1127625) and core funding MC-A760-5QX00 to the International Nutrition Group by the Medical Research Council UK and the UK Department for International Development (DfID) under the MRC/DfID Concordat agreement.

Further support was provided through a UKRI Future Leaders Fellowship (MR/S018425/1) held by Sarah Lloyd-Fox.

Declaration of Competing Interests

The authors have no conflict of interest to declare.

Acknowledgements

We would like to place on record our thanks to the children, mothers and wider families who took part in this study. We also thank the BRIGHT Project study team in Keneba for their invaluable contribution to data collection and management.

References

Aston, J. A. D., & Kirch, C. (2011). *Estimation of the distribution of change-points with application to fMRI data*.

Aston, J. A. D., & Kirch, C. (2012). Evaluating stationarity via change-point alternatives with applications to fMRI data. *The Annals of Applied Statistics*, *6*(4), 1906–1948. https://doi.org/10.1214/12-AOAS565

Aue, A., Rice, G., & Sönmez, O. (2018). Detecting and dating structural breaks in functional data without dimension reduction. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, *80*(3), 509–529.

Beaton, S., Blanco, B., Bulgarelli, C., Elwell, C. E., Lloyd-Fox, S., Mbye, E., McCann, S., Ribera, A. B., & Moore, S. E. (2025). *Investigating the effect of channel pruning on functional near-infrared spectroscopy data collected from children aged 5-24 months* (p. 2025.10.06.680288). bioRxiv. https://doi.org/10.1101/2025.10.06.680288

Begum Ali, J., & Holman, R. (2023). *Parent attitudes to data sharing*. Birkbeck College, University of London. https://doi.org/10.18743/DATA.00000319

Benerradi, J., Bulgarelli, C., Mbye, E., Blanco, B., Touray, E., McCann, S., Milosavljevic, B., Moore, S. E., Elwell, C. E., Blasi, A., Lloyd-Fox, S., & Team, T. B. P. (2025). *Functional specialisation across the first five years of life: A longitudinal characterisation of social perception with fNIRS* (p. 2025.07.28.667189). bioRxiv. https://doi.org/10.1101/2025.07.28.667189

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300.

Berkes, I., Gabrys, R., Horváth, L., & Kokoszka, P. (2009). Detecting Changes in the Mean of Functional Observations. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *71*(5), 927–946. https://doi.org/10.1111/j.1467-9868.2009.00713.x

Blasi Ribera, A., Blanco Maniega, B., McCann, S., Mbye, E., Touray, E., Rozhko, M., Milosavljevic, B., Katus, L., Saidykhan, M., Ceesay, M., Fadera, T., Ghillia, G., Perapoch Amado, M., Crespo-Llado, M. M., Moore, S. E., Elwell, C. E., & Lloyd-Fox, S. (2025). Longitudinal habituation and novelty detection neural responses from infancy to early childhood in The Gambia and UK. *Developmental Cognitive Neuroscience*, *76*, 101619. https://doi.org/10.1016/j.dcn.2025.101619

Bouchon, C., Nazzi, T., & Gervain, J. (2015). Hemispheric Asymmetries in Repetition Enhancement and Suppression Effects in the Newborn Brain. *PLOS ONE*, *10*(10), e0140160. https://doi.org/10.1371/journal.pone.0140160

Bulgarelli, C., Blasi, A., McCann, S., Milosavljevic, B., Ghillia, G., Mbye, E., Touray, E., Fadera, T., Acolatse, L., Moore, S. E., Lloyd-Fox, S., Elwell, C. E., Eggebrecht, A. T., & the BRIGHT Study Team. (2024). *Growth in early infancy drives optimal brain functional connectivity which predicts cognitive flexibility in later childhood* [Preprint]. Neuroscience. https://doi.org/10.1101/2024.01.02.573930

Cabrera, L., & Lau, B. K. (2022). The development of auditory temporal processing during the first year of life. *Hearing, Balance and Communication*, *20*(3), 155–165. https://doi.org/10.1080/21695717.2022.2029092

Collette, A. (2013). *Python and HDF5*. O'Reilly Media. https://www.h5py.org/

Collins-Jones, L. H., Cooper, R. J., Bulgarelli, C., Blasi, A., Katus, L., McCann, S., Mason, L., Mbye, E., Touray, E., Ceesay, M., Moore, S. E., Lloyd-Fox, S., & Elwell, C. E. (2021). Longitudinal infant fNIRS channel-space analyses are robust to variability parameters at the group-level: An image reconstruction investigation. *NeuroImage*, *237*, 118068. https://doi.org/10.1016/j.neuroimage.2021.118068

Cooper, R., Selb, J., Gagnon, L., Phillip, D., Schytz, H., Iversen, H., Ashina, M., & Boas, D. (2012). A Systematic Comparison of Motion Artifact Correction Techniques for Functional Near-Infrared Spectroscopy. *Frontiers in Neuroscience*, *6*, 147. https://doi.org/10.3389/fnins.2012.00147

Cuevas, A. (2014). A partial overview of the theory of statistics with functional data. *Journal of Statistical Planning and Inference*, *147*, 1–23. https://doi.org/10.1016/j.jspi.2013.04.002

Delpy, D. T., Cope, M., Zee, P. van der, Arridge, S., Wray, S., & Wyatt, J. (1988). Estimation of optical pathlength through tissue from direct time of flight measurement. *Physics in Medicine and Biology*, *33*(12), 1433–1442. https://doi.org/10.1088/0031-9155/33/12/008

Emberson, L. L., Cannon, G., Palmeri, H., Richards, J. E., & Aslin, R. N. (2017). Using fNIRS to examine occipital and temporal responses to stimulus repetition in young infants: Evidence of selective frontal cortex involvement. *Developmental Cognitive Neuroscience*, *23*, 26–38. https://doi.org/10.1016/j.dcn.2016.11.002

Fryzlewicz, P. (2014). Wild binary segmentation for multiple change-point detection. *The Annals of Statistics*, *42*(6). https://doi.org/10.1214/14-AOS1245

Gertheiss, J., Rügamer, D., Liew, B. X. W., & Greven, S. (2024). Functional Data Analysis: An Introduction and Recent Developments. *Biometrical Journal*, *66*(7), e202300363. https://doi.org/10.1002/bimj.202300363

Gonzalez-Gonzalez, N. L., Medina, V., Padron, E., Domenech, E., Diaz Gomez, N. M., Armas, H., & Bartha, J. L. (2009). Fetal and Neonatal Habituation in Infants of Diabetic Mothers. *The Journal of Pediatrics*, *154*(4), 492–497. https://doi.org/10.1016/j.jpeds.2008.10.020

Harris, C. R., Millman, K. J., Van Der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., Van Kerkwijk, M. H., Brett, M., Haldane, A., Del Río, J. F., Wiebe, M., Peterson, P., … Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, *585*(7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2

Hemmati Berivanlou, N., Setarehdan, S. K., & Ahmadi Noubari, H. (2014). Evoked hemodynamic response estimation using ensemble empirical mode decomposition based adaptive algorithm applied to dual channel functional near infrared spectroscopy (fNIRS). *Journal of Neuroscience Methods*, *224*, 13–25. https://doi.org/10.1016/j.jneumeth.2013.12.007

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 65–70.

Horváth, L., & Kokoszka, P. (2012a). 5. Two sample inference for the mean and covariance functions. In L. Horváth & P. Kokoszka, *Inference for Functional*

Data with Applications (Vol. 200, pp. 65–78). Springer New York. https://doi.org/10.1007/978-1-4614-3655-3_16

Horváth, L., & Kokoszka, P. (2012b). 6. Detection of changes in the mean function. In *Inference for Functional Data with Applications* (pp. 79–104). Springer.

Horváth, L., & Kokoszka, P. (2012c). 16. Functional time series. In L. Horváth & P. Kokoszka, *Inference for Functional Data with Applications* (Vol. 200, pp. 289–341). Springer New York. https://doi.org/10.1007/978-1-4614-3655-3_16

Horváth, L., & Kokoszka, P. (2012d). *Inference for functional data with applications* (Vol. 200). Springer Science & Business Media.

Horváth, L., Rice, G., & Zhao, Y. (2022). Change point analysis of covariance functions: A weighted cumulative sum approach. *J. Multivar. Anal.*, *189*(C). https://doi.org/10.1016/j.jmva.2021.104877

Hsing, T., & Eubank, R. (2015). *Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators*. Wiley.

Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*, *9*(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

Huppert, T. J. (2016). Commentary on the statistical properties of noise and its implication on general linear models in functional near-infrared spectroscopy. *Neurophotonics*, *3*(1), 010401. https://doi.org/10.1117/1.NPh.3.1.010401

Huppert, T. J., Diamond, S. G., Franceschini, M. A., & Boas, D. A. (2009). HomER: A review of time-series analysis methods for near-infrared spectroscopy of the brain. *Applied Optics*, *48*(10), D280–D298. https://doi.org/10.1364/AO.48.00D280

Jiao, S., Chan, N.-H., & Yau, C.-Y. (2022). *Enhanced Change-Point Detection in Functional Means* (arXiv:2205.04299). arXiv. https://doi.org/10.48550/arXiv.2205.04299

Katus, L., Blasi, A., McCann, S., Mason, L., Mbye, E., Touray, E., Ceesay, M., de Haan, M., Moore, S. E., Elwell, C. E., & Lloyd-Fox, S. (2023). Longitudinal fNIRS and EEG metrics of habituation and novelty detection are correlated in 1–18-month-old infants. *NeuroImage*, *274*, 120153. https://doi.org/10.1016/j.neuroimage.2023.120153

Kiefer, J. (1959). K-Sample Analogues of the Kolmogorov-Smirnov and Cramer-V. Mises Tests. *The Annals of Mathematical Statistics*, *30*(2), 420–447. https://doi.org/10.1214/aoms/1177706261

Killick, R., Fearnhead, P., & Eckley, I. A. (2012). Optimal Detection of Changepoints With a Linear Computational Cost. *Journal of the American Statistical Association*, *107*(500), 1590–1598. https://doi.org/10.1080/01621459.2012.737745

Kokoszka, P. (2012). Dependent Functional Data. *ISRN Probability and Statistics*, *2012*, 1–30. https://doi.org/10.5402/2012/958254

Kokoszka, P., & Reimherr, M. (2017a). Chapter 10, Elements of Hilbert space theory. In *Introduction to Functional Data Analysis* (pp. 211–232). CRC Press.

Kokoszka, P., & Reimherr, M. (2017b). *Introduction to Functional Data Analysis*. CRC Press.

Lindquist, M. A., Meng Loh, J., Atlas, L. Y., & Wager, T. D. (2009). Modeling the hemodynamic response function in fMRI: Efficiency, bias and mis-modeling.

NeuroImage, *45*(1), S187–S198. https://doi.org/10.1016/j.neuroimage.2008.10.065

Lloyd-Fox, S., Blasi, A., McCann, S., Rozhko, M., Katus, L., Mason, L., Austin, T., Moore, S. E., Elwell, C. E., & The BRIGHT project team. (2019). Habituation and novelty detection fNIRS brain responses in 5- and 8-month-old infants: The Gambia and UK. *Developmental Science*, *22*(5). https://doi.org/10.1111/desc.12817

Lloyd-Fox, S., Blasi, A., Volein, A., Everdell, N., Elwell, C. E., & Johnson, M. H. (2009). Social Perception in Infancy: A Near Infrared Spectroscopy Study. *Child Development*, *80*(4), 986–999. https://doi.org/10.1111/j.1467-8624.2009.01312.x

Lloyd-Fox, S., McCann, S., Milosavljevic, B., Katus, L., Blasi, A., Bulgarelli, C., Crespo-Llado, M., Ghillia, G., Fadera, T., Mbye, E., Mason, L., Njai, F., Njie, O., Perapoch-Amado, M., Rozhko, M., Sosseh, F., Saidykhan, M., Touray, E., Moore, S. E., … The BRIGHT Project team. (2023). The Brain Imaging for Global Health (BRIGHT) Project: Longitudinal cohort study protocol. *Gates Open Research*, *7*, 126. https://doi.org/10.12688/gatesopenres.14795.1

Lloyd-Fox, S., Richards, J. E., Blasi, A., Murphy, D. G. M., Elwell, C. E., & Johnson, M. H. (2014). Coregistering functional near-infrared spectroscopy with underlying cortical areas in infants. *Neurophotonics*, *1*(2), 025006. https://doi.org/10.1117/1.NPh.1.2.025006

Luke, R., Larson, E. D., Shader, M. J., Innes-Brown, H., Yper, L. V., Lee, A. K. C., Sowman, P. F., & McAlpine, D. (2021). Analysis methods for measuring passive auditory fNIRS responses generated by a block-design paradigm. *Neurophotonics*, *8*(2), 025008. https://doi.org/10.1117/1.NPh.8.2.025008

McCann, S. (2020). *Nutrition, Growth and Neurodevelopment Across Infancy and Early Childhood in Rural Gambia.* [PhD Thesis]. King's College London.

McKinney, W. (2010). Data structures for statistical computing in Python. *Proceedings of the 9th Python in Science Conference*, 51–56. https://pandas.pydata.org/

Molavi, B., & Dumont, G. A. (2012). *Wavelet-based motion artifact removal for functional near-infrared spectroscopy*. *33*(2), 259–270. https://doi.org/10.1088/0967-3334/33/2/259

Muller, H.-G. (2009). Functional Data Analysis. *International Handbook (Encyclopedia) of Statistical Sciences*, 5.

Nakano, T., Watanabe, H., Homae, F., & Taga, G. (2009). Prefrontal Cortical Involvement in Young Infants' Analysis of Novelty. *Cerebral Cortex*, *19*(2), 455–463. https://doi.org/10.1093/cercor/bhn096

Nordt, M., Hoehl, S., & Weigelt, S. (2016). The use of repetition suppression paradigms in developmental cognitive neuroscience. *Cortex*, *80*, 61–75. https://doi.org/10.1016/j.cortex.2016.04.002

Olsen, N. L., Pini, A., & Vantini, S. (2023). *Local inference for functional data on manifold domains using permutation tests* (arXiv:2306.07738). arXiv. http://arxiv.org/abs/2306.07738

Otto, S., & Breitung, J. (2023). Backward CUSUM for Testing and Monitoring Structural Change with an Application to COVID-19 Pandemic Data. *Econometric Theory*, *39*(4), 659–692. https://doi.org/10.1017/S0266466622000159

Padilla, C. M. M., Wang, D., Zhao, Z., & Yu, Y. (2022). Change-point Detection for Sparse and Dense Functional Data in General Dimensions. *Advances in Neural Information Processing Systems Datasets and Benchmarks 2022*, 13.

Pini, A., & Vantini, S. (2017). Interval-wise testing for functional data. *Journal of Nonparametric Statistics*, *29*(2), 407–424. https://doi.org/10.1080/10485252.2017.1306627

Pomann, G.-M., Staicu, A.-M., & Ghosh, S. (2016). A Two-Sample Distribution-Free Test for Functional Data with Application to a Diffusion Tensor Imaging Study of Multiple Sclerosis. *Journal of the Royal Statistical Society Series C: Applied Statistics*, *65*(3), 395–414. https://doi.org/10.1111/rssc.12130

Puertollano, M., Ribas-Prats, T., Gorina-Careta, N., Ijjou-Kadiri, S., Arenillas-Alcón, S., Mondéjar-Segovia, A., Dolores Gómez-Roig, M., & Escera, C. (2024). Longitudinal trajectories of the neural encoding mechanisms of speech-sound features during the first year of life. *Brain and Language*, *258*, 105474. https://doi.org/10.1016/j.bandl.2024.105474

Python Software Foundation. (2021). *Python 3.9.16*. https://www.python.org/

Python Software Foundation. (2025a). *Python Standard Library: Os—Miscellaneous operating system interfaces*. https://docs.python.org/3/library/os.html

Python Software Foundation. (2025b). *Python Standard Library: Pickle—Python object serialization*. https://docs.python.org/3/library/pickle.html

Python Software Foundation. (2025c). *Python Standard Library: Statistics—Mathematical statistics functions*. https://docs.python.org/3/library/statistics.html

Python Software Foundation. (2025d). *Python Standard Library: Sys—System-specific parameters and functions*. https://docs.python.org/3/library/sys.html

Ramos-Carreño, C., Torrecilla, J. L., Carbajo Berrocal, M., Marcos Manchón, P., & Suárez, A. (2024). scikit-fda: A Python Package for Functional Data Analysis. *Journal of Statistical Software*, *109*(2), 1–37. https://doi.org/10.18637/jss.v109.i02

Ramsay, J. O., & Silverman, B. W. (2005). *Functional data analysis* (2nd ed). Springer.

Rankin, C. H., Abrams, T., Barry, R. J., Bhatnagar, S., Clayton, D. F., Colombo, J., Coppola, G., Geyer, M. A., Glanzman, D. L., Marsland, S., McSweeney, F. K., Wilson, D. A., Wu, C.-F., & Thompson, R. F. (2009). Habituation revisited: An updated and revised description of the behavioral characteristics of habituation. *Neurobiology of Learning and Memory*, *92*(2), 135–138. https://doi.org/10.1016/j.nlm.2008.09.012

Rice, G., & Zhang, C. (2022). Consistency of binary segmentation for multiple change-point estimation with functional data. *Statistics & Probability Letters*, *180*, 109228. https://doi.org/10.1016/j.spl.2021.109228

Scholkmann, F., Spichtig, S., Muehlemann, T., & Wolf, M. (2010). How to detect and reduce movement artifacts in near-infrared imaging using moving standard deviation and spline interpolation. *Physiological Measurement*, *31*(5), 649–662. https://doi.org/10.1088/0967-3334/31/5/004

Scholkmann, F., & Wolf, M. (2013). General equation for the differential pathlength factor of the frontal human head depending on wavelength and age. *Journal*

of Biomedical Optics, *18*(10), 105004. https://doi.org/10.1117/1.JBO.18.10.105004

Seabold, S., & Perktold, J. (2010). statsmodels: Econometric and statistical modeling with python. *9th Python in Science Conference*.

Sicard-Cras, I., Rioualen, S., Pellae, E., Misery, L., Sizun, J., & Roué, J.-M. (2022). A review of the characteristics, mechanisms and clinical significance of habituation in foetuses and newborn infants. *Acta Paediatrica*, *111*(2), 245–258. https://doi.org/10.1111/apa.16115

Smith, S., & Nichols, T. (2009). Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage*, *44*(1), 83–98. https://doi.org/10.1016/j.neuroimage.2008.03.061

Tachtsidis, I., & Scholkmann, F. (2016). False positives and false negatives in functional near-infrared spectroscopy: Issues, challenges, and the way forward. *Neurophotonics*, *3*(3), 031405. https://doi.org/10.1117/1.NPh.3.3.031405

Tukey, J. W. (1949). Comparing individual means in the analysis of variance. *Biometrics*, 99–114.

Uchitel, J., Blanco, B., Vidal-Rosas, E., Collins-Jones, L., & Cooper, R. J. (2022). Reliability and similarity of resting state functional connectivity networks imaged using wearable, high-density diffuse optical tomography in the home setting. *NeuroImage*, *263*, 119663. https://doi.org/10.1016/j.neuroimage.2022.119663

Ullah, S., & Finch, C. F. (2013). Applications of functional data analysis: A systematic review. *BMC Medical Research Methodology*, *13*(1), 43. https://doi.org/10.1186/1471-2288-13-43

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, *17*, 261–272. https://doi.org/10.1038/s41592-019-0686-2

Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society*, *54*(3), 426–482.

Wang, J.-L., Chiou, J.-M., & Müller, H.-G. (2016). Functional Data Analysis. *Annual Review of Statistics and Its Application*, *3*(1), 257–295. https://doi.org/10.1146/annurev-statistics-041715-033624

Yücel, M. A., Lühmann, A. v., Scholkmann, F., Gervain, J., Dan, I., Ayaz, H., Boas, D., Cooper, R. J., Culver, J., Elwell, C. E., Eggebrecht, A., Franceschini, M. A., Grova, C., Homae, F., Lesage, F., Obrig, H., Tachtsidis, I., Tak, S., Tong, Y., … Wolf, M. (2021). Best practices for fNIRS publications. *Neurophotonics*, *8*(01). https://doi.org/10.1117/1.NPh.8.1.012101

Zhang, C., Peng, H., & Zhang, J.-T. (2010). Two Samples Tests for Functional Data. *Communications in Statistics - Theory and Methods*, *39*(4), 559–578. https://doi.org/10.1080/03610920902755839

Zhang, J.-T., Liang, X., & Xiao, S. (2010). On the Two-Sample Behrens-Fisher Problem for Functional Data. *Journal of Statistical Theory and Practice*, *4*(4), 571–587. https://doi.org/10.1080/15598608.2010.10412005

Appendices

Appendix A Mathematical detail of FCPt detection

In each channel, age-grouped data was obtained by averaging across participants for each trial, yielding a single group-averaged function per trial in each channel, represented by Equation 1. The error term in Equation 1 is typically modelled with zero mean and covariance function: $c(t,s) = E\{\epsilon_n(t)\epsilon_n(s)\}$. This describes the linear association between functional observations ϵ_n at two different time points, t and s, during fNIRS recording. The covariance function is a key component of FCPt detection. In practice, $c(t,s)$ is estimated using the sample covariance function:

$$\hat{c}(t,s) = \frac{1}{N}\sum_{1\leq n\leq N}\{Y_n(t) - \bar{Y}_N(t)\}\{Y_n(s) - \bar{Y}_N(s)\}$$

Equation 3

where $\bar{Y}_N(t)$ is the sample mean haemodynamic function:

$$\bar{Y}_N(t) = \frac{1}{N}\sum_{1\leq n\leq N} Y_n(t)$$

Equation 4

Potential change in the mean haemodynamic response function at all trials n_c was tested using the FF FCPt detection method (Aue et al., 2018), during which structural breaks in the mean function are examined using the covariance structure. This exploits the way deviations in the cumulative mean behaviour of functional data are expected to behave under the null hypothesis (Equation 2).

For a particular trial n in the HaND paradigm ($1 \leq n \leq N$) it is possible to adapt the sample mean in Equation 4 to give the sample mean hemodynamic response function of the first n trials (i), and the final $(N-n)$ trials (ii):

$$\text{(i)} \quad \bar{Y}_n(t) = \frac{1}{n}\sum_{i=1}^{n} Y_i(t) \qquad \text{(ii)} \quad \bar{Y}_{N-n}(t) = \frac{1}{(N-n)}\sum_{i=n+1}^{N} Y_i(t)$$

This can be used to give the weighted cumulative mean difference up to trial n:

$$P_n(t) = \frac{n(N-n)}{N}[\bar{Y}_n(t) - \bar{Y}_{N-n}(t)]$$

Equation 5

If the mean is constant, $P_n(t)$ will be small for all $1 \leq n \leq N$; if a structural change occurs it will be large. The parabolic weight function $\frac{n(N-n)}{N}$ stabilises random variability which can occur when n is large ($\approx N$) or small (Horváth & Kokoszka, 2012b, p. 84). Substituting the full definitions of the two partial sums $\bar{Y}_n(t)$ and $\bar{Y}_{N-n}(t)$ into Equation 4 and rearranging gives

$$P_n(t) = \sum_{i=1}^{n} Y_i(t) - \frac{n}{N}\sum_{i=1}^{N} Y_i(t),$$

Equation 6

a cumulative sum (CUSUM) statistic, so named because it involves the sums of the weighted cumulative mean differences in Equation 5. $P_n(t)$ represents the

deviation of the cumulative mean up to trial n from the overall mean across all N trials. This motivates the test statistic T_N^F:

$$T_N^F(n) = \max_{1 \le n \le N} \frac{1}{N} \int_T (P_n(t))^2 \, dt \; = \max_{1 \le n \le N} \frac{1}{N} \int_T \left(\sum_{i=1}^{n} Y_i(t) - \frac{n}{N} \sum_{i=1}^{N} Y_i(t) \right)^2 dt$$

Equation 7

T_N^F represents the trial number n which produces the largest absolute difference between the cumulative sum of Y_n and all its antecedents, and the mean function across all functional observations weighted by $\frac{n}{N}$, evaluated across the entire (time) domain. Normalising by N ensures T_N^F is well scaled.

Appendix B FCPt significance testing with Brownian Bridges

To assess the size of $T_N^F(n)$ values for the data and the significance of the associated change points, the distribution of the test statistic under H_0 is required, which relies on the estimated covariance, $\hat{c}(t,s)$.

An eigendecomposition of the estimated covariance operator $\int \hat{c}(t,s)\, ds$ provides its estimated eigenfunctions $\hat{v}_k(t)$ and eigenvalues $\hat{\lambda}_k$ (Horváth & Kokoszka, 2012c, p. 301; Kokoszka & Reimherr, 2017b, pp. 40–42; Muller, 2009):

$$\int \hat{c}(t,s)\, \hat{v}_k(s) ds = \hat{\lambda}_k \hat{v}_k(t)$$

Equation 8

Eigenfunctions characterise the dominant modes/shapes of variation in functional data. The kth eigenvalue represents the amount of variation corresponding to the kth eigenfunction. The importance of the eigenvalues $\hat{\lambda}_k$ for T_N^F is seen in its convergence in distribution, D, under H_0:

$$T_N^F(n) \xrightarrow{D} \sup_{1 \le n \le N} \sum_{k}^{\infty} \hat{\lambda}_k B_k^2(n) \, , \qquad (N \to \infty)$$

(Aue et al., 2018; Mansuy & Yor, 2008)

where $B_k(n)$ is a Brownian Bridge on the same values of n – a stochastic process $B_k(n) = W(n) - nW(n)$ such that $B_n(0) = B_n(N) = 0$ and $W_N(n)$ is a standard Brownian motion (Figure 5). During Monte Carlo simulation, for each n^*, the maximum value for the mth Brownian Bridge, B_m, was given by

$$B_m = \max_{0 \le n \le N_b} \sum_{k=1}^{d} \hat{\lambda}_k B_k^2(n) = \max_{0 \le n \le N_b} \sum_{k=1}^{d} \hat{\lambda}_k \big(B_k(n)\big)^2$$

where N_b is the length of the interval being assessed at each stage of the recursive wild binary segmentation process. The $B_m$model random fluctuations under the null hypothesis; $B_k^2(n) = \big(B_k(n)\big)^2$ quantifies the variability contributed by each component; multiplying by $\hat{\lambda}_k$ ensures more dominant modes of variability contribute more to the sum.