Theta-Beta Ratio in Attention Deficit Hyperactivity Disorder: A Multiverse Analysis

AUTHORS:

Dawid Strzelczyk[1,2,3*], Andrea Vetsch, Nicolas Langer[1,2,3]

AFFILIATIONS:

1: Methods of Plasticity Research, Department of Psychology, University of Zurich, Zurich, Switzerland

2: University Research Priority Program (URPP) Dynamics of Healthy Aging, Zurich, Switzerland

3: Neuroscience Center Zurich (ZNZ), Zurich, Switzerland

E-mails: n.langer@psychologie.uzh.ch, dawid.strzelczyk@psychologie.uzh.ch

CORRESPONDING AUTHOR (*):

Dawid Strzelczyk

Methods of Plasticity Research

Department of Psychology

University of Zurich

Binzmühlestrasse 14

CH-8050, Zürich, Switzerland

E-mail: dawid.strzelczyk@uzh.ch

Tel.: +41 75 421 34 00

Abstract

Attention Deficit Hyperactivity Disorder (ADHD) affects 5-7% of children worldwide. Current diagnostics rely on subjective assessments, highlighting the need for objective neurobiological markers. The theta/beta ratio (TBR) recorded with electroencephalography has been proposed as a potential ADHD biomarker, with affected children showing increased TBR compared to healthy controls. However, recent findings have raised concerns about TBR's diagnostic validity, suggesting it may only be reliable under specific methodological conditions. To address these inconsistencies, we employed multiverse analysis to systematically investigate how researcher degrees of freedom affect TBR estimates across the full spectrum of reasonable methodological choices. Using two large, independent samples (Healthy Brain Network: N=1,499; validation sample: N=381), we varied TBR computation methods to evaluate the consistency of group differences across the analytical space. Our results demonstrate that TBR differences are highly contingent on analytical choices. Critically, we observed virtually no main effects of diagnosis, indicating no reliable differences between healthy controls, ADHD-inattentive, and ADHD-combined subtypes. However, significant interactions emerged, particularly with age and individual alpha frequency (IAF), specifically when TBR was computed using frequency bands relative to IAF and from 1/f-uncorrected power or from the aperiodic signal itself. These findings suggest that previously reported TBR effects may be confounded by variations in the 1/f-slope and IAF rather than reflecting genuine differences in oscillatory power. Importantly, these interaction patterns replicated across both independent datasets and held for categorical and dimensional ADHD definitions. Taken together, these findings indicate that TBR does not provide a reliable standalone biomarker for ADHD.

1. Introduction

Attention Deficit Hyperactivity Disorder (ADHD) affects approximately 5-7% of children worldwide and represents one of the most prevalent neurodevelopmental disorders, with significant implications for public health and educational systems [1–3]. Core symptoms such as inattention, hyperactivity, and impulsivity can significantly impair academic performance, disrupt social relationships, and increase the risk of comorbid mental health conditions, with these challenges frequently persisting into adulthood [2,4,5]. Current diagnostic practices rely heavily on subjective clinical assessments, behavioral rating scales, and structured interviews, creating a pressing need for objective neurobiological markers that could enhance diagnostic precision and reduce the substantial variability in clinical decision-making [6].

Numerous studies have shown that children with ADHD often exhibit increased slow-wave activity, particularly in the theta range (4-8 Hz), alongside decreased beta activity (>14 Hz) [5,7–11]. In the context of ADHD, increased theta has been associated with fatigue and drowsiness, whereas lower beta has been linked to decreased mental activity and concentration [2,12]. To quantify this imbalance, Lubar (1991) proposed the theta/beta ratio (TBR) as a potential biomarker of ADHD, suggesting it could serve as a reliable tool for both diagnosis and treatment monitoring. This approach gained traction with evidence from Monastra et al. (1999), who reported high sensitivity (86%) and specificity (98%) for classifying ADHD in a multi-center study [13,14]. These promising results contributed to regulatory recognition, culminating in the U.S. FDA's approval of TBR-based assessment as an adjunctive tool for ADHD diagnosis [15–18]. Consequently, multiple commercial EEG-based diagnostic systems emerged, promising to enhance the objectivity and reliability of ADHD evaluations in clinical practice.

However, despite early promising results, more recent studies with heterogeneous samples have raised concerns about the validity and robustness of TBR as a diagnostic marker for ADHD [2,5,12,15,19,20]. As pointed out by Arns et al. (2013, 2016), the foundational studies by Monastra and colleagues (1999, 2001) contained several methodological limitations that may compromise their generalizability. Both studies employed unbalanced designs with ADHD groups approximately 5 times larger than controls, potentially inflating statistical power [21,22]. While manual artifact rejection was performed, no specific thresholds or rejection rates were reported across groups, a critical omission given that movement artifacts significantly affect the aperiodic (1/f) component of EEG and can substantially alter TBR

values [23,24]. The risk of biased TBR estimates is particularly high, because children with ADHD tend to move more than controls [25]. Furthermore, the ADHD samples excluded comorbid conditions such as anxiety, depression, or learning disorders, despite these being present in 60-80% of real-world ADHD cases, severely limiting ecological validity [1,3,22]. Importantly, the authors conflated statistically significant group differences with diagnostic utility, yet demonstrating higher average TBR in ADHD groups does not establish the measure's capacity for reliable individual classification.

Meta-analyses by Arns et al. (2013, 2016) have revealed significant heterogeneity across studies and a marked decline in effect sizes over time, indicating that more recent studies tend to find smaller or non-significant differences in TBR between individuals with ADHD and healthy controls. This growing inconsistency has led to the view that TBR may only be a reliable marker of ADHD under specific methodological conditions. Factors such as preprocessing pipelines, resting state conditions (eyes open vs. eyes closed), reference schemes, power computation methods, frequency band definitions, regions of interest, and participant characteristics (e.g., medication status, comorbidities) can all influence the resulting TBR values [19,26]. Supporting this, van Dijk et al. (2020) processed the same dataset with five alternative pipelines that varied aperiodic (1/f) handling, spectral estimation (i.e., Welch, Multi-taper, Wavelet), and the TBR formula. Although absolute TBR values differed significantly, the estimates were highly correlated and none reliably distinguished ADHD from controls. However, five pipelines cover only a small fraction of plausible specifications, and other reasonable choices could similarly shift the estimates.

The multitude of choices researchers face during data preprocessing and analysis is referred to as researcher degrees of freedom (RDF) [27]. When these decisions remain undisclosed or are made post-hoc based on desired outcomes, they can inflate false-positive rates and undermine the reproducibility of scientific findings [27,28]. To address these concerns, the scientific community has increasingly embraced multiverse analysis as a robust framework for evaluating the stability of research findings across the full spectrum of reasonable analytical choices [28–32]. Rather than presenting results from a single, potentially arbitrary analytical pathway, multiverse analysis systematically implements all plausible combinations of methodological decisions, providing a comprehensive assessment of effect robustness and identifying specific conditions under which effects emerge or disappear [30–32].

Therefore, in the present study, we employed a multiverse analysis to systematically address the methodological uncertainties surrounding the computation of TBR in children and adolescents with ADHD and healthy controls. Using two large, independent samples: the

Healthy Brain Network (N = 1499) and a validation sample (N = 381), we systematically varied preprocessing steps and TBR computation methods to evaluate the consistency of observed group differences across the full reasonable analytical space. Our choice of analytic specifications was informed by a systematic review of prior work on TBR in ADHD (see Methods 2.6). This comprehensive framework enabled us to determine whether TBR differences between children and adolescents with ADHD and healthy controls represent robust neurobiological phenomena or are effects of specific methodological choices, thereby providing crucial evidence for the field's ongoing evaluation of EEG-based ADHD biomarkers.

2. Methods

To ensure transparency and reproducibility, all analysis code, EEG features, and demographic data used in the statistical models are accessible on OSF via https://osf.io/u5yxv.

2. 1. Participants

To assess the generalizability of our findings, we analyzed data from two independent samples: a main sample from the Healthy Brain Network (HBN) project [33,34] and a validation sample taken from a large-scale multicenter clinical study [35–37]. As the multiverse analysis results were highly similar, and preprocessing and data analysis were identical across samples, we report demographic details, EEG acquisition information and all results for the validation sample in the Supplement (Supplementary Material 1.2).

The HBN project by the Child Mind Institute is an ongoing initiative that aims to generate a freely available biobank of multimodal datasets of children, adolescents and young adults aged 5-22 years. For the current study, we analyzed data available up to Release 11 (as of 11.23.2022), identifying a total of 1,499 participants with resting state EEG and either no psychiatric diagnosis (healthy controls) or a diagnosis of ADHD. The sample consisted of 271 healthy controls (mean age = 9.81 years, SD = 3.52, range = 5.02-21.67), 584 participants with ADHD-Combined Type (mean age = 9.19 years, SD = 2.86, range = 5.04-21.72), and 559 with ADHD-Inattentive Type (mean age = 10.88 years, SD = 3.07, range = 5.43-21.48). An additional 85 participants with ADHD-Hyperactive/Impulsive Type (mean age = 7.65 years, SD = 2.10, range = 5.01-13.67) were excluded from further analyses due to insufficient sample size across key subgroup factors used in the multiverse analysis (e.g., gender, comorbidity and medication status, and their interactions), which were central to the multiverse analysis framework. For detailed demographic information please refer to Table 1.

ADHD diagnoses in the HBN study were determined by licensed clinicians based on a structured clinical protocol [33]. The primary diagnostic instrument was the Kiddie Schedule for Affective Disorders and Schizophrenia (K-SADS; [38]), a semi-structured interview conducted with parents and, for participants aged 11 and older, also with the child. Additional clinical judgment was informed by behavioral observations during testing sessions, family history of psychiatric conditions, prior diagnoses, and responses on standardized parent-, child-, and

teacher-report questionnaires (e.g., CBCL [39], SWAN [40], Conners [41]). The decision also took into account whether the participant had previously received mental health support. For detailed information, please refer to the HBN Diagnostic Process (Version 3): (https://fcon_1000.projects.nitrc.org/indi/cmi_healthy_brain_network). The IQ was measured using the Wechsler Intelligence Scale for Children (WISC-V), Wechsler Abbreviated Scale of Intelligence (WASI), or Wechsler Adult Intelligence Scale (WAIS) according to the age of the participant. 280 ADHD participants were regularly treated with psychiatric medication in daily life (i.e., ADHD stimulants (Methylphenidate-based: Ritalin, Concerta; Amphetamine-based: Adderall, Vyvanse), Antidepressants (Prozac, Zoloft) and Atypical antipsychotics (Abilify)). Prior to participation, legal guardians or participants of legal age provided written informed consent. Study approval was given by the Chesapeake Institutional Review Board.

2. 2. EEG Acquisition

The HBN EEG data were recorded using a 128-channel Hydrogel net system (Electrical Geodesics Inc.) at a sampling rate of either 250 Hz or 500 Hz. To ensure consistency across recordings, all EEG data were downsampled to 250 Hz prior to further processing. The reference electrode was placed at Cz (vertex of the head), and electrode impedances were kept below 40 kΩ, lower than EGI's standard recommendation of 50 (Net Station Acquisition Technical Manual).

Participants from the HBN sample were seated comfortably in a sound-shielded room, 70 cm away from a 17-inch CRT monitor (Sony Trinitron Multiscan G220; display: 330 × 240 mm; resolution: 800 × 600 pixels; refresh rate: 100 Hz). A chinrest was used to minimize head movements. Participants were informed that EEG would be recorded while they alternated between eyes-open (EO) and eyes-closed (EC) resting-state conditions. Task instructions were presented on the screen, and a research assistant provided clarification via intercom from an adjacent control room. Compliance was monitored through a live video feed. The resting state task consisted of five EO-EC cycles. Each EO block lasted 20 s (totaling 1 min 40 s), and each EC block lasted 40 s (totaling 3 min 20 s), resulting in a total recording time of 5 minutes. Pre-recorded verbal cues were delivered via loudspeakers to instruct participants when to open or close their eyes. During EO blocks, participants were asked to maintain gaze on a central fixation cross. The EO-EC alternation aimed to prevent fatigue and maintain vigilance, with longer EC periods to capitalize on lower artifact rates in the EC condition.

2. 3. EEG Preprocessing

EEG data from both the main and validation datasets were preprocessed using an identical pipeline to ensure methodological consistency across samples. Data preprocessing and EEG feature extraction were conducted using MATLAB 2023b (The MathWorks, Inc., Natick, MA, USA) and RStudio 4.4.1 (R Core Team). The data were preprocessed in Automagic 3.1, a MATLAB based toolbox for automated, reliable and objective preprocessing of EEG datasets [42]. In a first step, the bad channels were detected using the PREP pipeline [43]. A channel was defined as bad based on 1) extreme amplitudes (z-score cutoff for robust channel deviation of more than 5), 2) lack of correlation (at least 0.4) with other channels with a window size of 1s to compute the correlation, 3) lack of predictability by other channels (channel is bad if the prediction falls below the absolute correlation of 0.75 in a fraction of 0.4 windows of a duration of 5s), 4) unusual high frequency noise using a z-score cutoff for SNR of 5. These channels were removed from the original EEG data. The data was filtered using a high-pass filter with 0.5 Hz cutoff using the EEGLAB function pop_eegfiltnew [44]. Line noise was removed using a ZapLine method with a passband edge of 60 Hz for HBN sample and 50 Hz for validation sample [45], removing 7 power line components. Next, independent component analysis (ICA) was performed. However, as the ICA is biased towards high amplitude and low frequency noise (i.e., sweating), the data was temporarily filtered with a high-pass filter of 2 Hz in order to improve the ICA decomposition. Using the pre-trained classifier IClabel [46] each independent component with a probability rating >0.8 of being an artifact such as muscle activity, heart artifacts, eye activity, line noise and channel noise were removed from the data. The remaining components were back-projected on the original 0.5 Hz high-pass filtered data. In the next step, the channels identified as bad were interpolated using the spherical interpolation method. Finally, the quality of the data was automatically and objectively assessed in Automagic, thus increasing research reproducibility by having objective measures for data quality. Using a selection of 4 quality measures, the data was classified into three categories: “good”, “ok” or “bad”. Data was classified as “bad”, if 1) the proportion of high-amplitude data points (>30μV) in the signal is greater than 0.3, or 2) more than 50% of time points show a variance greater than 50μV across all channels, or 3) 30% of the channels show variance greater than 40μV, or 4) the ratio of bad channels is greater than 0.3. For the further analysis only the datasets with “good” and “ok” ratings were used. In the HBN sample, classification was performed at the subject level (1,020 “good”, 430 “ok”, 45 “bad”; 45 subjects excluded) because EO and EC blocks were recorded as alternating segments within a single continuous file per participant. In the validation dataset, classification was performed at the file level (521 “good”, 230 “ok”, 12 “bad”; 12 files excluded) because EO and EC were recorded and stored as separate files.

Subsequently, in the HBN sample, which was recorded using a 128-channel Hydrogel net system, 23 channels were excluded from further analysis: 10 EOG channels and 13 channels located on the chin and neck, as they capture minimal brain activity and are typically contaminated with muscle artifacts [34,47,48].

Next, the data from both samples was re-referenced either to average reference or linked mastoid (see multiverse analysis tree) into 2 s long segments. The first and the last segment of each EO / EC block was discarded to exclude motor activity related to opening and closing the eyes and auditory activity due to the prompt from the speakers. Remaining EEG segments were inspected using an amplitude threshold of 90 µV. In the HBN sample, 197 subjects (12.6%) were excluded from further analysis because more than 60% of their trials exceeded this threshold. Among the remaining subjects, an average of 18% of trials were excluded based on this criterion. In the validation sample, 2 subjects (0.5%) were excluded for the same reason, with an average of 14% of trials removed among the remaining subjects.

2. 5. Feature Extraction

The features were extracted from EO and EC conditions, using both average and linked mastoid references. For each condition and reference combination, we computed three types of features: relative power, aperiodic-adjusted power, and the aperiodic component. Spectral analysis was performed using the Fast Fourier Transform (FFT) over the 1-40 Hz range after applying a single Hanning window, implemented in Fieldtrip [49]. First, to obtain relative power, total power spectra were computed using cfg.output = ‘pow’. These power spectra were then normalized by dividing each frequency bin by the mean power across all bins, yielding relative power estimates. Next, to compute aperiodic-adjusted power, the power spectra were decomposed into periodic and aperiodic components using the SpecParam algorithm [24], implemented in FieldTrip (i.e., cfg.output = 'fooof'). SpecParam was applied over the 1-40 Hz frequency range to minimize the risk of overfitting low-frequency noise as narrowband peaks. The algorithm settings were as follows: peak width limits set to [1, 8], maximum number of peaks of 6, minimum peak height of 0, border threshold of 5, peak threshold of 2 standard deviations above the mean, and the aperiodic mode set to ‘fixed’. The aperiodic component was reconstructed based on its fitted parameters and subtracted from the total power spectrum, resulting in an aperiodic-adjusted (1/f-corrected) power spectrum. To ensure data quality, we applied a model fit threshold of $R^2 = 0.85$, calculated as the mean R^2 across all subjects minus 2.5 SD, following established procedures from the seminal specParm publication [24,50]. A total of 69 subjects from the HBN

sample were excluded from further analyses due to poor model fit. For the remaining participants, the mean goodness-of-fit (R^2) of the aperiodic decomposition was 0.97 ± 0.02 (M ± SD). In the validation sample, 40 subjects were excluded due to poor model fit, with remaining subjects showing R^2 of 0.97 ± 0.03 (M ± SD). Summarized, we obtained 6 datasets per participant: EC and EO data, each with relative power (i. e., uncorrected power), aperiodic-adjusted power (i. e., 1/f corrected), and the isolated aperiodic component (i. e., aperiodic signal only).

Subsequently, we proceeded with the computation of the individual alpha peak frequency (IAF) as the shift of the IAF during brain maturation might introduce a bias when power is averaged within a fixed-frequency window [51,52]. Specifically, it has been shown that the IAF increases over children's development and the rate of increase can vary among individuals [52]. To account for this, we estimated IAF from aperiodic-adjusted power spectra obtained during EC recordings over posterior electrodes (HBN: E62 (Pz), E75 (Oz), E70 (O1), E83 (O2), E72, E71, E76; Validation Sample: P7/T5, P3, O1, Pz, O2, P4, P8/T6). IAF was defined as the frequency of maximum power within a pre-specified range of 7 to 14 Hz, consistent with prior work [53–55]. If the peak was located outside of the preselected frequency range or at the border, no alpha peak was extracted for that subject, and the corresponding data was excluded from all further analyses. In the HBN sample, IAF could not be identified in 13 subjects. For the remaining participants, the mean IAF (M ± SD) was 9.74 ± 1.08 Hz in the ADHD-Combined group, 9.96 ± 1.00 Hz in the ADHD-Inattentive group, and 9.88 ± 1.02 Hz in the group without a diagnosis. In the validation sample, IAF could not be determined in 10 subjects. For the remaining participants, the mean IAF was 9.58 ± 1.08 Hz in the ADHD-Combined group, 9.60 ± 1.11 Hz in the ADHD-Inattentive group and 9.82 ± 1.06 Hz in the healthy control group.

Subsequently, theta and beta power were computed for each participant, condition, and spectral estimate using two approaches: canonical frequency bands and individualized frequency bands. Canonical bands were defined as 4-8 Hz for theta and 13-30 Hz for beta [56]. Individualized bands were centered on the IAF, defined as theta = IAF-6 Hz to IAF-4 Hz, and beta = IAF+2 Hz to 30 Hz, consistent with previous research [55,56]. Finally, TBR was calculated by dividing theta power by beta power [7]. TBR values were extracted from multiple scalp regions of interest (ROIs), selected based on prior literature [2,5,7,12,13,15,20,57,58]. These ROIs are illustrated in Figure 1.

2. 6. Multiverse Analyses

Multiverse analysis was conducted in RStudio (R version 4.4.1) using the multiverse package (version 0.6.1; Sarma et al., 2021). The selection of analytical specifications was guided by a systematic review of prior research of TBR in ADHD [2,5,7,8,12,13,15,18,20,57–97]. Following analytical specifications were investigated: resting state condition (2 levels: EO or EC), reference scheme (2 levels: average or linked mastoids), frequency band definition (2 levels: canonical frequency range [theta: 4-8 Hz; beta: 13-30 Hz] or individualized frequency bands based on IAF [theta: IAF-6 Hz to IAF-4 Hz; beta: IAF+2 Hz to 30 Hz]), handling of aperiodic signal (3 levels: uncorrected, aperiodic-adjusted, or aperiodic signal only), region of interest (6 levels: frontal (Fz), left frontal (Fp1, F3, F7), right frontal (Fp2, F4, F8), midline (Fz, Cz, Pz), central (Cz), extended central (C3, Cz, C4)), inclusion of comorbid diagnoses (2 levels: ADHD participants with comorbid diagnoses included or excluded), inclusion of subjects with medication (2 levels: ADHD participants with medication included or excluded), group contrast (2 levels: ADHD-Inattentive vs. Healthy Controls or ADHD-Combined vs. Healthy Controls), regression model (8 levels: group only; group * age; group * gender; group *IAF; group * age * gender; group * age * IAF; group * IAF * gender; group * age * gender * IAF). We use Wilkinson notation [98], where * indicates inclusion of both main effects and their interaction. Importantly, we focused exclusively on regression terms that included the factor group, allowing us to assess how the association between diagnostic status and TBR was moderated by age, gender, and IAF across analytical specifications. In summary, this resulted in 576 distinct model specifications (i.e., universes) per group contrast. Each universe included 8 regression terms, yielding a total of 4,608 (i.e., 576 x 8) possible effects for each group contrast (i.e. HC vs. ADHD-Inattentive and HC vs. ADHD-Combined contrasts). A full overview of the analysis tree with analytical specification and variable labeling scheme is provided in Figure 1.

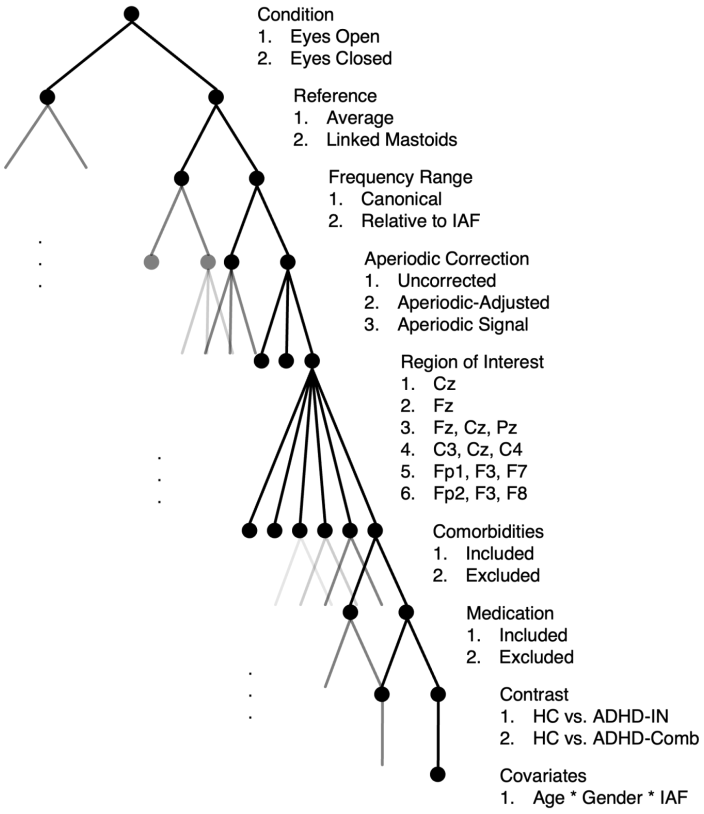


Figure 1. Final analysis tree illustrating all included analytical specifications. Each node represents a distinct analytical parameter in the multiverse analysis. Note that due to space limitations, not all parameter nodes are displayed in this visualization. Information on comorbid diagnoses and medication status was available only in the HBN sample, and these factors were therefore investigated exclusively in this dataset. HC = Healthy Control, IAF = Individual Alpha Frequency.

As an additional robustness check, we performed a bootstrap analysis to evaluate whether unequal group sizes influenced the multiverse results. For each specification, we identified the group with the fewest participants (e.g., ADHD-Inattentive, male, unmedicated, without comorbidities) and resampled all groups to match this minimum size. The multiverse analysis was then repeated on these balanced subsamples, and the procedure was iterated 1,000 times with independent random draws. Full details and aggregated results of the bootstrap analysis are provided in the Supplement (Supplementary Material 1. 1. 4).

To complement the categorical analyses based on the diagnostic group (i.e., HC, ADHD-Inattentive, ADHD-Combined), we conducted an identical multiverse analysis using the SWAN total scores [40] as a dimensional predictor of ADHD symptom severity in the HBN sample only (healthy controls as well as ADHD patients), as SWAN scores were not available for the validation sample. This dimensional approach tested the robustness of findings across the full spectrum of ADHD symptomatology, avoiding potential information

loss from diagnostic thresholds while increasing statistical power and capturing subclinical variability. The identical set of analytical specifications and model structures were applied (i.e., 576 universes), substituting the categorical Group factor with continuous SWAN total scores and their interactions with age, gender, and IAF. Full results are provided in Supplementary Material 1. 1. 5.

2. 6. 1. Specification curves plots

To visualize the impact of analytical specifications on the estimated effects, we constructed specification curves plots [99,100]. The specification curve displays the estimated effect size (regression coefficient) for a given regression term across all 576 analytical specifications per group contrast. This visualization provides a comprehensive overview of the consistency, direction, and statistical reliability of results across the full analytical multiverse. Separate specification curves were generated for each regression term within the HC vs. ADHD-Inattentive and HC vs. ADHD-Combined contrasts. If effect sizes remain consistent across a wide range of analytical decisions, it suggests that the findings are robust and not driven by arbitrary choices. In contrast, if results vary substantially across specifications, this indicates sensitivity to particular analytical decisions and warrants cautious interpretation. Notably, even consistently null results are informative, as they imply that no meaningful effect emerges regardless of the analytical path taken.

2. 6. 2. Proportions plots and possibility space

To facilitate the interpretation of the results and to highlight which analytical choices most consistently produced significant effects, we summarised all 576 universes per regression term as proportion of significantly positive (Supplementary Figure 2) and negative (Supplementary Figure 3) effects. For every regression term within each ADHD subtype we computed the proportions of significantly positive and significantly negative estimates to the total number of universes.

To assess whether the proportion of significant results within each subset of the multiverse analysis exceeded what would be expected by chance (i.e., 5%), we conducted one-sided binomial tests (i.e., binom.test()). The null hypothesis assumed that only 5% of models would yield significant results by random chance (H_0: $p = 0.05$), and the alternative hypothesis tested whether the observed proportion was greater than expected (H_1: $p > 0.05$). We computed the minimum proportion of significant results required for the binomial test to

return $p < 0.05$ for subsets with 288 (e.g., EO and EC), 192 (e.g., uncorrected power, aperiodic-adjusted power, aperiodic signal), and 96 (e.g., 6 regions of interest) universes. The thresholds obtained from the binomial tests were approximately 7.6% (for condition, reference, frequency range, comorbidities and medication), 8.3% (for aperiodic correction), and 10.4% (for region of interest). These values were used to interpret whether observed proportions within specific analytical branches could be considered statistically robust.

When interpreting the results of a multiverse analysis, it is important to distinguish between probabilistic and possibilistic uncertainty (i.e., incertitude) [29,30]. Probabilistic uncertainty reflects the variability within each individual model specification, arising from sampling error and limited data. It allows for inferences about the likelihood of parameter estimates and is typically quantified through p-values, confidence intervals, or consonance curves [29,30,101,102]. In contrast, possibilistic uncertainty captures the range of results that emerge across all defensible analytic decisions within the multiverse [29,30]. This type of uncertainty does not imply probability. Instead, each result is treated as a valid possibility, regardless of how frequently it occurs. It is therefore essential to consider both forms of uncertainty, as they provide complementary information: probabilistic measures quantify the reliability of estimates within a given specification, while possibilistic measures reveal the robustness of conclusions across alternative specifications. To capture the range of possibilities, we developed a ShinyApp in RStudio that summarizes estimates for each regression term, visualizing the full distribution of results and highlighting significant ones through color-coding. An online version of the app is available via shinyapps.io (HBN: https://dstrze.shinyapps.io/HBN-sample/ and https://dstrze.shinyapps.io/hbn-sample-swan/; Validation Sample: https://dstrze.shinyapps.io/validation-sample/), and the corresponding R code is openly shared on the OSF under https://osf.io/u5yxv. Users can interactively filter model specifications to explore which analytic choices most frequently yield significant outcomes. Key plots summarizing the most important results are presented in the main text of the manuscript (Figures 5 & 7).

3. Results

3. 1. Healthy Brain Network Sample

3. 1. 1. Demographics

Table 1 and Figure 2 summarize the demographic characteristics of the Healthy Brain Network sample. Gender was compared across all three groups (i.e., HC, ADHD-Combined, ADHD-Inattentive) using chi-square tests of independence. The chi-square test revealed a significant association between those groups and gender (χ^2 = 47.47, p = 4.9e-11). Comorbid conditions and medication use were analyzed only within the ADHD sample (ADHD-Combined vs ADHD-Inattentive), because these variables were not applicable (NA) for controls. Both showed significant group differences: comorbid conditions (χ^2 = 7.33, p = 0.007) and medication use (χ^2 = 40.08, p = 2.43e-10), indicating that the distributions of those those variables differed across both ADHD groups. Next, we conducted one-way ANOVAs to investigate differences in continuous variables. The ANOVAs revealed significant group effects for age (F = 43.23, p = 5.9e-19), all SWAN subscales (inattention (F = 288.35, p = 2.6e-105), hyperactivity (F = 369.75, p = 3.2e-129), and total scores (F = 394.42, p = 4.1e-136)), as well as WISC (F = 27.11, p = 3e-12), and IAF (F = 5.23, p = 0.006), indicating that the HC, ADHD-Combined and ADHD-Inattentive differed significantly across all measured variables.

Table 1. Demographics information of the Healthy Brain Network sample.

	HC	***Combined***	***Inattentive***	***X²***	***p-value***
N	271	584	559		
Male	148	454	382		
Female	123	130	177	47.47	4.9e-11***
Comorbidities					
Yes	NA	465	407		
No	271	119	152	7.33	0.007**
Medication					
Yes	NA	176	86		

No	268		408		473		40.08	2.43e-10***

	Mean	***SD***	***Mean***	***SD***	***Mean***	***SD***	***F-value***	***p-value***
Age (years)	9.81	3.52	9.19	2.86	10.87	3.07	43.23	5.9e-19***
SWAN								
IN	-0.32	1.11	1.24	0.85	1.13	0.85	288.35	2.6e-105***
HY	-0.44	1.11	1.23	0.76	0.27	0.85	369.75	3.2e-129***
Total	-0.38	1.02	1.24	0.68	0.70	0.70	394.42	4.1e-136***
WISC FSIQ	106.92	14.63	98.66	15.97	97.91	15.43	27.11	3e-12***
IAF	9.88	1.02	9.74	1.08	9.96	1.00	5.23	0.006**

Note. HC = Healthy control. SD = Standard deviation. IN = Inattentive. IAF = Individual Alpha Frequency.

$*p < 0.05$. $**p < 0.01$. $***p < 0.001$

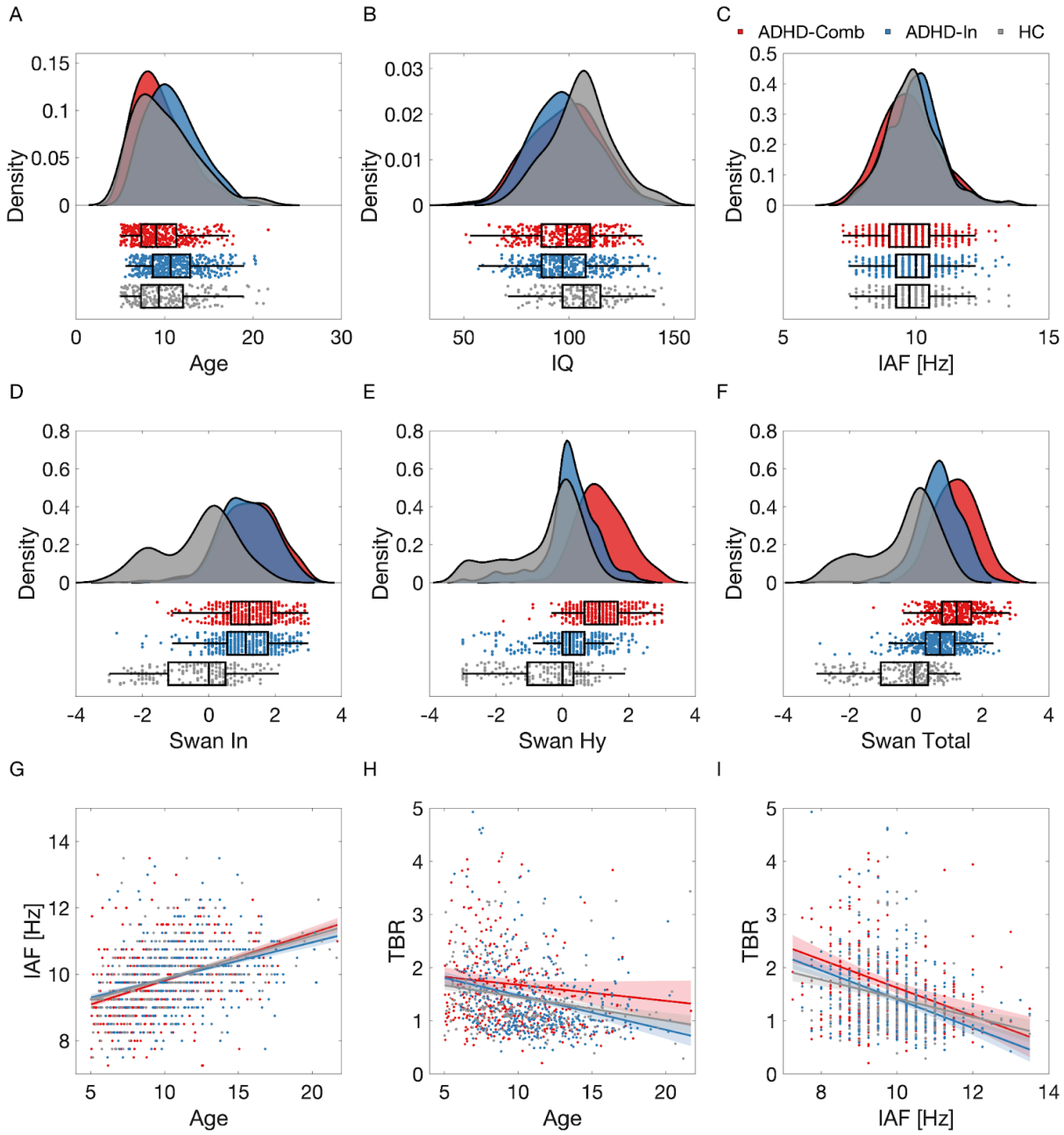


Figure 2. Overview of sample characteristics and variable distributions. Density plots show the distribution of (A) age, (B) IQ, (C) IAF, (D-F) and SWAN scores (Inattention, Hyperactivity, and Total) across the three groups. (G-I) Scatter plots display the relationship between (G) IAF and age, (H) between TBR and age, (I) and TBR and IAF across all participants. The regression line is shown with standard error bands. Note. IAF = Individual Alpha Frequency. TBR = Theta Beta Ratio. HC = Healthy Control.

3. 1. 2. EEG Features

To visualize the neurophysiological data, we plotted the scalp topographies of theta, beta and TBR, power spectra and aperiodic signal for HC, ADHD-Combined, and

ADHD-Inattentive groups (Figure 3). The figure displays aperiodic-adjusted power during EO condition, computed using a canonical frequency range (theta: 4-8 Hz; beta: 13-30 Hz). For completeness, the corresponding 1/f-uncorrected power spectra and topographies are provided in the Supplementary Materials (1. 1. 1).

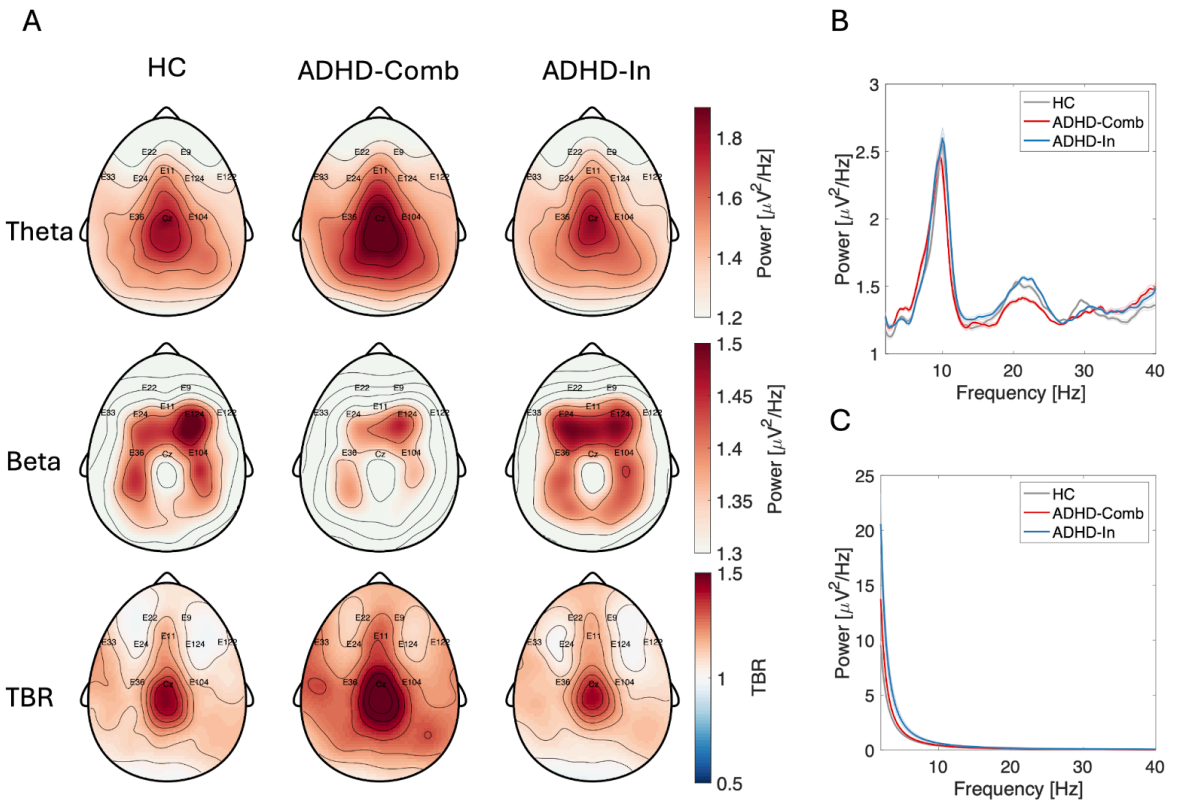


Figure 3. Neurophysiological data from the Healthy Brain Network sample. (A) Scalp topographies, (B) power spectra and (C) aperiodic signal for HC, ADHD-Combined, and ADHD-Inattentive groups. The figure displays aperiodic-adjusted power during EO condition, computed using a canonical frequency range (theta: 4-8 Hz; beta: 13-30 Hz) and TBR. Electrodes highlighted on the topographies correspond to the six regions of interest derived from literature used across different branches of multiverse analysis. The power spectra and aperiodic signal were computed by averaging across all electrodes within each respective region of interest. Note. TBR = theta-beta ratio. HC = healthy controls. Comb = Combined. In = Inattentive.

3. 1. 3. Multiverse Analysis Results

To examine the robustness of results across all analytical specifications, we conducted a multiverse analysis using the following regression model:

$$TBR \sim Group * Age * Gender * IAF$$

Here, Group was a factor with three levels: HC (reference), ADHD-Inattentive, and ADHD-Combined. Gender was a factor with 2 levels: male (i.e., reference) and female. Accordingly, the intercept represents healthy control males. The multiverse analysis also computed contrasts between HC and ADHD-Inattentive, as well as between HC and ADHD-Combined types. Table 2 provides a summary of the results across all multiverse specifications. In the context of the multiverse, positive universes refer to analytical specifications that yielded positive beta estimates, meaning that TBR was higher in the ADHD group than in HC. Conversely, negative universes reflect negative beta estimates, indicating lower TBR in the ADHD group relative to HC.

Table 2. Results of the multiverse analysis in the HBN sample showing the proportion of significant effects across 576 universes.

	Significantly Positive	***Significantly Negative***	***Not Significant***	***Total***
ADHD[IN]	0 (0.00 %)	0 (0.00 %)	576 (100.00 %)	576
ADHD[IN] * Age	0 (0.00 %)	9 (1.56 %)	567 (98.44 %)	576
ADHD[IN] * Gender[Female]	13 (2.26%)	1 (0.17%)	562 (97.57 %)	576
ADHD[IN] * IAF	15 (2.60 %)	76 (13.19 %)	485 (84.20 %)	576
ADHD[IN] * Age * Gender[Female]	0 (0.00 %)	4 (0.69 %)	572 (99.31 %)	576
ADHD[IN] * Age * IAF	62 (10.76 %)	2 (0.35 %)	512 (88.89 %)	576
ADHD[IN] * Gender[Female] * IAF	40 (6.94 %)	5 (0.87 %)	531 (92.19 %)	576
ADHD[IN] * Age * Gender[Female] * IAF	5 (0.87 %)	10 (1.73 %)	561 (97.40 %)	576
ADHD[Comb]	11 (1.91 %)	0 (0.00 %)	565 (98.09 %)	576
ADHD[Comb] * Age	64 (11.11 %)	0 (0.00 %)	512 (88.89 %)	576
ADHD[Comb] * Gender[Female]	61 (10.59 %)	3 (0.52 %)	512 (88.89 %)	576

ADHD[Comb] * IAF	4 (0.69 %)	62 (10.76 %)	510 (88.54 %)	576
ADHD[Comb] * Age * Gender[Female]	42 (7.29 %)	25 (4.34 %)	509 (88.36 %)	576
ADHD[Comb] * Age * IAF	43 (7.46 %)	1 (0.17 %)	532 (92.36 %)	576
ADHD[Comb] * Gender[Female] * IAF	10 (1.73 %)	67 (11.63 %)	499 (86.63 %)	576
ADHD[Comb] * Age * Gender[Female] * IAF	9 (1.56 %)	85 (14.76 %)	482 (83.68 %)	576

Note. Positive universe = higher TBR in ADHD compared to HC. Negative universe = lower TBR in ADHD compared to HC. IN = Inattentive. Comb = Combined. IAF = Individual alpha frequency.

3. 1. 3. 1. Healthy controls vs. ADHD-Inattentive comparison

The multiverse analysis revealed 0 significantly positive (i.e., higher TBR in ADHD-Inattentive compared to HC) and 0 significantly negative (i.e., lower TBR in ADHD-Inattentive compared to HC) universes out of 576 for the main effect of ADHD-Inattentive, indicating that TBR did not differ between ADHD-IN males compared to healthy control males in any specification (Figure 4).

Figure 4. Specification curve representing all universes for the healthy control ADHD-Inattentive contrast. The top panel shows regression estimates for each specification, sorted by regression estimate, with 95% confidence intervals. Statistically significant positive estimates are shown in green, negative estimates in red, and non-significant estimates in gray. The bottom panel indicates which analytical choices were associated with significantly positive (i.e., higher TBR for ADHD compared to HC) or negative (i.e., lower TBR for ADHD compared to HC) effects across specifications. Note. CI = confidence interval. f = female. EC = eyes closed. EO = eyes open. AVG = average reference. LM = linked mastoid reference. canonical = canonical frequency range. relative to IAF = bandwidth relative to individual alpha frequency.

Notably, the most frequently observed effects in the ADHD-Inattentive vs. healthy control comparisons involved the IAF, specifically the interactions ADHD-Inattentive * IAF, ADHD-Inattentive * Age * IAF, and ADHD-Inattentive * Gender * IAF. As illustrated in the Figure 4, the negative ADHD-Inattentive * IAF effect emerged almost exclusively when using frequency bands relative to IAF and was most frequently observed for the aperiodic signal and when children with comorbidities were included. To further explore the commonalities across universes, we used the ShinyApp to interactively filter the analytical space. When restricting the multiverse to the models with frequency bands relative to IAF, aperiodic signal as spectral measure and including subjects with comorbidities, all 48 remaining universes (i.e., across 2 resting-state conditions × 2 references × 6 ROIs × 2 medication options)

yielded negative estimates, with 46 of them reaching statistical significance (Figure 5 A). In contrast, the positive ADHD-Inattentive * Age * IAF effect was almost exclusively found when TBR was computed using canonical frequency bands and aperiodic-adjusted power, and in analyses that included participants with comorbid diagnoses (37 significant universes out of 48 possibilities; Figure 5 B). Finally, the positive ADHD-Inattentive * Gender * IAF effect was observed exclusively when using frequency bands relative to IAF, primarily in the aperiodic signal, and again in analyses that included children with comorbid diagnoses (28 significant universes out of 48 possibilities; Figure 5 C).

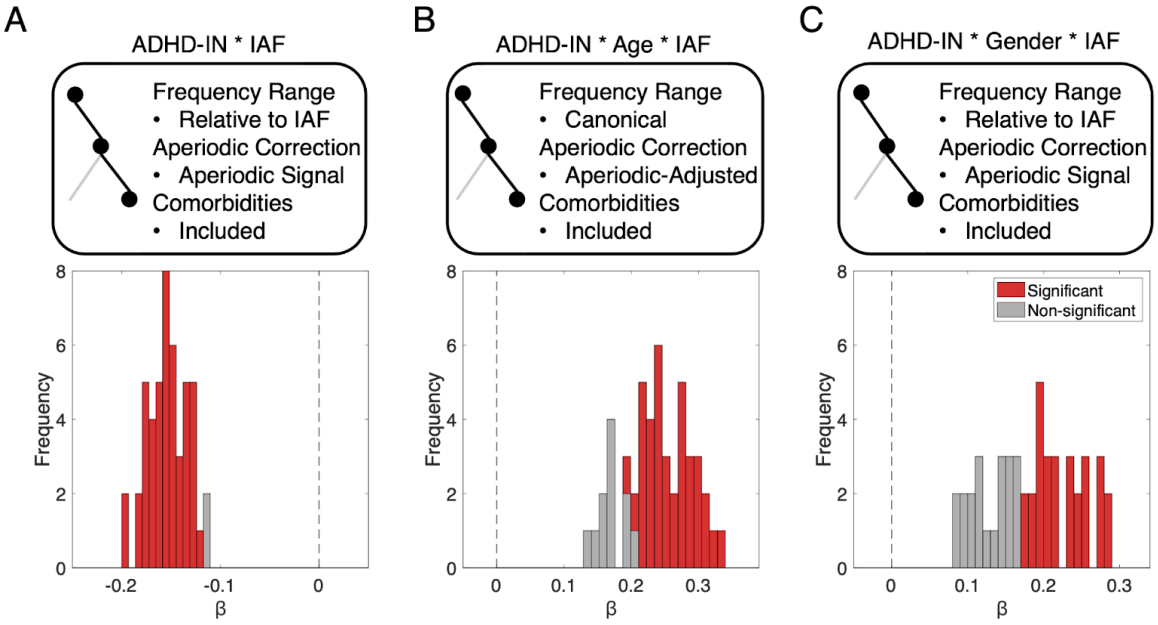


Figure 5. Possibility space of regression estimates from the subset of analytical choices yielding the highest number of significant results. (A) ADHD-Inattentive * IAF interaction: Significant effects (46 significant out of 48 possibilities) were predominantly observed when using frequency bands relative to IAF, the aperiodic signal, and including children with comorbid diagnoses. (B) ADHD-Inattentive * Age * IAF interaction: This positive effect emerged almost exclusively when TBR was computed using canonical frequency bands, aperiodic-adjusted power, and analyses included participants with comorbidities (37 out of 48 significant). (C) ADHD-Inattentive * Gender * IAF interaction: The effect was found exclusively under frequency bands relative to IAF, primarily in the aperiodic signal, and when children with comorbid diagnoses were included (28 out of 48 significant). Note. The significant estimates are highlighted in red and non-significant in grey.

3. 1. 3. 2. Healthy controls vs. ADHD-Combined comparison

Next, we compared the ADHD-Combined group to healthy controls across all analytical specifications. The multiverse analysis revealed 11 significantly positive (i.e., higher TBR in ADHD-Combined compared to HC) and 0 significantly negative (i.e., lower TBR in

ADHD-Combined compared to HC) universes out of 576 for the main effect of ADHD-Combined, indicating that TBR was higher in ADHD-Combined males compared to healthy control males in 11 specifications (Figure 6).

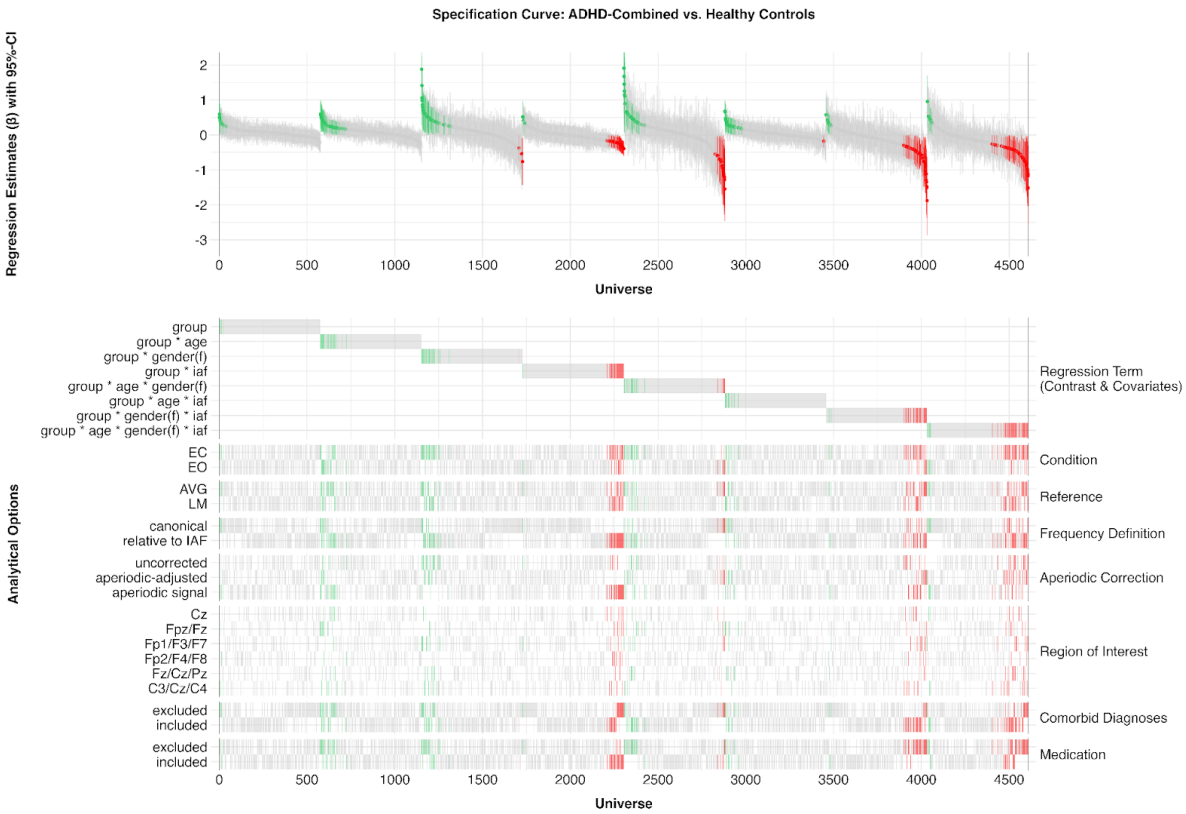


Figure 6. Specification curve representing all universes for the healthy control ADHD-Combined contrast. The top panel shows regression estimates for each specification, sorted by regression estimate, with 95% confidence intervals. Statistically significant positive estimates are shown in green, negative estimates in red, and non-significant estimates in gray. The bottom panel indicates which analytical choices were associated with significantly positive (i.e., higher TBR for ADHD compared to HC) or negative (i.e., lower TBR for ADHD compared to HC) effects across specifications. Note. CI = confidence interval. f = female. EC = eyes closed. EO = eyes open. AVG = average reference. LM = linked mastoid reference. canonical = canonical frequency range. relative to IAF = bandwidth relative to individual alpha frequency.

Notably, the most frequently observed effects in the ADHD-Combined vs. healthy control comparisons involved interactions with Age, Gender, and IAF. As illustrated in Figure 6, the positive ADHD-Combined * Age effect emerged mostly when using frequency bands relative to IAF and was most frequently observed in the aperiodic signal (24 significant universes out of 96 possibilities; Figure 7 A). Similarly, the negative ADHD-Combined * IAF effect occurred almost exclusively when using frequency bands relative to IAF, was predominantly observed in the aperiodic signal, and emerged mainly in analyses that included children on medication

(47 significant universes out of 48 possibilities; Figure 7 B). The positive ADHD-Combined * Gender effect appeared almost exclusively in the EC condition, primarily when using frequency bands relative to IAF, and was most often found in 1/f uncorrected power (27 significant universes out of 48 possibilities; Figure 7 C). The positive ADHD-Combined * Age * IAF effect showed no clear pattern across specifications and was relatively evenly distributed. Finally, the three interactions including gender (i.e., positive ADHD-Combined * Age * Gender, ADHD-Combined * Gender * IAF and ADHD-Combined * Age * Gender * IAF) appeared most frequently in the EC condition, almost exclusively when children with comorbid diagnoses were included, but those on medication were excluded (26 significant universes out of 72 possibilities, 35 significant universes out of out of 72 possibilities and 31 significant universes out of out of 72 possibilities, respectively; Figure 7 D-F).

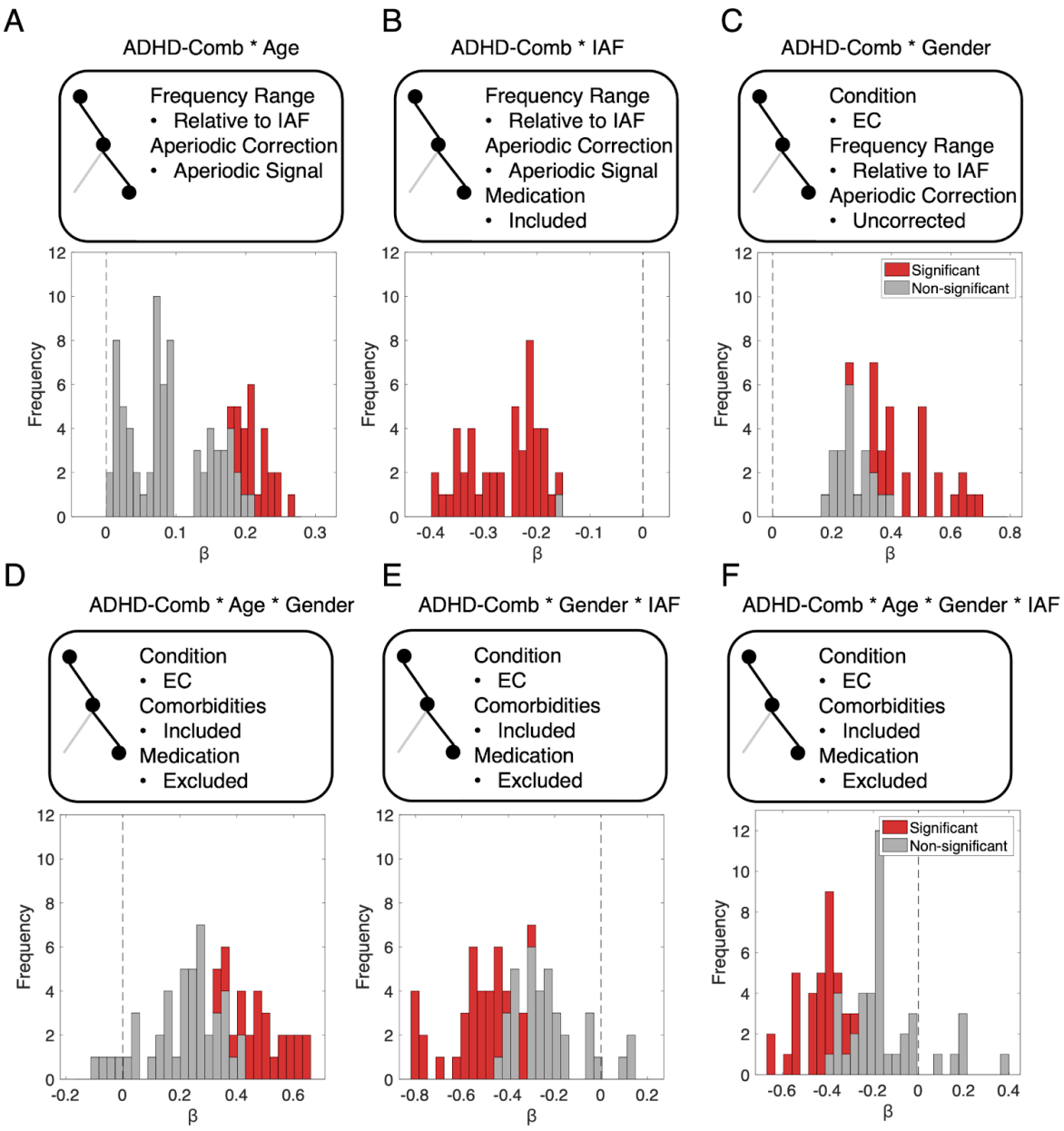


Figure 7. Possibility space of regression estimates from the subset of analytical choices yielding the highest number of significant results in the ADHD-Combined vs. healthy control comparisons. (A) ADHD-Combined * Age interaction: Significant effects (24 out of 96) were primarily observed when using frequency bands relative to IAF and the aperiodic signal. (B) ADHD-Combined * IAF interaction: A negative effect emerged almost exclusively under frequency bands relative to IAF, predominantly in the aperiodic signal, and in analyses excluding children on medication (47 out of 48 significant). (C) ADHD-Combined * Gender interaction: This positive effect appeared almost exclusively in the EC condition, with frequency bands relative to IAF and 1/f-uncorrected power (27 out of 48 significant). (D) ADHD-Combined * Age * Gender, (E) ADHD-Combined * Gender * IAF and (F) ADHD-Combined * Age * Gender * IAF interactions: These effects appeared mainly in the EC condition, almost exclusively when comorbidities were included and medication excluded (26, 35 and 31 out of 72 significant, respectively). Note. Significant estimates are highlighted in red and non-significant estimates are shown in grey.

3. 1. 4. Robustness analyses

To assess whether our multiverse findings were influenced by unequal group sizes across predictors, we conducted a bootstrap analysis in which all groups were resampled to the size of the smallest subgroup within each specification. For example, we identified the group with the lowest number of participants (e.g., ADHD-Inattentive, female, medication and comorbidities excluded) and randomly resampled all groups to this minimum size. This procedure was repeated 1,000 times with independent random draws. The results closely mirrored the main multiverse findings: the same interaction patterns involving IAF, age, and gender consistently emerged, and the frequency and direction of significant effects were relatively stable across iterations, as illustrated by low standard deviations. These results indicate that the observed patterns are robust and not a bias introduced by group size imbalance (see Supplementary Material 1. 1. 4).

To examine whether our results depended on categorical diagnosis definitions, we repeated the multiverse analysis using dimensional SWAN scores as continuous predictors. The dimensional multiverse analysis yielded results that were highly similar to the categorical findings (Supplementary Material 1. 1. 5). The main effect of SWAN showed no significantly positive universes (i.e., higher TBR in ADHD-Combined compared to HC) and only 3.65% significantly negative universes (i.e., lower TBR in ADHD-Combined compared to HC) across all specifications. Consistent with the categorical approach, the most frequent significant interactions involved age, gender, and IAF: SWAN * Gender interactions were significantly positive in 11.28% of specifications, SWAN * IAF interactions yielded significantly negative effects in 23.26% of specifications, SWAN * Gender * IAF showed significantly positive effects in 13.71% of specifications, and the four-way SWAN * Age * Gender * IAF interaction demonstrated significantly negative effects in 11.63% of cases. The convergence between categorical and dimensional approaches validates the robustness of the observed TBR effects in ADHD, demonstrating that these associations are consistent regardless of whether ADHD is operationalized as discrete diagnostic categories or as a continuous symptom dimension. Detailed analyses are presented in Supplementary Material 1. 1. 5.

4. Discussion

The present study examined the robustness of theta-beta ratio (TBR) differences between ADHD subtypes and healthy controls using a multiverse analysis across two independent samples. By systematically varying reasonable analytical parameters, including resting-state condition, frequency band definitions, aperiodic correction, reference scheme, region of interest, and inclusion/exclusion of comorbidities and medication, we quantified both the likelihood of detecting ADHD-TBR associations and the specific analytic conditions under which they arise. This approach directly addresses long-standing inconsistencies in EEG biomarker research and reveals why previous findings on TBR in ADHD have been so heterogeneous. Our results show that TBR differences are highly contingent on analytic choices. Robust effects did not emerge as simple main effects of diagnosis but rather as higher-order interactions with individual alpha frequency (IAF), age, and gender, particularly in universes without 1/f correction and those using IAF-relative frequency bands. Importantly, these interaction patterns were strikingly similar across the two independent datasets and held for both categorical and dimensional definitions of ADHD. Below, we first discuss the role of covariates in shaping TBR, then place these findings in a developmental context, and finally consider their clinical implications.

<u>No Reliable Main Effects of Diagnosis on TBR</u>

Across the 576 universes in the HBN sample and the 144 universes in the validation sample, we found no consistent TBR differences between healthy controls and either ADHD subtypes. For the HC-ADHD-Inattentive comparison, positive universes (indicating higher TBR in ADHD than in HC) and negative universes (indicating lower TBR in ADHD compared to HC) were entirely absent in the HBN sample (0% positive, 0% negative) and similarly rare in the validation sample (1.39% positive, 7.64% negative). The same pattern emerged for ADHD-Combined, with only minimal proportions of positive or negative universes (HBN: 1.91% positive, 0% negative; validation: 0.69% positive, 1.39% negative). Importantly, the results from the categorical approach were corroborated by the dimensional analyses using SWAN scores, indicating convergence across methodological frameworks (Supplementary Material 1. 1. 5). These null results align with a growing body of work challenging the diagnostic value of TBR [2,5,15,19,57]. Earlier reports of large group differences have increasingly been attributed to methodological limitations, sample imbalances, and analytic flexibility, while more recent studies have repeatedly failed to identify reliable ADHD-related TBR effects. Our multiverse results thus converge with this broader literature, providing further

evidence that TBR lacks the reliability and discriminative validity required for clinical utility. Taken together, these findings demonstrate that, regardless of whether ADHD is operationalized as discrete diagnostic categories or as a continuous symptom dimension, TBR alone shows insufficient diagnostic reliability for differentiating healthy controls from individuals with ADHD.

The role of IAF and 1/f on TBR differences

Although we found no consistent main effects of ADHD diagnosis on TBR, our multiverse analyses revealed recurring interaction patterns involving IAF, age, and gender. These interactions were highly dependent on analytic decisions, most notably, the choice between canonical versus IAF-relative frequency bands and whether oscillatory and aperiodic components were separated. Across both samples, significant effects occurred most frequently in models using IAF-relative bands combined with either the aperiodic slope or uncorrected 1/f signal. In contrast, when canonical frequency bands or aperiodic-adjusted power were used, these effects largely disappeared. These results suggest that apparent TBR differences may reflect properties of the aperiodic background signal interacting with individual variability in IAF rather than true oscillatory theta or beta activity.

Two mechanisms are likely to contribute to these effects. First, subtle differences in aperiodic slope disproportionately affect low frequencies, where the 1/f structure inflates amplitude and can mimic elevated theta power. Second, differences in IAF shift the frequency bands such that individuals with slower IAF have theta bands defined at lower frequencies (e.g., individual with slow IAF = 8 Hz has theta band defined at 2-4 Hz), which are more strongly influenced by the aperiodic signal. It artificially elevates TBR, whereas faster IAF is associated with lower TBR estimates. Importantly, both mechanisms can operate simultaneously and amplify each other. Even if group differences in IAF or aperiodic slope alone are insufficient to distinguish healthy controls from ADHD subtypes, their combined influence can produce statistically significant interaction effects. For example, when fixed canonical bands were used, IAF correlates only modestly with 1/f-uncorrected ($r = -0.17$, $p = 1.28e-7$) or aperiodic-adjusted power ($r = -0.18$, $p = 1.65e-9$). However, when frequency bands are defined relative to IAF, the correlation between IAF and 1/f-uncorrected power becomes markedly stronger ($r = -0.70$, $p = 2e-308$), and even reverses direction for aperiodic-adjusted power ($r = 0.11$, $p = 1.47e-4$), demonstrating how the interaction between subtle changes in 1/f-slope and IAF might greatly affect TBR estimates. Furthermore, our data show overall group differences in IAF (Table 1), although post-hoc comparisons indicate no significant differences between healthy controls and either ADHD-Combined ($t(655) =$

1.55, p = 0.122) or ADHD-Inattentive (t(691) = -1.09, p = 0.275) subgroups. The only statistically significant difference was a higher IAF in ADHD-Inattentive compared to ADHD-Combined (t(892) = -3.21, p = 0.001). Thus, while neither IAF nor aperiodic slope significantly differentiates HC from ADHD subtypes, their interaction magnified by analytic choices such as IAF-relative band definitions can produce group-by-IAF interactions in TBR.

An additional explanation previously proposed in the literature is that individuals with slow IAF may appear to have elevated theta power because the lower alpha range overlaps with canonical theta frequencies. In such cases, alpha power may leak into the theta band when fixed frequency ranges are used [2,53]. This interpretation is consistent with findings showing that increased TBR in boys with ADHD was replicated only when canonical frequency bands were applied, but disappeared once IAF-relative frequency ranges were used [82]. However, our results do not support the idea that low IAF produces artificially elevated theta power through alpha leakage. We observed no such effect in either the 1/f-uncorrected power or in the aperiodic-adjusted power analyses, when canonical frequency bands were used. Instead, the pattern of results suggests that IAF-related variation in TBR is driven primarily by differences in aperiodic spectral slope, as discussed above, rather than by overlap between alpha and theta frequencies.

IAF also has developmental and clinical relevance that further contextualizes these findings. With maturation, IAF typically increases, shifting theta bands closer toward the canonical 4-8 Hz range [52]. As a result, theta power derived from 1/f-uncorrected power tends to decrease with age, since the influence of the aperiodic component diminishes at higher frequencies. These dynamics may account for the recurring Age * IAF interactions observed in our multiverse analyses, as both oscillatory and aperiodic contributions vary systematically across the lifespan. These developmental considerations highlight the importance of explicitly modeling the aperiodic component when studying neural oscillations over lifespan. Oscillatory and aperiodic activity represent distinct neurophysiological phenomena, yet they can easily contaminate each other if not separated. The functional meaning of the aperiodic component remains debated: it may reflect neural processes related to excitatory-inhibitory balance [24,103], but it may also partly capture measurement artifacts without direct biological relevance [23,104]. Importantly, recent evidence suggests that 1/f intercept and slopes can be systematically affected by various artifacts (i.e., heart, eye, muscle). If one group contributes systematically more artifacts to the EEG signal, for example, due to increased movement in participants with ADHD [25], this can artificially steepen the estimated 1/f slope and elevate low-frequency power. Such differences in artifact susceptibility make it difficult to determine whether observed TBR effects reflect genuine neurophysiological variation or shifts in

spectral composition caused by noise. Future work should therefore apply methods that separate oscillatory from aperiodic activity across development and clinical groups, while ideally also measuring movement artifacts (e.g., via EMG), in order to clarify whether apparent TBR differences reflect true neural alterations or shifts in spectral composition and artifact susceptibility.

While TBR appears to lack diagnostic reliability, emerging evidence suggests that other EEG features may be more informative for clinical decision-making. In particular, IAF has been proposed as a marker for predicting treatment response, with IAF-based stratification improving remission rates when selecting between methylphenidate and neurofeedback interventions [2,105,106]. Lower frontal IAF has also been reported in adolescent non-responders [57], potentially reflecting a disruption or delay in the normative developmental increase of IAF. However, our findings indicate that IAF alone is insufficient. Accurate interpretation of spectral biomarkers must also account for individual differences in the aperiodic 1/f slope, which strongly influence power estimates and their apparent associations with behavioral or clinical measures. Together, our results indicate that TBR is not a reliable biomarker of ADHD. Instead, TBR effects are strongly shaped by analytic decisions that alter how oscillatory and aperiodic components are represented. Such analytic dependence may help explain the long-standing inconsistency in TBR findings across the ADHD literature, as unaccounted differences in IAF and aperiodic slope can systematically bias power estimates, particularly when frequency bands are defined relative to IAF or when aperiodic activity is not adequately separated from oscillatory power.

Conclusion

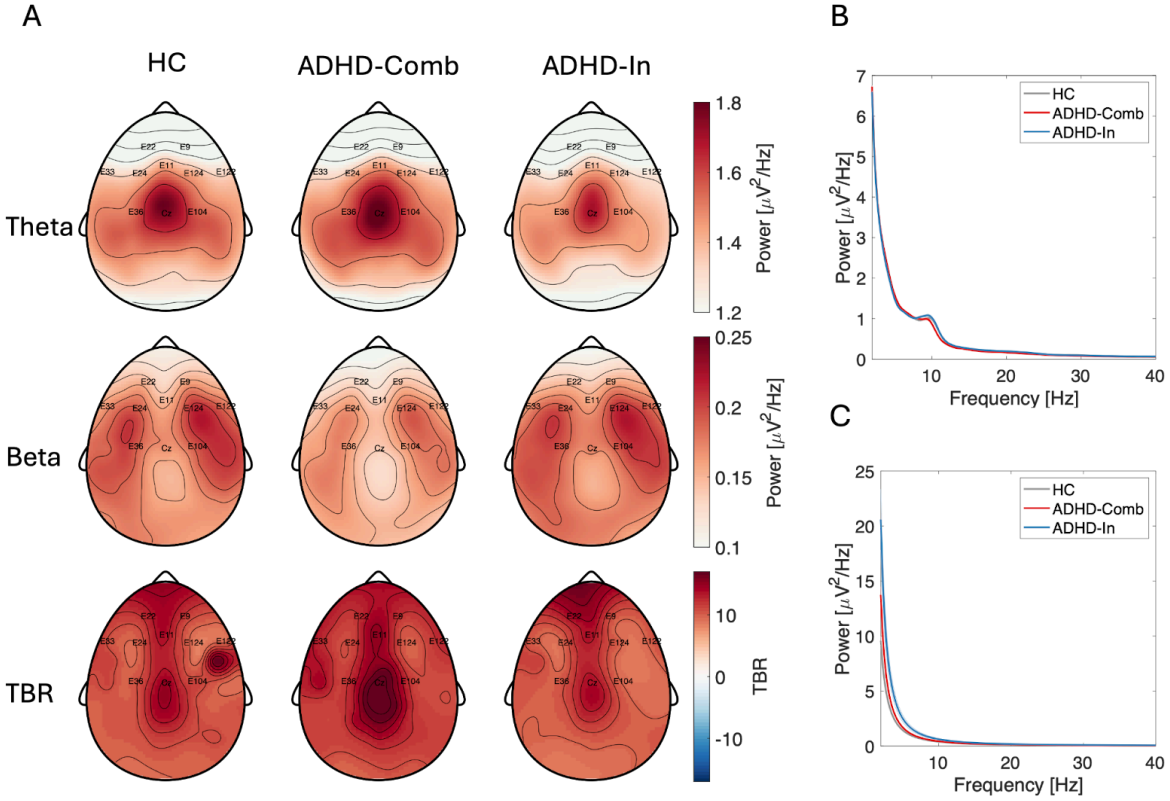
In this study, we used a multiverse analysis across two large EEG datasets to systematically examine the robustness of TBR as an ADHD biomarker. Across hundreds of plausible analytical specifications, we found no consistent evidence that TBR differentiates individuals with ADHD from healthy controls. Instead, TBR estimates varied substantially depending on analytic choices, most notably whether frequency bands were defined relative to IAF and how oscillatory versus aperiodic components were handled. Our findings suggest that individual differences in IAF and the aperiodic 1/f slope strongly influence TBR values and can create the appearance of group differences even when underlying oscillatory activity is similar. Although TBR itself does not provide a reliable diagnostic marker, individualized spectral features, particularly IAF in combination with the aperiodic slope, may offer more promising and clinically informative alternatives. Future work should therefore prioritize

analytic frameworks that explicitly separate oscillatory and aperiodic activity and incorporate IAF when developing EEG-based biomarkers for ADHD.

1. Supplementary Material

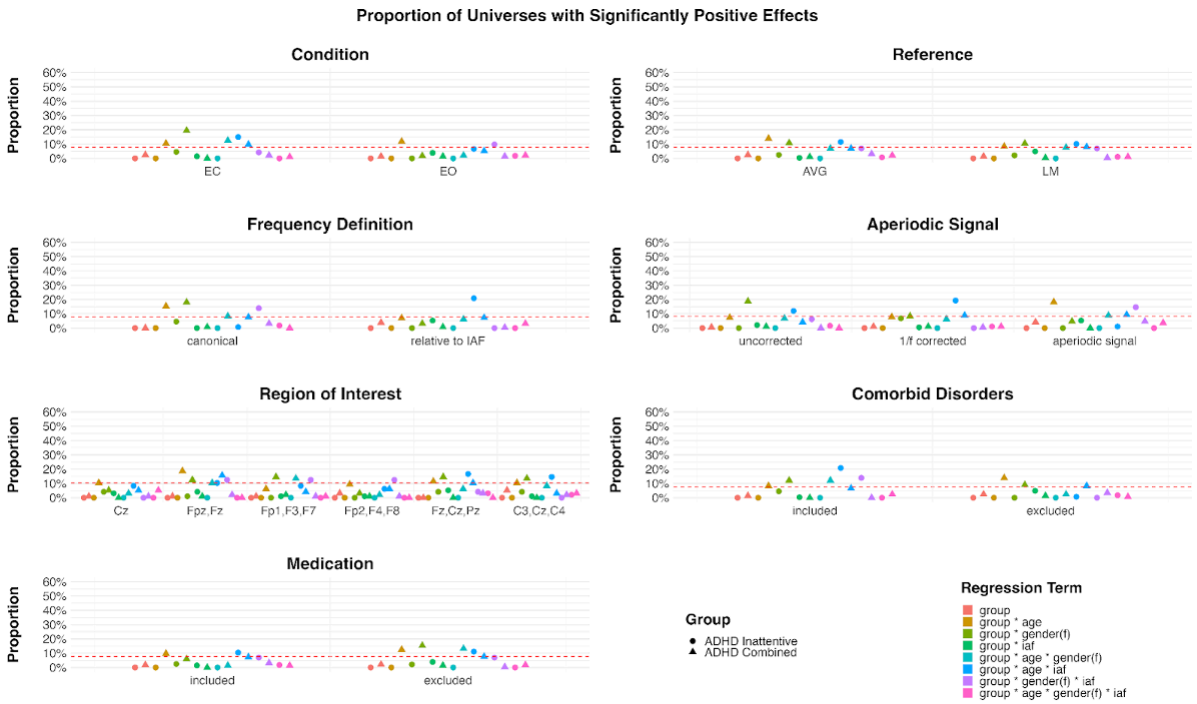
1.1. Healthy Brain Network Sample

1.1.1. 1/f uncorrected power

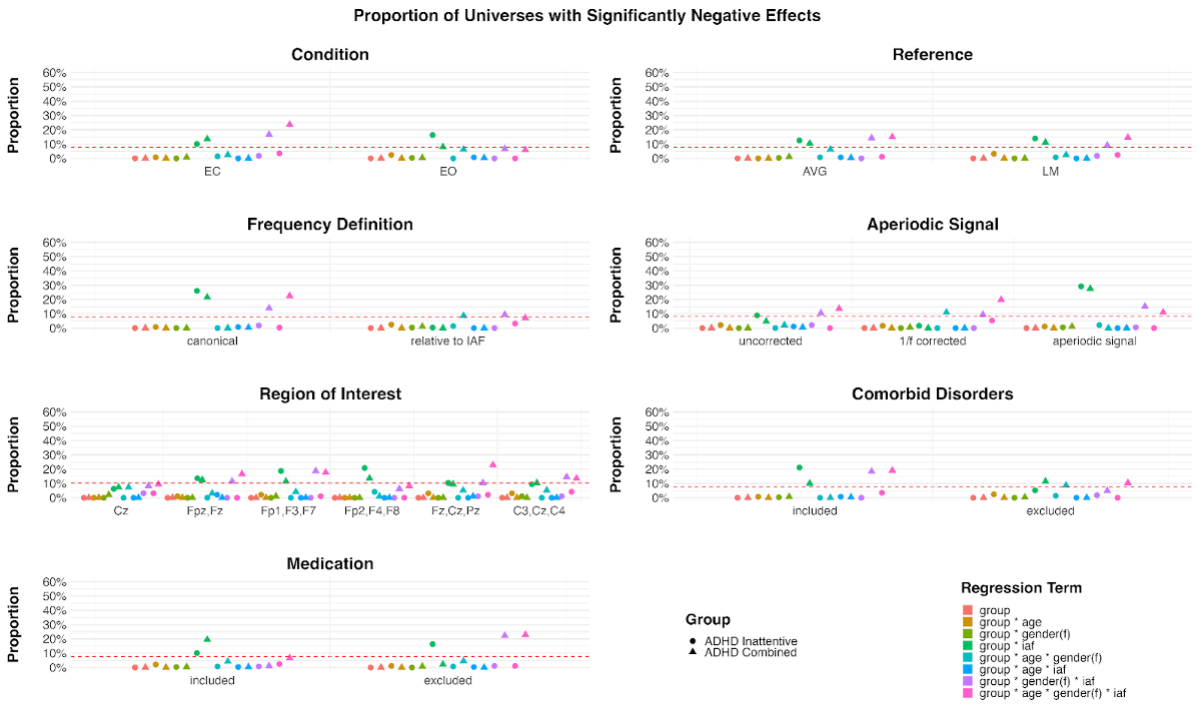


Supplementary Figure 1. Neurophysiological data from the Healthy Brain Network sample. (A) Scalp topographies, (B) power spectra and (C) Scalp topographies for HC, ADHD-Combined, and ADHD-Inattentive groups. The figure displays 1/f uncorrected power during EO condition, computed using a fixed frequency range (theta: 4-8 Hz; beta: 13-30 Hz) and TBR. Electrodes highlighted on the topographies correspond to the six regions of interest derived from literature used across different branches of multiverse analysis. The power spectra and aperiodic signal were computed by averaging across all electrodes within each respective region of interest. Note. HC = healthy controls. Comb = Combined. In = Inattentive.

1.1.2. Proportion Plots Across All Analytical Specifications



Supplementary Figure 2. Proportion of specifications showing a significantly positive effect by regression term and ADHD subtype across all analytical specifications. Each level of Condition, Reference, Frequency Definition, Comorbid Disorder, and Medication includes 288 analytical universes. The Aperiodic Signal factor includes 192 universes per level, and the Region of Interest includes 96 universes per level. The red line in the plot indicates the exact proportion of significant results required for the binomial test to reach statistical significance ($p < .05$) for each subset size.



Supplementary Figure 3. Proportion of specifications showing a significantly negative effect by regression term and ADHD subtype across all analytical specifications. Each level of Condition, Reference, Frequency Definition, Comorbid Disorder, and Medication includes 288 analytical universes. The Aperiodic Signal factor includes 192 universes per level, and the Region of Interest includes 96 universes per level. The red line in the plot indicates the exact proportion of significant results required for the binomial test to reach statistical significance ($p < .05$) for each subset size.

1.1.3. Supplementary Tables with Selected Models

To illustrate results from individual analysis universes in the HBN sample, we present six representative models. In all cases, TBR was computed from EO recordings at Cz with an average reference, including participants with comorbidities and medication. The models vary only in the choice of frequency band definition and spectral measure, thereby highlighting how these analytic decisions shape the observed effects.

Supplementary Table 1. Effects of Group, Age, Gender, and IAF on TBR calculated from 1/f uncorrected power using canonical frequency bands.

Variable	β	SE	CI	t-value	p-value
Intercept	0.00	0.10	-0.2 – 0.2	0.00	p = 1.000
Age	-0.10	0.09	-0.29 – 0.08	-1.09	p = 0.277
Gender[Female]	-0.09	0.15	-0.38 – 0.2	-0.59	p = 0.557
IAF	-0.13	0.11	-0.36 – 0.09	-1.16	p = 0.246
Age * Gender[Female]	0.00	0.14	-0.27 – 0.27	0.00	p = 1.000
Age * IAF	0.00	0.10	-0.2 – 0.2	0.02	p = 0.986
Gender[Female] * IAF	0.08	0.16	-0.23 – 0.39	0.51	p = 0.611
Age * Gender[Female] * IAF	-0.01	0.13	-0.28 – 0.25	-0.11	p = 0.912
ADHD[IN]	0.00	0.12	-0.23 – 0.23	-0.01	p = 0.995
ADHD[IN] * Age	-0.07	0.11	-0.29 – 0.16	-0.58	p = 0.563
ADHD[IN] * Gender[Female]	-0.12	0.18	-0.47 – 0.24	-0.65	p = 0.516
ADHD[IN] * IAF	-0.02	0.13	-0.27 – 0.24	-0.13	p = 0.898
ADHD[IN] * Age * Gender[Female]	0.01	0.18	-0.34 – 0.36	0.04	p = 0.965
ADHD[IN] * Age * IAF	0.09	0.12	-0.15 – 0.32	0.72	p = 0.474
ADHD[IN] * Gender[Female] * IAF	-0.01	0.19	-0.38 – 0.37	-0.03	p = 0.972
ADHD[IN] * Age * Gender[Female] * IAF	0.01	0.18	-0.35 – 0.36	0.04	p = 0.964
ADHD[Comb]	0.14	0.12	-0.09 – 0.37	1.23	p = 0.220
ADHD[Comb] * Age	0.13	0.11	-0.09 – 0.36	1.18	p = 0.238
ADHD[Comb] * Gender[Female]	-0.11	0.20	-0.49 – 0.28	-0.54	p = 0.588
ADHD[Comb] * IAF	-0.01	0.13	-0.26 – 0.25	-0.05	p = 0.959
ADHD[Comb] * Age * Gender[Female]	-0.13	0.19	-0.51 – 0.25	-0.67	p = 0.502
ADHD[Comb] * Age * IAF	-0.01	0.12	-0.25 – 0.23	-0.09	p = 0.931
ADHD[Comb] * Gender[Female] * IAF	-0.06	0.19	-0.44 – 0.31	-0.33	p = 0.741
ADHD[Comb] * Age * Gender[Female] * IAF	0.06	0.17	-0.28 – 0.4	0.37	p = 0.709

Note. In = Inattentive. Comb = Combined. IAF = Individual Alpha Frequency. SE = Standard Error. CI = 95% Confidence Intervals.
$^*p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$

Supplementary Table 2. Effects of Group, Age, Gender, and IAF on TBR calculated from 1/f uncorrected power using frequency bands relative to IAF.

Variable	β	SE	CI	t-value	p-value
Intercept	-0.09	0.07	-0.22 – 0.05	-1.22	p = 0.222
Age	-0.24	0.07	-0.37 – -0.11	-3.59	p = 3.43e-04***
Gender[Female]	-0.05	0.10	-0.25 – 0.15	-0.49	p = 0.623
IAF	-0.53	0.08	-0.69 – -0.37	-6.66	p = 4.37e-11***
Age * Gender[Female]	0.06	0.10	-0.13 – 0.24	0.58	p = 0.559
Age * IAF	0.29	0.07	0.15 – 0.43	4.09	p = 4.69e-05***
Gender[Female] * IAF	-0.03	0.11	-0.25 – 0.18	-0.32	p = 0.751
Age * Gender[Female] * IAF	-0.14	0.09	-0.32 – 0.04	-1.50	p = 0.135
ADHD[IN]	0.04	0.08	-0.12 – 0.21	0.54	p = 0.591
ADHD[IN] * Age	0.02	0.08	-0.13 – 0.18	0.26	p = 0.793
ADHD[IN] * Gender[Female]	-0.04	0.13	-0.28 – 0.21	-0.28	p = 0.779
ADHD[IN] * IAF	-0.10	0.09	-0.28 – 0.08	-1.14	p = 0.256
ADHD[IN] * Age * Gender[Female]	0.01	0.13	-0.24 – 0.25	0.04	p = 0.965
ADHD[IN] * Age * IAF	-0.11	0.08	-0.27 – 0.05	-1.31	p = 0.189
ADHD[IN] * Gender[Female] * IAF	0.07	0.13	-0.2 – 0.33	0.49	p = 0.624
ADHD[IN] * Age * Gender[Female] * IAF	0.16	0.13	-0.09 – 0.41	1.25	p = 0.213
ADHD[Comb]	0.08	0.08	-0.08 – 0.24	1.00	p = 0.316
ADHD[Comb] * Age	0.17	0.08	0.02 – 0.33	2.18	p = 0.030*
ADHD[Comb] * Gender[Female]	0.02	0.14	-0.25 – 0.29	0.18	p = 0.860
ADHD[Comb] * IAF	-0.12	0.09	-0.3 – 0.06	-1.35	p = 0.177
ADHD[Comb] * Age * Gender[Female]	-0.05	0.14	-0.32 – 0.21	-0.39	p = 0.696
ADHD[Comb] * Age * IAF	-0.17	0.08	-0.34 – -0.01	-2.06	p = 0.040*
ADHD[Comb] * Gender[Female] * IAF	0.02	0.14	-0.25 – 0.28	0.14	p = 0.888
ADHD[Comb] * Age * Gender[Female] * IAF	0.15	0.12	-0.09 – 0.39	1.26	p = 0.209

Note. In = Inattentive. Comb = Combined. IAF = Individual Alpha Frequency. SE = Standard Error. CI = 95% Confidence Intervals.

**p < 0.05. **p < 0.01. ***p < 0.001*

Supplementary Table 3. Effects of Group, Age, Gender, and IAF on TBR calculated from aperiodic-adjusted power using canonical frequency bands.

Variable	β	SE	CI	t-value	p-value
Intercept	-0.01	0.10	-0.21 – 0.18	-0.15	p = 0.882
Age	-0.05	0.10	-0.23 – 0.14	-0.50	p = 0.618
Gender[Female]	-0.08	0.15	-0.37 – 0.21	-0.54	p = 0.591
IAF	-0.12	0.11	-0.34 – 0.1	-1.05	p = 0.292
Age * Gender[Female]	-0.02	0.14	-0.28 – 0.25	-0.13	p = 0.899
Age * IAF	-0.06	0.10	-0.26 – 0.14	-0.58	p = 0.559
Gender[Female] * IAF	0.03	0.16	-0.27 – 0.34	0.22	p = 0.825
Age * Gender[Female] * IAF	0.10	0.13	-0.16 – 0.36	0.73	p = 0.466
ADHD[IN]	0.00	0.12	-0.23 – 0.23	0.01	p = 0.994
ADHD[IN] * Age	-0.06	0.11	-0.28 – 0.16	-0.55	p = 0.585
ADHD[IN] * Gender[Female]	-0.05	0.18	-0.41 – 0.3	-0.28	p = 0.780
ADHD[IN] * IAF	-0.09	0.13	-0.35 – 0.16	-0.73	p = 0.467
ADHD[IN] * Age * Gender[Female]	0.07	0.18	-0.28 – 0.42	0.40	p = 0.691
ADHD[IN] * Age * IAF	0.15	0.12	-0.08 – 0.38	1.26	p = 0.208
ADHD[IN] * Gender[Female] * IAF	0.11	0.19	-0.26 – 0.49	0.59	p = 0.552
ADHD[IN] * Age * Gender[Female] * IAF	-0.15	0.18	-0.51 – 0.2	-0.84	p = 0.402
ADHD[Comb]	0.10	0.12	-0.13 – 0.33	0.84	p = 0.400
ADHD[Comb] * Age	0.08	0.11	-0.15 – 0.3	0.67	p = 0.503
ADHD[Comb] * Gender[Female]	0.03	0.20	-0.35 – 0.42	0.18	p = 0.859
ADHD[Comb] * IAF	-0.02	0.13	-0.28 – 0.23	-0.16	p = 0.873
ADHD[Comb] * Age * Gender[Female]	-0.06	0.19	-0.44 – 0.32	-0.32	p = 0.751
ADHD[Comb] * Age * IAF	0.09	0.12	-0.14 – 0.33	0.78	p = 0.434
ADHD[Comb] * Gender[Female] * IAF	-0.15	0.19	-0.52 – 0.23	-0.75	p = 0.452
ADHD[Comb] * Age * Gender[Female] * IAF	-0.11	0.17	-0.45 – 0.24	-0.61	p = 0.544

Note. In = Inattentive. Comb = Combined. IAF = Individual Alpha Frequency. SE = Standard Error. CI = 95% Confidence Intervals.

$^{*}p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$

Supplementary Table 4. Effects of Group, Age, Gender, and IAF on TBR calculated from aperiodic-adjusted power using frequency bands relative to IAF.

Variable	β	SE	CI	t-value	p-value
Intercept	-0.06	0.10	-0.26 – 0.13	-0.63	p = 0.528
Age	-0.23	0.09	-0.42 – -0.04	-2.43	p = 0.015*
Gender[Female]	0.09	0.15	-0.19 – 0.38	0.64	p = 0.522
IAF	0.25	0.11	0.03 – 0.47	2.18	p = 0.029*
Age * Gender[Female]	0.08	0.14	-0.19 – 0.34	0.55	p = 0.580
Age * IAF	-0.03	0.10	-0.23 – 0.16	-0.34	p = 0.736
Gender[Female] * IAF	-0.23	0.16	-0.54 – 0.08	-1.47	p = 0.142
Age * Gender[Female] * IAF	0.09	0.13	-0.17 – 0.35	0.66	p = 0.508
ADHD[IN]	0.03	0.12	-0.2 – 0.26	0.26	p = 0.793
ADHD[IN] * Age	0.01	0.11	-0.22 – 0.23	0.05	p = 0.960
ADHD[IN] * Gender[Female]	0.05	0.18	-0.3 – 0.41	0.29	p = 0.773
ADHD[IN] * IAF	-0.09	0.13	-0.34 – 0.17	-0.67	p = 0.502
ADHD[IN] * Age * Gender[Female]	0.07	0.18	-0.27 – 0.42	0.42	p = 0.678
ADHD[IN] * Age * IAF	0.02	0.12	-0.22 – 0.25	0.13	p = 0.896
ADHD[IN] * Gender[Female] * IAF	0.05	0.19	-0.32 – 0.42	0.26	p = 0.797
ADHD[IN] * Age * Gender[Female] * IAF	-0.04	0.18	-0.39 – 0.32	-0.21	p = 0.837
ADHD[Comb]	0.12	0.12	-0.11 – 0.35	1.00	p = 0.319
ADHD[Comb] * Age	0.21	0.11	-0.02 – 0.43	1.82	p = 0.069
ADHD[Comb] * Gender[Female]	-0.12	0.20	-0.5 – 0.27	-0.60	p = 0.548
ADHD[Comb] * IAF	0.04	0.13	-0.22 – 0.29	0.29	p = 0.774
ADHD[Comb] * Age * Gender[Female]	-0.15	0.19	-0.52 – 0.23	-0.76	p = 0.449
ADHD[Comb] * Age * IAF	0.03	0.12	-0.2 – 0.27	0.27	p = 0.785
ADHD[Comb] * Gender[Female] * IAF	0.05	0.19	-0.32 – 0.43	0.29	p = 0.775
ADHD[Comb] * Age * Gender[Female] * IAF	-0.13	0.17	-0.47 – 0.21	-0.75	p = 0.451

Note. In = Inattentive. Comb = Combined. IAF = Individual Alpha Frequency. SE = Standard Error. CI = 95% Confidence Intervals.

$*p < 0.05$. $**p < 0.01$. $***p < 0.001$

Supplementary Table 5. Effects of Group, Age, Gender, and IAF on TBR calculated from aperiodic signal power using canonical frequency bands.

Variable	β	SE	CI	t-value	p-value
Intercept	0.13	0.09	-0.05 – 0.32	1.40	p = 0.162
Age	-0.20	0.09	-0.37 – -0.02	-2.20	p = 0.028*
Gender[Female]	-0.30	0.14	-0.57 – -0.03	-2.15	p = 0.032*
IAF	-0.27	0.11	-0.48 – -0.06	-2.51	p = 0.012*
Age * Gender[Female]	0.01	0.13	-0.24 – 0.26	0.06	p = 0.953
Age * IAF	0.10	0.10	-0.09 – 0.28	1.02	p = 0.309
Gender[Female] * IAF	0.24	0.15	-0.04 – 0.53	1.66	p = 0.098
Age * Gender[Female] * IAF	-0.25	0.12	-0.49 – -0.01	-2.01	p = 0.045*
ADHD[IN]	-0.01	0.11	-0.23 – 0.2	-0.11	p = 0.911
ADHD[IN] * Age	-0.08	0.11	-0.29 – 0.13	-0.75	p = 0.456
ADHD[IN] * Gender[Female]	-0.33	0.17	-0.66 – 0	-1.95	p = 0.052
ADHD[IN] * IAF	0.16	0.12	-0.08 – 0.4	1.33	p = 0.185
ADHD[IN] * Age * Gender[Female]	-0.08	0.17	-0.41 – 0.25	-0.49	p = 0.622
ADHD[IN] * Age * IAF	-0.03	0.11	-0.25 – 0.18	-0.31	p = 0.757
ADHD[IN] * Gender[Female] * IAF	-0.22	0.18	-0.57 – 0.13	-1.23	p = 0.220
ADHD[IN] * Age * Gender[Female] * IAF	0.33	0.17	0 – 0.66	1.93	p = 0.053
ADHD[Comb]	0.10	0.11	-0.12 – 0.31	0.90	p = 0.367
ADHD[Comb] * Age	0.02	0.11	-0.19 – 0.23	0.15	p = 0.882
ADHD[Comb] * Gender[Female]	-0.27	0.18	-0.63 – 0.09	-1.49	p = 0.136
ADHD[Comb] * IAF	0.13	0.12	-0.11 – 0.37	1.07	p = 0.283
ADHD[Comb] * Age * Gender[Female]	-0.03	0.18	-0.38 – 0.33	-0.14	p = 0.886
ADHD[Comb] * Age * IAF	-0.14	0.11	-0.36 – 0.08	-1.27	p = 0.205
ADHD[Comb] * Gender[Female] * IAF	-0.13	0.18	-0.48 – 0.23	-0.71	p = 0.475
ADHD[Comb] * Age * Gender[Female] * IAF	0.43	0.16	0.11 – 0.75	2.66	p = 0.008**

Note. In = Inattentive. Comb = Combined. IAF = Individual Alpha Frequency. SE = Standard Error. CI = 95% Confidence Intervals.

**p < 0.05. **p < 0.01. ***p < 0.001*

Supplementary Table 6. Effects of Group, Age, Gender, and IAF on TBR calculated from aperiodic signal power using frequency bands relative to IAF.

Variable	β	SE	CI	t-value	p-value
Intercept	-0.14	0.06	-0.25 – -0.03	-2.54	p = 0.011*
Age	-0.14	0.05	-0.24 – -0.04	-2.65	p = 0.008**
Gender[Female]	-0.03	0.08	-0.19 – 0.13	-0.34	p = 0.737
IAF	-0.69	0.06	-0.81 – -0.56	-10.80	p = 6.23e-26***
Age * Gender[Female]	0.04	0.08	-0.11 – 0.19	0.56	p = 0.576
Age * IAF	0.33	0.06	0.22 – 0.44	5.80	p = 8.66e-09***
Gender[Female] * IAF	0.11	0.09	-0.06 – 0.28	1.27	p = 0.205
Age * Gender[Female] * IAF	-0.17	0.07	-0.32 – -0.02	-2.29	p = 0.022*
ADHD[IN]	0.08	0.07	-0.05 – 0.21	1.23	p = 0.220
ADHD[IN] * Age	0.04	0.06	-0.08 – 0.17	0.67	p = 0.505
ADHD[IN] * Gender[Female]	-0.06	0.10	-0.26 – 0.14	-0.57	p = 0.567
ADHD[IN] * IAF	-0.11	0.07	-0.25 – 0.03	-1.55	p = 0.121
ADHD[IN] * Age * Gender[Female]	-0.02	0.10	-0.22 – 0.18	-0.20	p = 0.841
ADHD[IN] * Age * IAF	-0.09	0.07	-0.22 – 0.04	-1.31	p = 0.190
ADHD[IN] * Gender[Female] * IAF	0.08	0.11	-0.13 – 0.29	0.76	p = 0.446
ADHD[IN] * Age * Gender[Female] * IAF	0.06	0.10	-0.14 – 0.26	0.61	p = 0.545
ADHD[Comb]	0.10	0.07	-0.03 – 0.23	1.50	p = 0.134
ADHD[Comb] * Age	0.10	0.06	-0.03 – 0.22	1.49	p = 0.136
ADHD[Comb] * Gender[Female]	0.14	0.11	-0.08 – 0.35	1.25	p = 0.213
ADHD[Comb] * IAF	-0.16	0.07	-0.3 – -0.01	-2.14	p = 0.033*
ADHD[Comb] * Age * Gender[Female]	0.07	0.11	-0.15 – 0.28	0.61	p = 0.540
ADHD[Comb] * Age * IAF	-0.13	0.07	-0.26 – 0	-1.96	p = 0.051
ADHD[Comb] * Gender[Female] * IAF	-0.11	0.11	-0.32 – 0.1	-1.00	p = 0.317
ADHD[Comb] * Age * Gender[Female] * IAF	0.06	0.10	-0.13 – 0.25	0.62	p = 0.536

Note. In = Inattentive. Comb = Combined. IAF = Individual Alpha Frequency. SE = Standard Error. CI = 95% Confidence Intervals.

**p < 0.05. **p < 0.01. ***p < 0.001*

1.1.4. Bootstrap Robustness Analysis

To evaluate the robustness of our multiverse findings to sample size imbalances, we conducted a bootstrap-based sensitivity analysis. For each universe, we identified the group with the lowest number of participants (e.g., ADHD-Inattentive, female, no medication, no comorbidities) and randomly resampled all groups to this minimum size. The multiverse analysis was then re-run on these balanced subsamples. This procedure was repeated 1000 times with independent random draws. The bootstrap analysis yielded results that closely mirrored the main multiverse findings. The frequency and direction of significant effects were consistent across resampled iterations, indicating that our conclusions were not driven by unequal group sizes or by idiosyncratic subsamples (Supplementary Table 8).

Supplementary Table 8. Results of the bootstrap robustness analysis across 1000 resampled multiverse iterations.

	Significantly Positive	***Significantly Negative***	***Not Significant***	***Total***
ADHD[IN]	13.28 ± 3.5 (2.31%)	15.86 ± 3.58 (2.75 %)	546.86 ± 5.13 (94.94 %)	576
ADHD[IN] * Age	10.17 ± 3.15 (1.77 %)	17.15 ± 3.77 (2.98 %)	548.68 ± 4.92 (95.26 %)	576
ADHD[IN] * Gender[Female]	20.68 ± 3.33 (3.59 %)	6.68 ± 2.43 (1.16 %)	548.64 ± 4.17 (95.25 %)	576
ADHD[IN] * IAF	23.90 ± 4.75 (4.15 %)	61.40 ± 6.42 (10.66 %)	490.70 ± 7.78 (85.19 %)	576
ADHD[IN] * Age * Gender[Female]	4.20 ± 1.99 (0.73 %)	6.16 ± 2.25 (1.07 %)	565.64 ± 2.96 (98.20 %)	576
ADHD[IN] * Age * IAF	62.80 ± 6.63 (10.90 %)	30.16 ± 5.03 (5.24 %)	483.04 ± 8.01 (83.86 %)	576
ADHD[IN] * Gender[Female] * IAF	40.34 ± 4.75 (7.00 %)	18.59 ± 4.03 (3.23 %)	517.07 ± 6.12 (89.77 %)	576
ADHD[IN] * Age * Gender[Female] * IAF	22.90 ± 4.47 (3.98 %)	27.34 ± 4.67 (4.75 %)	525.76 ± 6.46 (91.28 %)	576

ADHD[Comb]	28.03 ± 4.76 (4.87 %)	4.06 ± 2.00 (0.70 %)	543.92 ± 5.13 (94.43 %)	576
ADHD[Comb] * Age	51.01 ± 6.32 (8.86 %)	6.30 ± 2.44 (1.09 %)	518.70 ± 6.80 (90.05 %)	576
ADHD[Comb] * Gender[Female]	56.06 ± 5.82 (9.73 %)	4.25 ± 1.97 (0.74 %)	515.69 ± 6.10 (89.53 %)	576
ADHD[Comb] * IAF	12.01 ± 3.42 (2.09 %)	46.67 ± 6.07 (8.10 %)	517.32 ± 6.87 (89.81 %)	576
ADHD[Comb] * Age * Gender[Female]	44.40 ± 4.94 (7.71 %)	15.09 ± 3.30 (2.62 %)	516.51 ± 5.95 (89.67 %)	576
ADHD[Comb] * Age * IAF	44.38 ± 5.76 (7.70 %)	27.80 ± 4.87 (4.83 %)	503.83 ± 7.10 (87.47 %)	576
ADHD[Comb] * Gender[Female] * IAF	12.76 ± 3.25 (2.22 %)	64.78 ± 5.87 (11.25 %)	498.46 ± 6.68 (86.54 %)	576
ADHD[Comb] * Age * Gender[Female] * IAF	9.99 ± 2.80 (1.74 %)	87.31 ± 7.41 (15.16 %)	478.70 ± 7.85 (83.11 %)	576

Note. Values reported represent the mean across 1000 iterations ± standard deviation, as well as the percentage reflecting the mean value. Positive universe = higher TBR in ADHD compared to HC. Negative universe = lower TBR in ADHD compared to HC. IN = Inattentive. Comb = Combined. IAF = Individual alpha frequency.

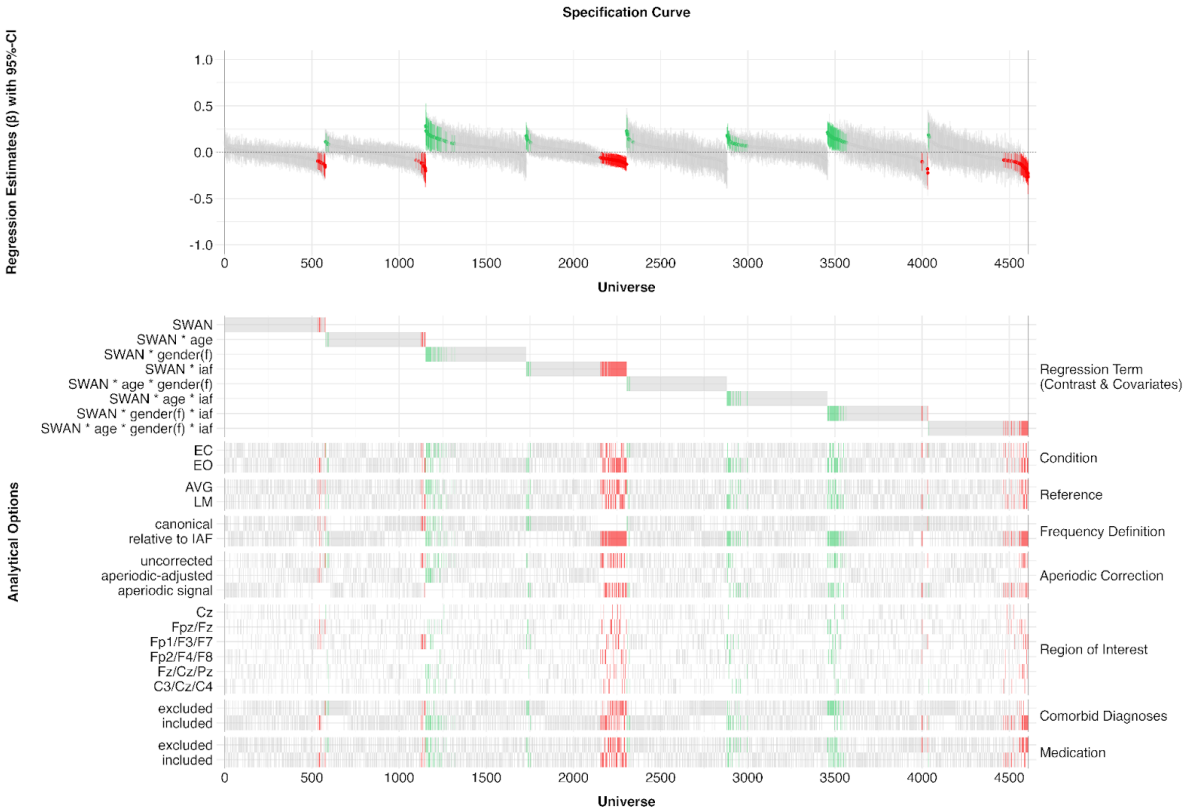
1.1.5. Dimensional multiverse analysis with SWAN Total scores

To complement the categorical analyses based on the diagnostic group (i.e., HC, ADHD-Inattentive, ADHD-Combined), we conducted an identical multiverse analysis using the SWAN scores [40] as a dimensional predictor of ADHD symptom severity in the HBN sample only, as SWAN scores were not available for the validation sample. This dimensional approach tested the robustness of findings across the full spectrum of ADHD symptomatology, avoiding potential information loss from diagnostic thresholds while increasing statistical power and capturing subclinical variability. The identical set of analytical specifications and model structures were applied, substituting the categorical Group factor with continuous SWAN scores and their interactions with age, gender, and IAF.

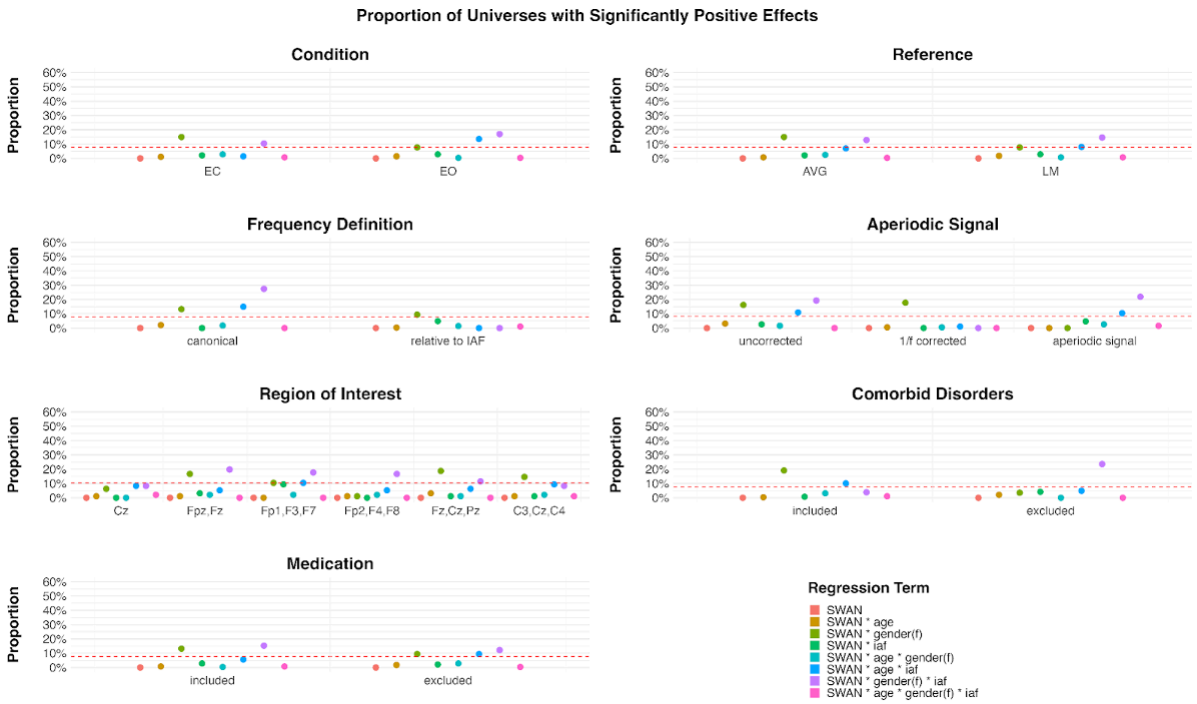
Supplementary Table 7. Results of the multiverse analysis using dimensional SWAN scores instead of categorical ADHD diagnoses.

	Significantly Positive	***Significantly Negative***	***Not Significant***	***Total***
SWAN	0 (0.00 %)	21 (3.65 %)	555 (96.35 %)	576
SWAN * Age	7 (1.21 %)	20 (3.47 %)	549 (95.31 %)	576
SWAN * Gender[Female]	65 (11.28%)	0 (0.00%)	562 (88.72 %)	576
SWAN * IAF	14 (2.43 %)	134 (23.26 %)	428 (74.31 %)	576
SWAN * Age * Gender[Female]	9 (1.56 %)	0 (0.00 %)	567 (98.44 %)	576
SWAN * Age * IAF	43 (7.47%)	0 (0.00 %)	533 (92.53 %)	576
SWAN * Gender[Female] * IAF	79 (13.71 %)	5 (0.87 %)	492 (85.42 %)	576
SWAN * Age * Gender[Female] * IAF	3 (0.52 %)	67 (11.63 %)	506 (87.85 %)	576

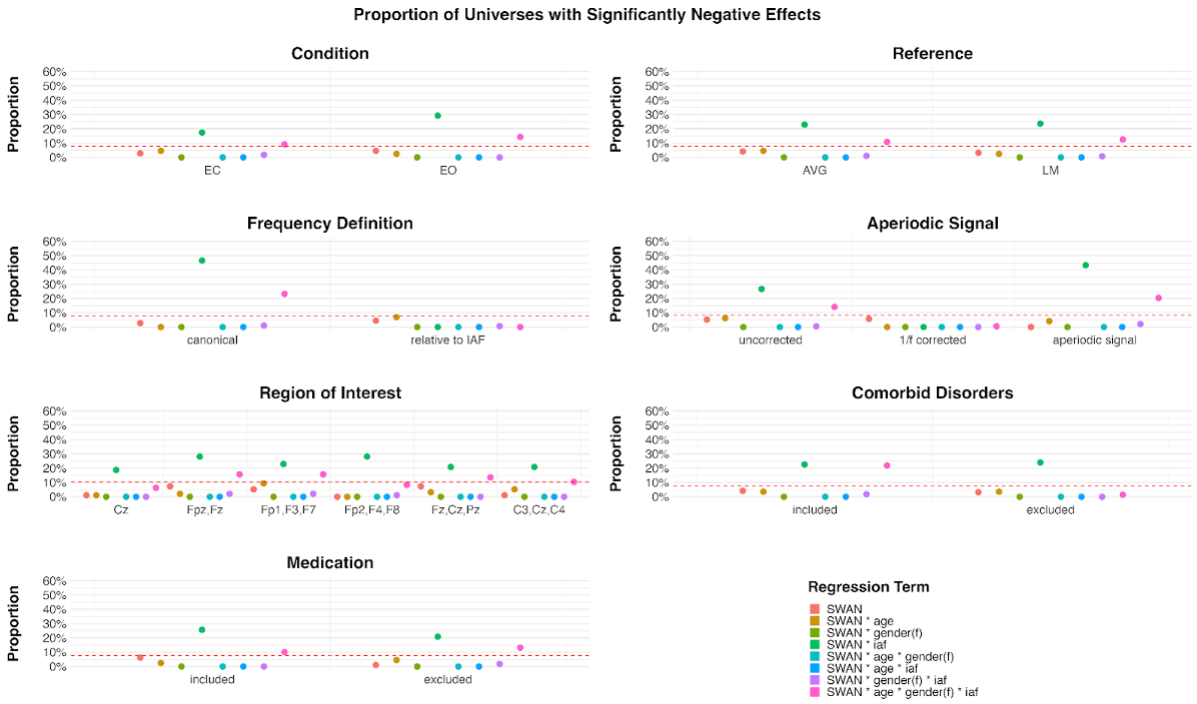
Note. Positive universe = higher TBR in ADHD compared to HC. Negative universe = lower TBR in ADHD compared to HC. IAF = Individual alpha frequency.



Supplementary Figure 3. Specification curve representing all universes for the dimensional SWAN score analysis. The top panel shows regression estimates for each specification, sorted by regression estimate, with 95% confidence intervals. Statistically significant positive estimates are shown in green, negative estimates in red, and non-significant estimates in gray. The bottom panel indicates which analytical choices were associated with significantly positive (i.e., higher TBR for ADHD compared to HC) or negative (i.e., lower TBR for ADHD compared to HC) effects across specifications. Note. CI = confidence interval. f = female. EC = eyes closed. EO = eyes open. AVG = average reference. LM = linked mastoid reference. canonical = canonical frequency range. relative to IAF = bandwidth relative to individual alpha frequency.



Supplementary Figure 4. Proportion of specifications showing a significantly positive effect by regression term across all analytical specifications. Each level of Condition, Reference, Frequency Definition, Comorbid Disorder, and Medication includes 288 analytical universes. The Aperiodic Signal factor includes 192 universes per level, and the Region of Interest includes 96 universes per level. The red line in the plot indicates the exact proportion of significant results required for the binomial test to reach statistical significance ($p < .05$) for each subset size.



Supplementary Figure 5. Proportion of specifications showing a significantly negative effect by regression term across all analytical specifications. Each level of Condition, Reference, Frequency Definition, Comorbid Disorder, and Medication includes 288 analytical universes. The Aperiodic Signal factor includes 192 universes per level, and the Region of Interest includes 96 universes per level. The red line in the plot indicates the exact proportion of significant results required for the binomial test to reach statistical significance ($p < .05$) for each subset size.

1.2. Validation Sample

1.2.1. Methods

1.2.1.1. Participants

The second dataset used for validation was obtained from a previously conducted multicenter study carried out as part of the project coordinated by the Brain and Trauma Foundation Grisons, Switzerland [35–37]. The dataset comprised 381 participants, including both individuals with ADHD and healthy controls. The sample consisted of 191 healthy controls (mean age = 14.23 years, SD = 4.39, range = 6.28-21.90), 110 participants with ADHD-Inattentive Type (mean age = 12.88 years, SD = 3.35, range = 6.48-21.81), 65 with ADHD-Combined Type (mean age = 11.19 years, SD = 3.73, range = 6.28-21.35), and 15 with ADHD-Hyperactive/Impulsive Type (mean age = 10.34 years, SD = 3.55, range = 6.39-19.69). Due to the small sample size of the ADHD-Hyperactive/Impulsive group, it was excluded from further analyses.

Prior to participation, interested individuals completed an ADHD screening questionnaire. During the initial appointment, informed consent was obtained and a DSM-5-based ADHD diagnosis was made or confirmed by board-certified psychiatrists or clinical psychologists. The first testing session included structured clinical interviews (SKID; [107]), IQ testing using age-appropriate instruments (CFT 1-R or CFT 20-R: [108]; WMT-2: [109]), and a battery of neuropsychological tests (candit.com). EEG data were recorded in a subsequent session. General exclusion criteria included a history of traumatic brain injury, loss of consciousness, major psychiatric disorders (e.g., depression or schizophrenia), drug abuse, epilepsy, pregnancy, and an IQ below 80. All participants were required to be fluent in German, capable of understanding study procedures, and able to provide informed consent. While all participants were medication-free at the time of EEG recording, 139 ADHD participants were regularly treated with methylphenidate in daily life (e.g., Ritalin, Concerta, Elvanse). For detailed demographic information, see Supplementary Table 9. The study was approved by the cantonal ethics committee of Zurich (KEK-ZH 2013-0327 / EKNZ 2014-160), and data collection took place between July 2014 and July 2017 at five locations across Switzerland (Zürich, Chur, Lausanne, Lucerne, and Rapperswil).

Please note that detailed information on comorbid diagnoses and medication status was available only in the HBN sample; accordingly, these factors were examined exclusively in HBN. In the validation dataset, we knew that some participants were medicated but lacked details on compound, dose, or timing because this was a secondary dataset not collected by

us and the information was not accessible. As a result, the specification set for the validation sample omitted the comorbidity and medication dimensions, reducing the number of analytic universes to 144 per group contrast (vs. 576 in HBN).

1.2.1.2. EEG Acquisition

In the validation dataset, resting state EEG was recorded using a 19-channel system (NeuroAmp ×23) in combination with the ERPrec software (BEE Medic GmbH). EEG data were sampled at 500 Hz and recorded during a total of 8 minutes of resting-state recording, consisting of 4 minutes with EO and 4 minutes with EC. Electrodes were positioned according to the international 10–20 system using an electrode cap equipped with tin electrodes (Electro-Cap International Inc., Eaton, OH, USA). The recording reference was set to the linked earlobes, and electrode impedances were maintained below 5 kΩ. To ensure consistency with the main dataset (i.e. Healthy Brain Network), all EEG recordings were downsampled to 250 Hz prior to further processing.

1.2.2. Results

1.2.2.1. Demographics

Supplementary Table 9 and Figure 6 summarize the demographic characteristics of the validation sample. Similar to the main sample, we first compared demographic characteristics of all three groups (i.e., HC, ADHD-Combined, ADHD-Inattentive). The chi-square test revealed a significant association between gender and diagnosis group, indicating that gender distribution differed across the three groups (χ^2 = 33.63, p = 4.9e-8). Next, the ANOVAs revealed significant group effects for age (F = 12.24, p = 7.8e-6), and IQ (F = 9.86, p = 7.2e-5), indicating that the HC, ADHD-Combined and ADHD-Inattentive differed significantly across these variables. The effect of the group for IAF was not statistically significant (F = 2.47, p = 0.087), suggesting that the groups did not differ in their IAF.

Supplementary Table 9. Demographics information of the validation sample.

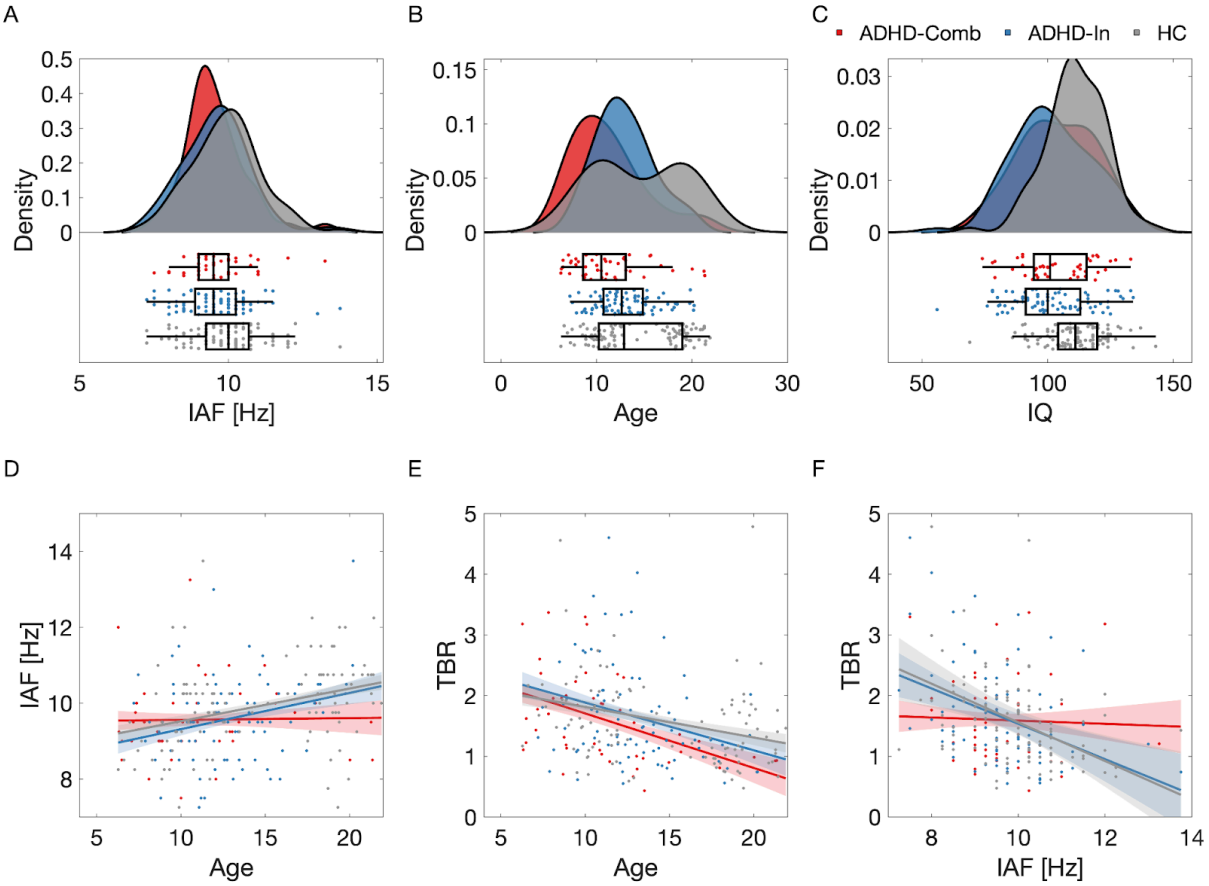
	HC	***Combined***	***Inattentive***	***X^2***	***p-value***
N	123	64	108		
Male	55	44	89		

Female	68		20		19		33.63	5e-8***

	Mean	***SD***	***Mean***	***SD***	***Mean***	***SD***	***F-value***	***p-value***
Age (years)	14.24	4.76	11.23	3.75	12.80	3.27	12.24	7.8e-6***
IQ	110.00	12.70	102.47	14.35	102.93	15.71	9.86	7.2e-5***
IAF	9.91	1.15	9.47	1.04	9.60	1.13	2.47	0.087

Note. HC = Healthy control. SD = Standard deviation. IN = Inattentive. HY = Hyperactive.

**p < 0.05. **p < 0.01. ***p < 0.001*

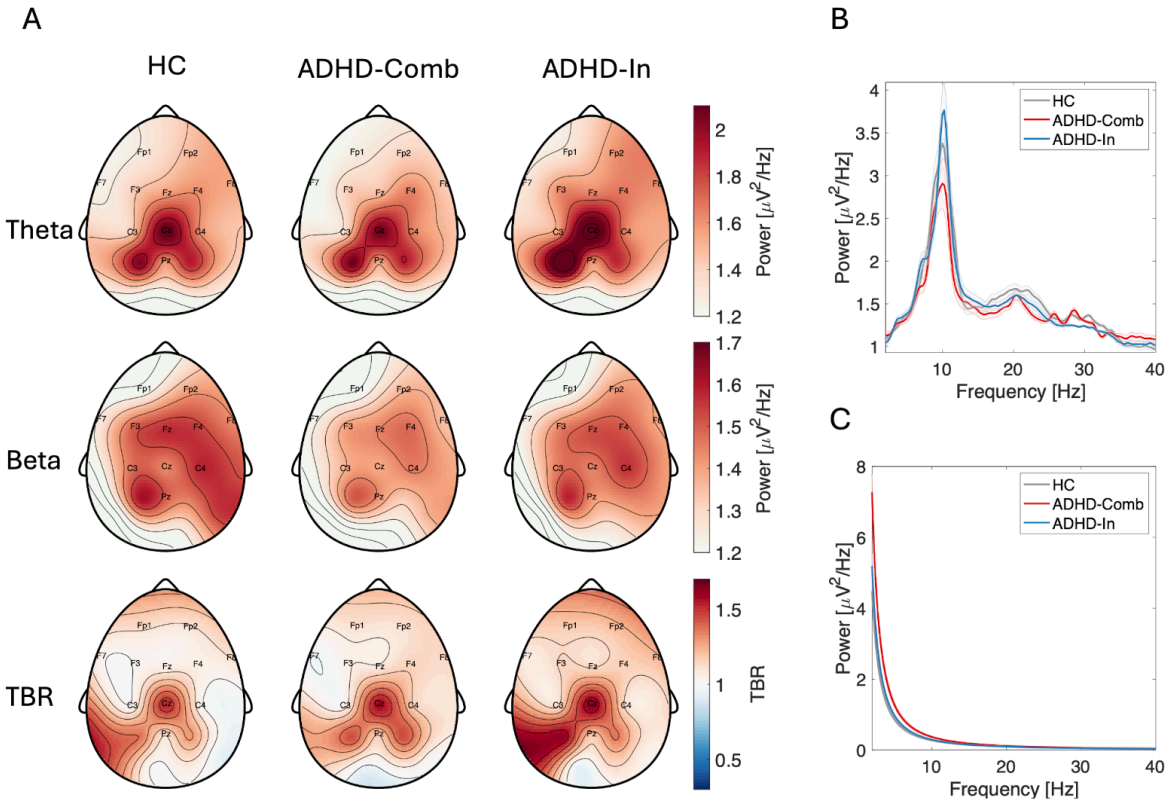


Supplementary Figure 6. Overview of sample characteristics and variable distributions. Density plots show the distribution of (A) age, (B) IQ and (C) IAF across the three groups. (D-F) Scatter plots display the relationship between (D) IAF and age, (E) TBR and age, (F) and TBR and IAF across all participants. The regression line is shown with standard error bands. Note. IAF = Individual Alpha Frequency. TBR = Theta-Beta Ratio. HC = Healthy Control.

1.2.2.2. EEG features

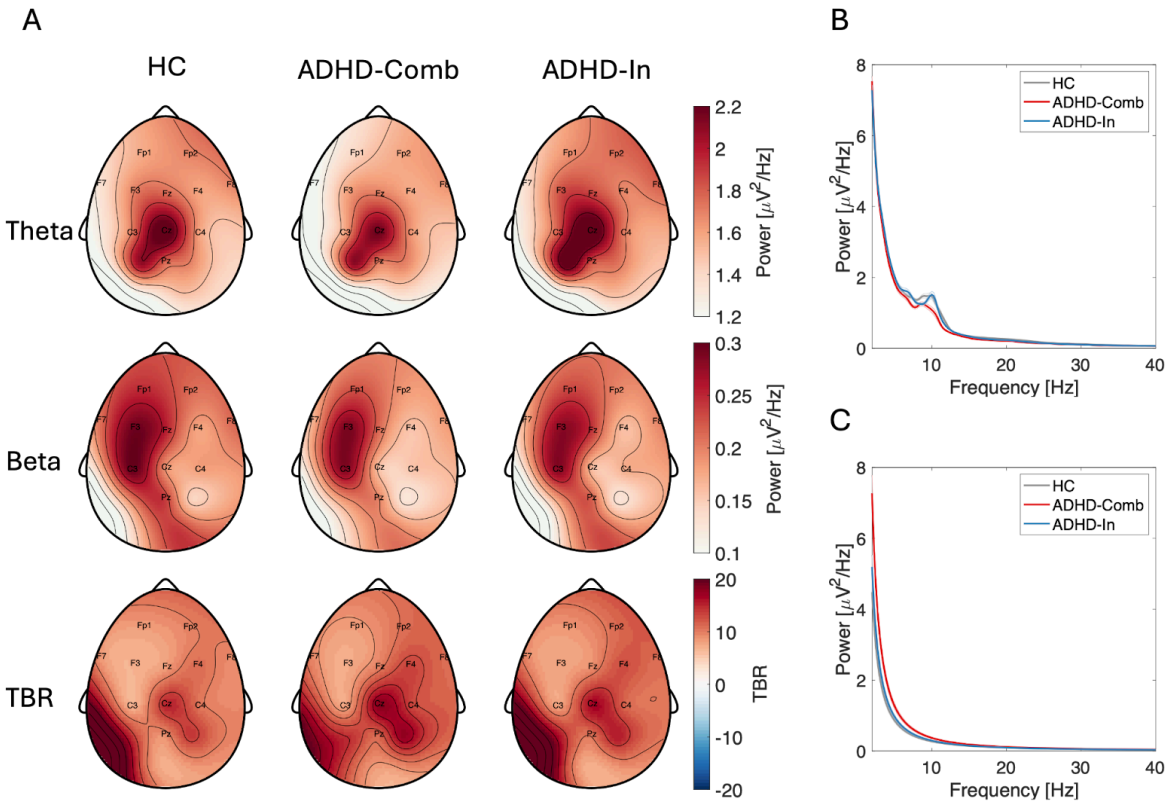
1.2.2.2.1. Aperiodic-adjusted power

To visualize the neurophysiological data, we plotted the scalp topographies of theta, beta and TBR as well as power spectra and aperiodic signal for HC, ADHD-Combined, and ADHD-Inattentive groups (Supplementary Figure 7). The figure displays aperiodic-adjusted power during EO condition, computed using a canonical frequency range (theta: 4-8 Hz; beta: 13-30 Hz).



Supplementary Figure 7. Neurophysiological data from the validation sample. (A) Scalp topographies, (B) power spectra and (C) for HC, ADHD-Combined, and ADHD-Inattentive groups. The figure displays aperiodic-adjusted power during EO condition, computed using a canonical frequency range (theta: 4-8 Hz; beta: 13-30 Hz) and TBR. Electrodes highlighted on the topographies correspond to the six regions of interest derived from literature used across different branches of multiverse analysis. The power spectra and aperiodic signal were computed by averaging across all electrodes within each respective region of interest. Note. HC = healthy controls. Comb = Combined. In = Inattentive.

1.2.2.2.2. 1/f uncorrected power



Supplementary Figure 8. Neurophysiological data from the validation sample. (A) Scalp topographies, (B) power spectra and (C) for HC, ADHD-Combined, and ADHD-Inattentive groups. The figure displays 1/f uncorrected power during EO condition, computed using a fixed frequency range (theta: 4-8 Hz; beta: 13-30 Hz) and TBR. Electrodes highlighted on the topographies correspond to the six regions of interest derived from literature used across different branches of multiverse analysis. The power spectra and aperiodic signal were computed by averaging across all electrodes within each respective region of interest. Note. HC = healthy controls. Comb = Combined. In = Inattentive.

1.2.2.3. Multiverse analysis results

To evaluate the robustness of effects in the validation sample, we conducted a multiverse analysis using the same regression model:

TBR ~ Group * Age * Gender * IAF

Group was coded identically as in the HBN sample, with HC as the reference category, and ADHD-Inattentive and ADHD-Combined as the comparisons groups. Gender was coded with males as the reference group. Supplementary Table 10 summarizes the number of significant effects for each term across the 144 model specifications. The multiverse analysis in the validation sample included 144 rather than 576 specifications, as information on comorbid diagnoses and medication status was not available (144 * 2 * 2 = 576 in the HBN sample).

Supplementary Table 10. Results of the multiverse analysis in the validation sample showing the proportion of significant effects across 144 universes.

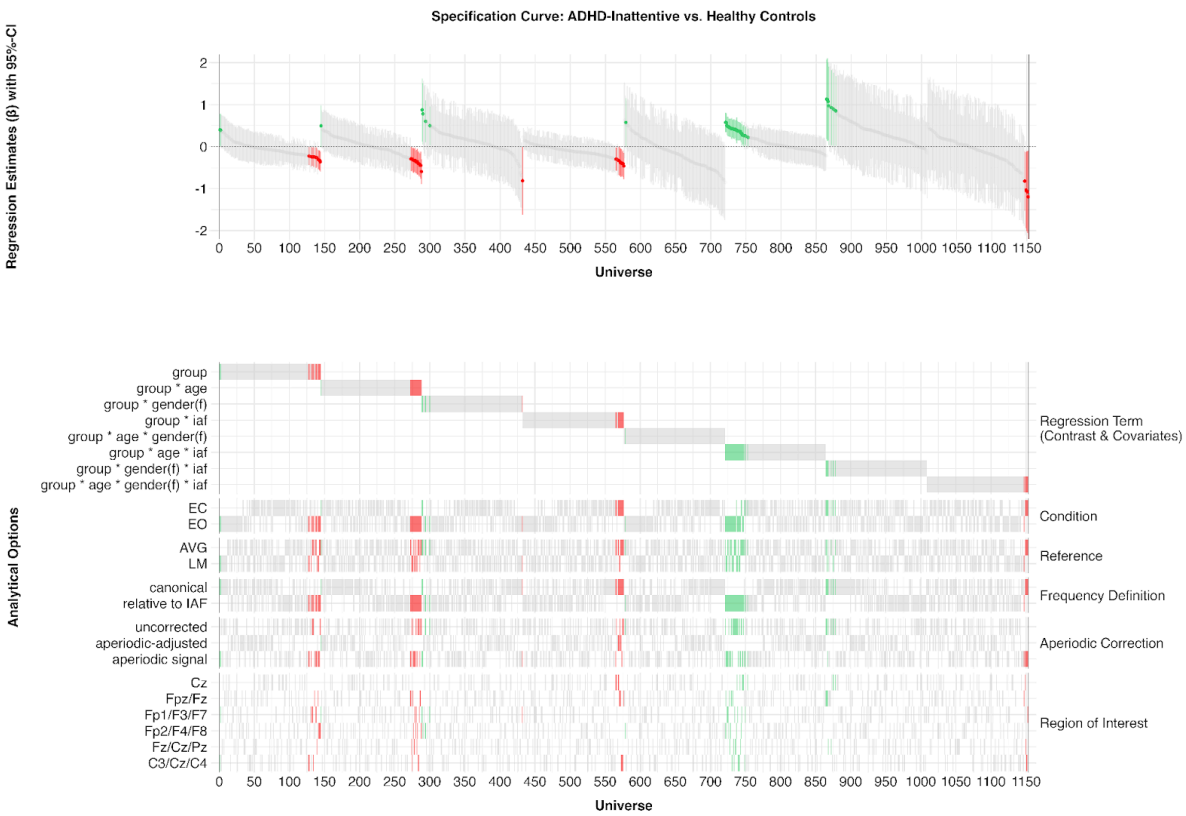
	Significantly Positive	***Significantly Negative***	***Not Significant***	***Total***
ADHD[IN]	2 (1.39 %)	11 (7.64 %)	131 (91.00 %)	144
ADHD[IN] * Age	1 (0.69 %)	16 (11.10 %)	127 (88.20 %)	144
ADHD[IN] * Gender[Female]	4 (2.78 %)	1 (0.69 %)	139 (96.50 %)	144
ADHD[IN] * IAF	0 (0.00 %)	10 (6.94 %)	134 (94.10 %)	144
ADHD[IN] * Age * Gender[Female]	1 (0.69 %)	0 (0.00 %)	143 (99.30 %)	144
ADHD[IN] * Age * IAF	30 (20.80 %)	0 (0.00 %)	114 (79.20 %)	144
ADHD[IN] * Gender[Female] * IAF	7 (4.86 %)	0 (0.00 %)	137 (95.10 %)	144
ADHD[IN] * Age * Gender[Female] * IAF	0 (0.87 %)	5 (3.47 %)	139 (96.50 %)	144

ADHD[Comb]	1 (0.69 %)	2 (1.39 %)	141 (97.90 %)	144
ADHD[Comb] * Age	4 (2.78 %)	19 (13.20 %)	121 (84.00 %)	144
ADHD[Comb] * Gender[Female]	0 (0.00 %)	0 (0.00 %)	144 (100.00 %)	144
ADHD[Comb] * IAF	0 (0.00 %)	1 (0.69 %)	143 (99.30 %)	144
ADHD[Comb] * Age * Gender[Female]	0 (0.00 %)	8 (5.56 %)	136 (94.40 %)	144
ADHD[Comb] * Age * IAF	14 (9.72 %)	6 (4.17 %)	124 (86.10 %)	144
ADHD[Comb] * Gender[Female] * IAF	2 (1.39 %)	0 (0.00 %)	142 (98.60 %)	144
ADHD[Comb] * Age * Gender[Female] * IAF	0 (0.00 %)	27 (18.80 %)	117 (81.20 %)	144

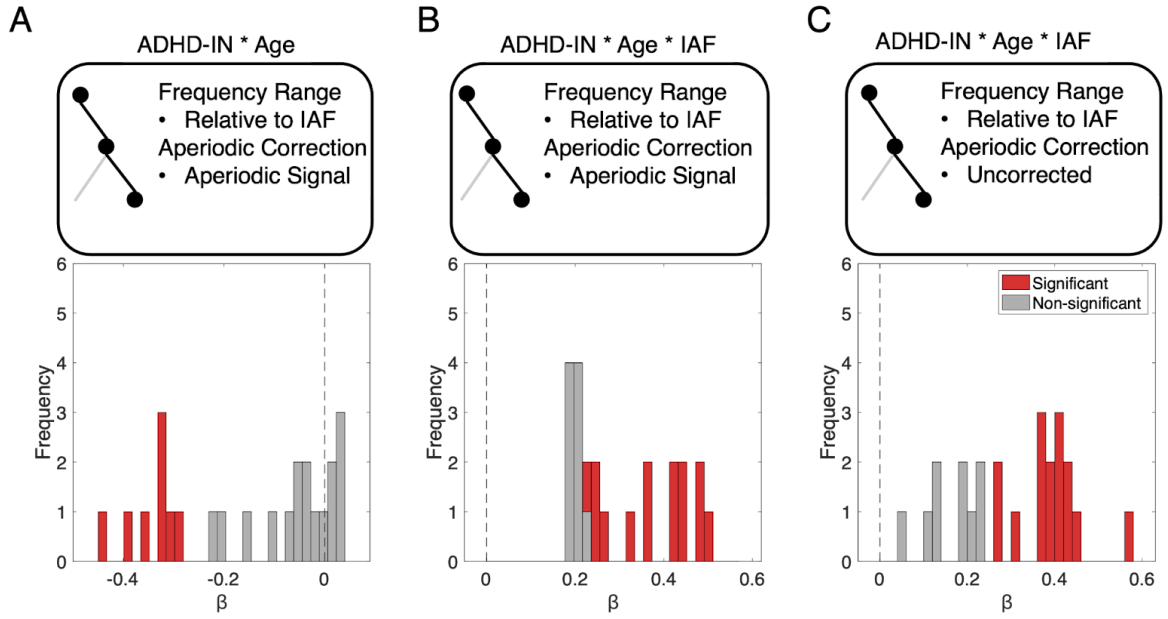
Note. Positive universe = higher TBR in ADHD compared to HC. Negative universe = lower TBR in ADHD compared to HC. IN = Inattentive. Comb = Combined. IAF = Individual alpha frequency.

1.2.2.4. Healthy Controls vs. ADHD-Inattentive Comparison

The results largely mirrored the HBN sample (Supplementary Figure 9). No consistent main effect of ADHD-Inattentive was observed, with only 2 positive and 11 negative universes out of 144 possibilities. Out of those 11 negative universes, 8 were observed when IAF-related frequency band and aperiodic signal. Next, the ADHD-Inattentive * Age interaction resulted in 16 significantly negative universes, which were again observed when using IAF-related frequency band and aperiodic signal (8 out of 24 possible universes; Supplementary Figure 10 A) and 1/f uncorrected power (8 out of 24 possible universes). Finally, the ADHD-Inattentive * Age * IAF interaction was significantly positive in 30 universes, especially when frequency bands were defined relative to IAF and when using aperiodic signal (15 out of 24 possible universes; Supplementary Figure 10 B) and 1/f uncorrected power (15 out of 24 possible universes; Supplementary Figure 10 C).



Supplementary Figure 9. Specification curve representing all universes for the healthy control vs. ADHD-Inattentive comparison in the validation sample. The top panel shows regression estimates sorted by effect size with 95% confidence intervals. Statistically significant positive estimates are shown in green, negative in red, and non-significant in gray. The bottom panel maps the analytical choices that produced each estimate. Note. CI = confidence interval. f = female. EC = eyes closed. EO = eyes open. AVG = average reference. LM = linked mastoid reference. canonical = canonical frequency range. relative to IAF = bandwidth relative to individual alpha frequency.

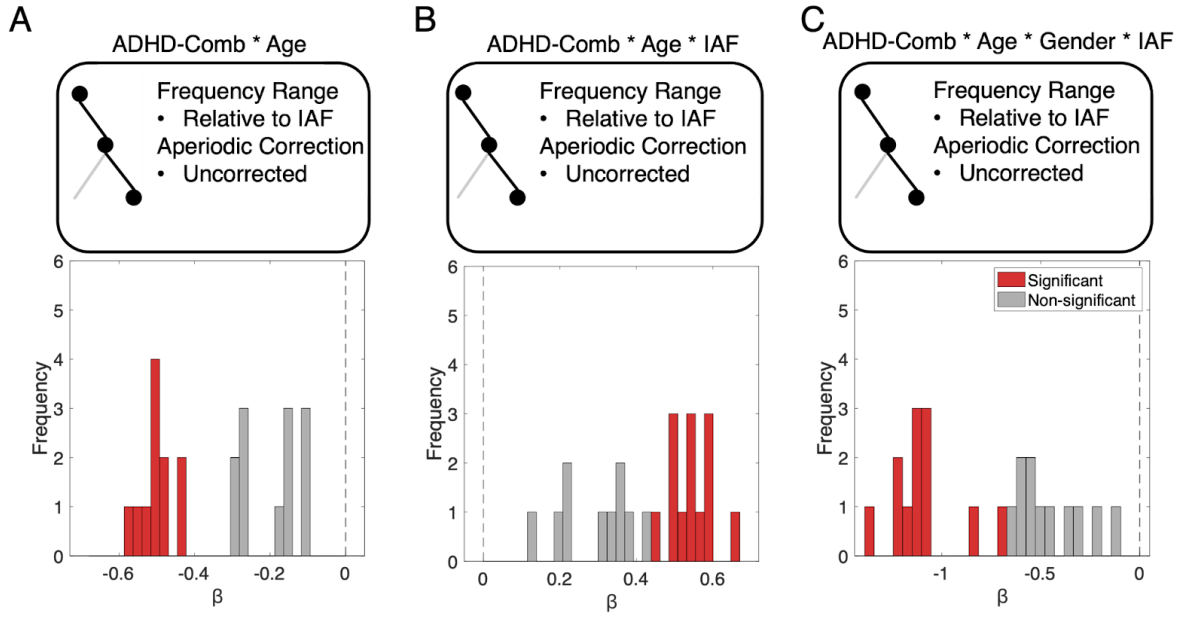


Supplementary Figure 10. Possibility space for the ADHD-Inattentive * Age and ADHD-Inattentive * Age * IAF interactions. (A) Negative ADHD-Inattentive * Age effect: 8 out of 24 significant universes with IAF-relative bands and aperiodic signal. (B) Positive ADHD-Inattentive * Age * IAF effect: 15/24 significant under aperiodic signal with IAF-relative bands (C) Positive ADHD-Inattentive * Age * IAF effect: 15/24 significant with IAF-relative bands and uncorrected power. Note. Significant beta estimates are highlighted in red and non-significant estimates are shown in grey.

1.2.2.5. Healthy Controls vs. ADHD-Combined Comparison

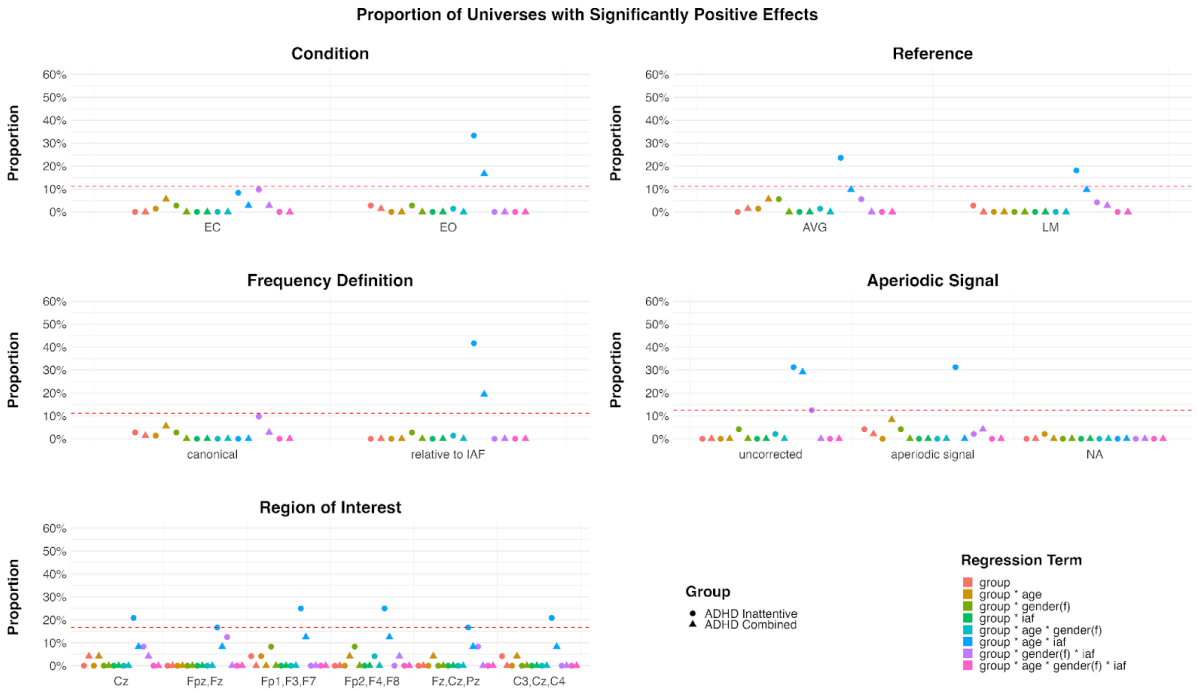
Lastly, we computed the contrast between the HC and ADHD-Combined groups. The specification curve for this analysis is shown in Supplementary Figure 11. In the HC vs. ADHD-Combined contrast, we observed an ADHD-Combined * Age interaction that was significantly negative in 19 universes, specifically in 12 out of 24 specifications involving IAF-related bands and the 1/f uncorrected signal (Supplementary Figure 12 A). Furthermore, the ADHD-Combined * Age * IAF interaction was significantly positive in 14 universes, all of which were based on IAF-related band definitions and the 1/f uncorrected signal (Supplementary Figure 12 B). Finally, the ADHD-Combined * Age * Gender * IAF interaction was significantly negative in 27 out of 144 universes, with 13 of these effects arising from specifications using IAF-related bands and the 1/f-uncorrected signal (Supplementary Figure 12 C).

Supplementary Figure 11. Specification curve representing all universes for the healthy control vs. ADHD-Combined comparison in the validation sample. The top panel shows regression estimates sorted by effect size with 95% confidence intervals. Statistically significant positive estimates are shown in green, negative in red, and non-significant in gray. The bottom panel maps the analytical choices that produced each estimate. Note. CI = confidence interval. f = female. EC = eyes closed. EO = eyes open. AVG = average reference. LM = linked mastoid reference. canonical = canonical frequency range. relative to IAF = bandwidth relative to individual alpha frequency.

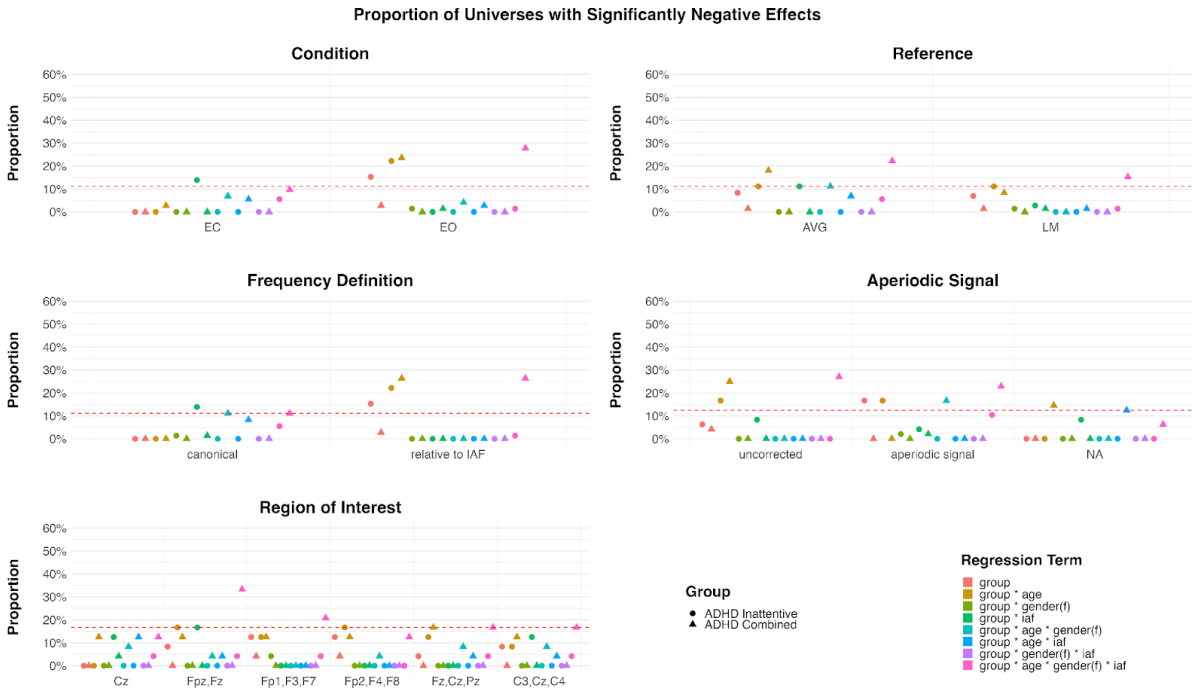


Supplementary Figure 12. Possibility space for the ADHD-Combined * Age, ADHD-Combined * Age * IAF and ADHD-Combined * Age * Gender * IAF interactions. (A) Negative ADHD-Combined * Age effect: 12 out of 24 significant universes with IAF-relative bands and uncorrected power. (B) Positive ADHD-Combined * Age * IAF effect: 14/24 significant under aperiodic signal with IAF-relative bands (C) Positive ADHD-Combined * Age *Gender * IAF effect: 13/24 significant with IAF-relative bands and uncorrected power. Note. Significant beta estimates are highlighted in red and non-significant estimates are shown in grey.

1.2.2.6. Proportion Plots Across All Analytical Specifications



Supplementary Figure 13. Proportion of specifications showing a significantly positive effect by regression term and ADHD subtype across all analytical specifications. Each level of Condition, Reference, Frequency Definition includes 72 analytical universes. The Aperiodic Signal factor includes 48 universes per level, and the Region of Interest includes 24 universes per level. The red line in the plot indicates the exact proportion of significant results required for the binomial test to reach statistical significance ($p < .05$) for each subset size.



Supplementary Figure 14. Proportion of specifications showing a significantly negative effect by regression term and ADHD subtype across all analytical specifications. Each level of Condition, Reference, Frequency Definition includes 72 analytical universes. The Aperiodic Signal factor includes 48 universes per level, and the Region of Interest includes 24 universes per level. The red line in the plot indicates the exact proportion of significant results required for the binomial test to reach statistical significance ($p < .05$) for each subset size.

References

1. Sinishaw, Y. D., Dachew, B. A., Ayano, G., Betts, K. & Alati, R. The association between maternal diabetes and the risk of attention deficit/hyperactivity disorder in offspring: Updated systematic review and meta-analysis. *Eur. Psychiatry* **67**, S139–S139 (2024).
2. Boxum, M. *et al.* Challenging the diagnostic value of theta/beta ratio: Insights from an EEG subtyping meta-analytical approach in ADHD. *Appl. Psychophysiol. Biofeedback* (2024) doi:10.1007/s10484-024-09649-y.
3. Polanczyk, G., de Lima, M. S., Horta, B. L., Biederman, J. & Rohde, L. A. The worldwide prevalence of ADHD: a systematic review and metaregression analysis. *Am. J. Psychiatry* **164**, 942–948 (2007).
4. Faraone, S. V., Biederman, J. & Mick, E. The age-dependent decline of attention deficit hyperactivity disorder: a meta-analysis of follow-up studies. *Psychol. Med.* **36**, 159–165 (2006).
5. Arns, M., Conners, C. K. & Kraemer, H. C. A decade of EEG Theta/Beta Ratio Research in ADHD: a meta-analysis. *J. Atten. Disord.* **17**, 374–383 (2013).
6. Young, S., Bramham, J., Gray, K. & Rose, E. The experience of receiving a diagnosis and treatment of ADHD in adulthood: a qualitative study of clinically referred patients using interpretative phenomenological analysis. *J. Atten. Disord.* **11**, 493–503 (2008).
7. Lubar, J. F. Discourse on the development of EEG diagnostics and biofeedback for attention-deficit/hyperactivity disorders. *Biofeedback Self. Regul.* **16**, 201–225 (1991).
8. Bresnahan, S. M., Anderson, J. W. & Barry, R. J. Age-related changes in quantitative EEG in attention-deficit/hyperactivity disorder. *Biol. Psychiatry* **46**, 1690–1697 (1999).
9. Callaway, E., Halliday, R. & Naylor, H. Hyperactive children's event-related potentials fail to support underarousal and maturational-lag theories. *Arch. Gen. Psychiatry* **40**, 1243–1248 (1983).
10. Matsuura, M. *et al.* A cross-national EEG study of children with emotional and behavioral problems: a WHO collaborative study in the Western Pacific Region. *Biol.*

Psychiatry **34**, 59–65 (1993).

11. Mann, C. A., Lubar, J. F., Zimmerman, A. W., Miller, C. A. & Muenchen, R. A. Quantitative analysis of EEG in boys with attention-deficit-hyperactivity disorder: Controlled study with clinical implications. *Pediatr. Neurol.* **8**, 30–36 (1992).
12. Loo, S. K. & Arns, M. Should the EEG–based theta to beta ratio be used to diagnose ADHD? *ADHD Rep.* **23**, 8–13 (2015).
13. Monastra, V. J. *et al.* Assessing attention deficit hyperactivity disorder via quantitative electroencephalography: An initial validation study. *Neuropsychology* **13**, 424–433 (1999).
14. Monastra, V. J., Lubar, J. F. & Linden, M. The development of a quantitative electroencephalographic scanning process for attention deficit-hyperactivity disorder: reliability and validity studies. *Neuropsychology* **15**, 136–144 (2001).
15. Arns, M. *et al.* Editorial Perspective: How should child psychologists and psychiatrists interpret FDA device approval? Caveat emptor. *J. Child Psychol. Psychiatry* **57**, 656–658 (2016).
16. Dolgin, E. FDA clearance paves way for computerized ADHD monitoring. *Nat. Med.* **20**, 454–455 (2014).
17. Snyder, S. M., Rugino, T. A., Hornig, M. & Stein, M. A. Integration of an EEG biomarker with a clinician's ADHD evaluation. *Brain Behav.* **5**, e00330 (2015).
18. Snyder, S. M. *et al.* Blinded, multi-center validation of EEG and rating scales in identifying ADHD within a clinical sample. *Psychiatry Res.* **159**, 346–358 (2008).
19. van Dijk, H. *et al.* Different spectral analysis methods for the theta/Beta Ratio calculate different ratios but do not distinguish ADHD from controls. *Appl. Psychophysiol. Biofeedback* **45**, 165–173 (2020).
20. Lenartowicz, A. & Loo, S. K. Use of EEG to diagnose ADHD. *Curr. Psychiatry Rep.* **16**, 498 (2014).
21. Brysbaert, M. How many participants do we have to include in properly powered experiments? A tutorial of power analysis with reference tables. *J. Cogn.* **2**, 16 (2019).

22. Ioannidis, J. P. A. Why most published research findings are false. *Chance (N. Y.)* **18**, 40–47 (2005).
23. Schmidt, F. *et al.* Age-related changes in ‘cortical’ 1/f dynamics are linked to cardiac activity. (2025) doi:10.7554/elife.100605.2.
24. Donoghue, T. *et al.* Parameterizing neural power spectra into periodic and aperiodic components. *Nat. Neurosci.* **23**, 1655–1665 (2020).
25. Dziemian, S., Barańczuk-Turska, Z. & Langer, N. Association between attention-deficit/hyperactivity disorder symptom severity and white matter integrity moderated by in-scanner head motion. *Transl. Psychiatry* **12**, 434 (2022).
26. Kerson, C. *et al.* EEG theta/beta ratio calculations differ between various EEG neurofeedback and assessment software packages: Clinical interpretation. *Clin. EEG Neurosci.* **51**, 114–120 (2020).
27. Simmons, J. P., Nelson, L. D. & Simonsohn, U. False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychol. Sci.* **22**, 1359–1366 (2011).
28. Götz, M., Sarma, A. & O’Boyle, E. H. The multiverse of universes: A tutorial to plan, execute and interpret multiverses analyses using the R package multiverse. *Int. J. Psychol.* **59**, 1003–1014 (2024).
29. Sarma, A., Hwang, K., Hullman, J. & Kay, M. Milliways: Taming multiverses through principled evaluation of data analysis paths. in *Proceedings of the CHI Conference on Human Factors in Computing Systems* 1–15 (ACM, New York, NY, USA, 2024). doi:10.1145/3613904.3642375.
30. Sarma, A., Hedayati, M. & Kay, M. More forecasts, more (decision) problems: How uncertainty representations for multiple forecasts impact decision making. in *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* 1–14 (ACM, New York, NY, USA, 2025). doi:10.1145/3706598.3713725.
31. Del Giudice, M. & Gangestad, S. W. A traveler’s guide to the multiverse: Promises, pitfalls, and a framework for the evaluation of analytic decisions. *Adv. Methods Pract.*

Psychol. Sci. **4**, 251524592095492 (2021).

32. Steegen, S., Tuerlinckx, F., Gelman, A. & Vanpaemel, W. Increasing transparency through a multiverse analysis. *Perspect. Psychol. Sci.* **11**, 702–712 (2016).
33. Alexander, L. M. *et al.* An open resource for transdiagnostic research in pediatric mental health and learning disorders. *Sci Data* **4**, 170181 (2017).
34. Langer, N. *et al.* A resource for assessing information processing in the developing brain using EEG and eye tracking. *Sci Data* **4**, 170040 (2017).
35. Münger, M. *et al.* Behavioral and neurophysiological markers of ADHD in children, adolescents, and adults: A large-scale clinical study. *Clin. EEG Neurosci.* **52**, 311–320 (2021).
36. Münger, M. *et al.* Longitudinal analysis of self-reported symptoms, behavioral measures, and event-related potential components of a cued Go/NoGo task in adults with attention-deficit/hyperactivity disorder and controls. *Front. Hum. Neurosci.* **16**, 767789 (2022).
37. Müller, A. *et al.* EEG/ERP-based biomarker/neuroalgorithms in adults with ADHD: Development, reliability, and application in clinical practice. *World J. Biol. Psychiatry* **21**, 172–182 (2020).
38. Chambers, W., Puig-Antich, J. & Hirsch, M. *The Assessment of Affective Disorders in Children and Adolescents by Semi-Structured Interview: Test-Retest Reliability of the Schedule for Affective Disorders and Schizophrenia for School*. (1985).
39. Achenbach, T. M. & Dumenci, L. Advances in empirically based assessment: Revised cross-informant syndromes and new DSM-oriented scales for the CBCL, YSR, and TRF: Comment on Lengua, Sadowski, Friedrich, and Fisher (2001). *J. Consult. Clin. Psychol.* **69**, 699–702 (2001).
40. Swanson, J. *et al. Over- Identification of Extreme Behavior in Evaluation and Diagnosis of ADHD/HKD*. (2001).
41. Conners, K. C. Conners 3rd edition manual. New York: Multi-Health Systems. *Inc.[Google Scholar]* (2008).

42. Pedroni, A., Bahreini, A. & Langer, N. Automagic: Standardized preprocessing of big EEG data. *Neuroimage* **200**, 460–473 (2019).
43. Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K.-M. & Robbins, K. A. The PREP pipeline: standardized preprocessing for large-scale EEG analysis. *Front. Neuroinform.* **9**, 16 (2015).
44. Widmann, A. & Schröger, E. Filter effects and filter artifacts in the analysis of electrophysiological data. *Front. Psychol.* **3**, 233 (2012).
45. de Cheveigné, A. ZapLine: A simple and effective method to remove power line artifacts. *Neuroimage* **207**, 116356 (2020).
46. Pion-Tonachini, L., Kreutz-Delgado, K. & Makeig, S. ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *Neuroimage* **198**, 181–197 (2019).
47. Tröndle, M., Popov, T. & Langer, N. Decomposing the role of alpha oscillations during brain maturation. *bioRxiv* 2020.11.06.370882 (2020) doi:10.1101/2020.11.06.370882.
48. Tröndle, M. *et al.* Decomposing age effects in EEG alpha power. *bioRxiv* 2021.05.26.445765 (2021) doi:10.1101/2021.05.26.445765.
49. Oostenveld, R., Fries, P., Maris, E. & Schoffelen, J.-M. FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* **2011**, 156869 (2011).
50. Finley, A. J., Angus, D. J., van Reekum, C. M., Davidson, R. J. & Schaefer, S. M. Periodic and aperiodic contributions to theta-beta ratios across adulthood. *Psychophysiology* **59**, e14113 (2022).
51. Corcoran, A. W., Alday, P. M., Schlesewsky, M. & Bornkessel-Schlesewsky, I. Toward a reliable, automated method of individual alpha frequency (IAF) quantification. *Psychophysiology* **55**, e13064 (2018).
52. Tröndle, M., Popov, T., Dziemian, S. & Langer, N. Decomposing the role of alpha oscillations during brain maturation. *ELife* **11**, e77571 (2022).
53. Klimesch, W. EEG alpha and theta oscillations reflect cognitive and memory

performance: a review and analysis. *Brain Res. Rev.* **29**, 169–195 (1999).

54. Tröndle, M., Popov, T., Dziemian, S. & Langer, N. Decomposing the role of alpha oscillations during brain maturation. *Elife* **11**, (2022).
55. Babiloni, C. *et al.* International Federation of Clinical Neurophysiology (IFCN) - EEG research workgroup: Recommendations on frequency and topographic analysis of resting state EEG rhythms. Part 1: Applications in clinical research studies. *Clin. Neurophysiol.* **131**, 285–307 (2020).
56. Klimesch, W. A-band oscillations, attention, and controlled access to stored information. *Trends Cogn. Sci.* **16**, 606–617 (2012).
57. Arns, M. *et al.* Electroencephalographic biomarkers as predictors of methylphenidate response in attention-deficit/hyperactivity disorder. *Eur. Neuropsychopharmacol.* **28**, 881–891 (2018).
58. Loo, S. K. *et al.* Characterization of the theta to beta ratio in ADHD: identifying potential sources of heterogeneity. *J. Atten. Disord.* **17**, 384–392 (2013).
59. Ahmadi, M. *et al.* Cortical source analysis of resting state EEG data in children with attention deficit hyperactivity disorder. *Clin. Neurophysiol.* **131**, 2115–2130 (2020).
60. Aldemir, R. *et al.* Investigation of attention deficit hyperactivity disorder (ADHD) sub-types in children via EEG frequency domain analysis. *Int. J. Neurosci.* **128**, 349–360 (2018).
61. Barry, R. J., Clarke, A. R., Johnstone, S. J., McCarthy, R. & Selikowitz, M. Electroencephalogram theta/beta ratio and arousal in attention-deficit/hyperactivity disorder: evidence of independent processes. *Biol. Psychiatry* **66**, 398–401 (2009).
62. Buyck, I. & Wiersema, J. R. Resting electroencephalogram in attention deficit hyperactivity disorder: developmental course and diagnostic value. *Psychiatry Res.* **216**, 391–397 (2014).
63. Buyck, I. & Wiersema, J. R. State-related electroencephalographic deviances in attention deficit hyperactivity disorder. *Res. Dev. Disabil.* **35**, 3217–3225 (2014).
64. Buyck, I. & Wiersema, J. R. Electroencephalographic activity before and after cognitive

effort in children with attention deficit/hyperactivity disorder. *Clin. EEG Neurosci.* **46**, 88–93 (2015).

65. Chow, J. C. *et al.* Entropy-based quantitative electroencephalogram analysis for diagnosing attention-deficit hyperactivity disorder in girls. *Clin. EEG Neurosci.* **50**, 172–179 (2019).
66. Clarke, A. R. *et al.* Childhood EEG as a predictor of adult attention-deficit/hyperactivity disorder. *Clin. Neurophysiol.* **122**, 73–80 (2011).
67. Clarke, A. R., Barry, R. J. & Johnstone, S. Resting state EEG power research in Attention-Deficit/Hyperactivity Disorder: A review update. *Clin. Neurophysiol.* **131**, 1463–1479 (2020).
68. Clarke, A. R., Barry, R. J., McCarthy, R. & Selikowitz, M. Age and sex effects in the EEG: differences in two subtypes of attention-deficit/hyperactivity disorder. *Clin. Neurophysiol.* **112**, 815–826 (2001).
69. Clarke, A. R., Barry, R. J., McCarthy, R. & Selikowitz, M. EEG-defined subtypes of children with attention-deficit/hyperactivity disorder. *Clin. Neurophysiol.* **112**, 2098–2105 (2001).
70. Clarke, A. R., Barry, R. J., McCarthy, R. & Selikowitz, M. EEG analysis of children with attention-deficit/hyperactivity disorder and comorbid reading disabilities. *J. Learn. Disabil.* **35**, 276–285 (2002).
71. Dupuy, F. E., Barry, R. J., Clarke, A. R., McCarthy, R. & Selikowitz, M. Sex differences between the combined and inattentive types of attention-deficit/hyperactivity disorder: an EEG perspective. *Int. J. Psychophysiol.* **89**, 320–327 (2013).
72. Dupuy, F. E., Clarke, A. R., Barry, R. J., McCarthy, R. & Selikowitz, M. EEG differences between the Combined and Inattentive types of attention-deficit/hyperactivity disorder in girls: A further investigation. *Clin. EEG Neurosci.* **45**, 231–237 (2014).
73. González-Castro, P., Rodríguez, C., López, Á., Cueli, M. & Álvarez, L. Attention Deficit Hyperactivity Disorder, differential diagnosis with blood oxygenation, beta/theta ratio, and attention measures. *Int. J. Clin. Health Psychol.* **13**, 101–109 (2013).

74. Halawa, I. F., El Sayed, B. B., Amin, O. R., Meguid, N. A. & Abdel Kader, A. A. Frontal theta/beta ratio changes during TOVA in Egyptian ADHD children. *Neurosciences (Riyadh)* **22**, 287–291 (2017).

75. Heinrich, H. *et al.* EEG spectral analysis of attention in ADHD: implications for neurofeedback training? *Front. Hum. Neurosci.* **8**, 611 (2014).

76. Hermens, D. F., Kohn, M. R., Clarke, S. D., Gordon, E. & Williams, L. M. Sex differences in adolescent ADHD: findings from concurrent EEG and EDA. *Clin. Neurophysiol.* **116**, 1455–1463 (2005).

77. Hobbs, M. J., Clarke, A. R., Barry, R. J., McCarthy, R. & Selikowitz, M. EEG abnormalities in adolescent males with AD/HD. *Clin. Neurophysiol.* **118**, 363–371 (2007).

78. Huang, C.-J. *et al.* Effects of acute exercise on resting EEG in children with attention-deficit/hyperactivity disorder. *Child Psychiatry Hum. Dev.* **49**, 993–1002 (2018).

79. Janzen, T., Graap, K., Stephanson, S., Marshall, W. & Fitzsimmons, G. Differences in baseline EEG measures for ADD and normally achieving preadolescent males. *Biofeedback Self. Regul.* **20**, 65–82 (1995).

80. Jarrett, M. A. *et al.* An EEG study of children with and without ADHD symptoms: Between-group differences and associations with sluggish cognitive tempo symptoms. *J. Atten. Disord.* **24**, 1002–1010 (2020).

81. Kitsune, G. L. *et al.* A matter of time: The influence of recording context on EEG spectral power in adolescents and young adults with ADHD. *Brain Topogr.* **28**, 580–590 (2015).

82. Lansbergen, M. M., Arns, M., van Dongen-Boomsma, M., Spronk, D. & Buitelaar, J. K. The increase in theta/beta ratio on resting-state EEG in boys with attention-deficit/hyperactivity disorder is mediated by slow alpha peak frequency. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* **35**, 47–52 (2011).

83. Liechti, M. D. *et al.* Diagnostic value of resting electroencephalogram in attention-deficit/hyperactivity disorder across the lifespan. *Brain Topogr.* **26**, 135–151 (2013).

84. Luo, N. *et al.* Aberrant brain dynamics and spectral power in children with ADHD and its subtypes. *Eur. Child Adolesc. Psychiatry* **32**, 2223–2234 (2023).

85. Markovska-Simoska, S. & Pop-Jordanova, N. Quantitative EEG in children and adults with attention deficit hyperactivity disorder: Comparison of absolute and relative power spectra and theta/beta ratio. *Clin. EEG Neurosci.* **48**, 20–32 (2017).

86. Martín-Brufau, R. & Nombela-Gómez, M. Bioelectrical markers of ADHD: enhancement of direct EEG analysis. *Rev. Electron. Investig. Psicoeduc. Psigopedag.* **15**, 185–200 (2017).

87. Nazari, M. A., Wallois, F., Aarabi, A. & Berquin, P. Dynamic changes in quantitative electroencephalogram during continuous performance test in children with attention-deficit/hyperactivity disorder. *Int. J. Psychophysiol.* **81**, 230–236 (2011).

88. Ogrim, G., Kropotov, J. & Hestad, K. The quantitative EEG theta/beta ratio in attention deficit/hyperactivity disorder and normal controls: sensitivity, specificity, and behavioral correlates. *Psychiatry Res.* **198**, 482–488 (2012).

89. Poil, S.-S. *et al.* Age dependent electroencephalographic changes in attention-deficit/hyperactivity disorder (ADHD). *Clin. Neurophysiol.* **125**, 1626–1638 (2014).

90. Rezaeezadeh, M., Shamekhi, S. & Shamsi, M. Attention Deficit Hyperactivity Disorder Diagnosis using non-linear univariate and multivariate EEG measurements: a preliminary study. *Phys. Eng. Sci. Med.* **43**, 577–592 (2020).

91. Robertson, M. M. *et al.* EEG power spectral slope differs by ADHD status and stimulant medication exposure in early childhood. *J. Neurophysiol.* **122**, 2427–2437 (2019).

92. Sangal, R. B. & Sangal, J. M. Use of EEG beta-1 power and theta/beta ratio over Broca's area to confirm diagnosis of attention deficit/hyperactivity disorder in children. *Clin. EEG Neurosci.* **46**, 177–182 (2015).

93. Shi, T. *et al.* EEG characteristics and visual cognitive function of children with attention deficit hyperactivity disorder (ADHD). *Brain Dev.* **34**, 806–811 (2012).

94. Skalski, S., Pochwatko, G. & Balas, R. Impact of motivation on selected aspects of

attention in children with ADHD. *Child Psychiatry Hum. Dev.* **52**, 586–595 (2021).

95. Woltering, S., Jung, J., Liu, Z. & Tannock, R. Resting state EEG oscillatory power differences in ADHD college students and their peers. *Behav. Brain Funct.* **8**, 60 (2012).
96. Zhang, D.-W. *et al.* Time effects on resting EEG in children with/without AD/HD. *Brain Topogr.* **32**, 286–294 (2019).
97. Zhang, D.-W. *et al.* Electroencephalogram theta/beta ratio and spectral power correlates of executive functions in children and adolescents with AD/HD. *J. Atten. Disord.* **23**, 721–732 (2019).
98. Wilkinson, G. N. & Rogers, C. E. Symbolic description of factorial models for analysis of variance. *J. R. Stat. Soc. Ser. C Appl. Stat.* **22**, 392 (1973).
99. Simonsohn, U., Simmons, J. P. & Nelson, L. D. Specification curve: Descriptive and inferential statistics on all reasonable specifications. *SSRN Electron. J.* (2015) doi:10.2139/ssrn.2694998.
100. Simonsohn, U., Nelson, L. D. & Simmons, J. P. P-curve won't do your laundry, but it will distinguish replicable from non-replicable findings in observational research: Comment on Bruns & Ioannidis (2016). *PLoS One* **14**, e0213454 (2019).
101. Fernandes, M., Walls, L., Munson, S., Hullman, J. & Kay, M. Uncertainty displays using quantile dotplots or CDFs improve transit decision-making. in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* 1–12 (ACM, New York, NY, USA, 2018). doi:10.1145/3173574.3173718.
102. Yang, F., Mortenson, C., Nisbet, E., Diakopoulos, N. & Kay, M. In dice we trust: Uncertainty displays for maintaining trust in election forecasts over time. (2024) doi:10.31219/osf.io/9x4nr.
103. Gao, R., Peterson, E. J. & Voytek, B. Inferring synaptic excitation/inhibition balance from field potentials. *Neuroimage* **158**, 70–78 (2017).
104. Tröndle, M. & Langer, N. Effects of artifacts on aperiodic component in EEG. *in press* (2026).
105. Voetterl, H. *et al.* Brainmarker-I differentially predicts remission to various

attention-deficit/hyperactivity disorder treatments: A discovery, transfer, and blinded validation study. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **8**, 52–60 (2023).

106. Krepel, N. *et al.* A multicenter effectiveness trial of QEEG-informed neurofeedback in ADHD: Replication and treatment prediction. *NeuroImage Clin.* **28**, 102399 (2020).

107. Wittchen, H.-U., Zaudig, M. & Fydrich, T. SKID. Strukturiertes Klinisches Interview für DSM-IV. Achse I und II. Handanweisung. (1997).

108. Weiß. *Wiener Matrizen-Test 2: Ein Rasch-Skaldierter Sprachfreier Kurztest Zu Erfassung Der Intelligenz*. (2011).

109. Forman, Poswanger & Waldherr. *Grundintelligenztest Skala 2 (CFT 20-R) Mit Wortschatztest (WS) Und Zahlenfolgentest (ZF)*. (2006).