NEW RESEARCH

Check for updates

Saliency Models Reveal Reduced Top-Down Attention in Attention-Deficit/Hyperactivity Disorder: A Naturalistic Eye-Tracking Study

Sabine Dziemian, PhD, Gaston Bujia, MSc, Paul Prasse, PhD, Zofia Barańczuk-Turska, PhD, Lena A. Jäger, PhD, Juan E. Kamienkowski, PhD, Nicolas Langer, PhD

Objective: Attention-deficit/hyperactivity disorder (ADHD) is a highly prevalent neurodevelopmental disorder. It is associated with deficits in executive functions, especially in visual attention. Deviant visual attention in ADHD is suspected to arise from imbalances between top-down and bottom-up mechanisms. However, it is unclear which of these mechanisms propels the aberrant visual attention.

Method: In 815 medication-naïve children and adolescents (age range 5-21 years), differences in visual attention in participants with ADHD and neurotypical controls were investigated using eye tracking in a naturalistic video viewing task. Two opposing saliency models were used. Finegrained, based on low-level image features, was chosen to estimate bottom-up visually relevant areas. ViNet, a higher-level saliency model based on deep neural networks and trained on the gaze of neurotypical controls, was selected to determine top-down visually relevant regions. Correspondence between gaze and both saliency maps was calculated using normalized scanpath saliency, thus measuring the extent of coherence to bottom-up and top-down relevant contents.

Results: Participants with combined ADHD showed lower mean normalized scanpath saliency for the top-down saliency map, but not the bottom-up one, compared with neurotypical controls. This contrast indicates poorer top-down control as a major contributor to impaired visual attention in combined ADHD. There was no significant effect for the predominantly inattentive ADHD group.

Conclusion: This study demonstrated the use of eye tracking for differentiating between top-down and bottom-up visual attention. It shows that in combined ADHD, a reduction of top-down visual attention is key to an impaired competition between bottom-up and top-down visual attention.

Plain language summary: This study used eye-tracking to investigate visual attention differences during naturalistic video viewing among 815 medication-naïve children and adolescents with attention-deficit/hyperactivity disorder (ADHD) and neurotypical controls. Computational models were used to quantify bottom-up and top-down mechanisms of visual attention to video content. The study found that individuals with ADHD had reduced control over top-down visual attention, suggesting an imbalance in visual attention mechanisms that may be targeted during intervention.

Key words: ADHD; eye-tracking; naturalistic task-free viewing; saliency models; visual attention

JAACAP Open 2025;3(2):192-204.

Attention-deficit/hyperactivity disorder (ADHD) is a neurodevelopmental disorder that is subdivided based on symptoms into a predominantly hyperactive-impulsive presentation, a predominantly inattentive presentation (ADHD-IN), and a combined presentation (ADHD-C).[1] With a prevalence estimate of approximately 5% based on population surveys, ADHD is a very common neurodevelopmental disorder.[1] ADHD has also been associated with lifelong psychological and social consequences that severely impact quality of life and achievement.[1,2]

To date, the exact pathophysiology of ADHD and its presentations is poorly understood, and no biological marker or distinctive genetic factor has yet been identified to aid diagnosis.[1,3] The lack of an objective biomedical indicator for ADHD constrains early diagnosis and limits the potential of treatments and interventions.[1,3] Despite this lack, the symptoms of ADHD have been argued to arise primarily because of deficits in executive functions.[4–6] These deficits manifest in difficulties in focusing and maintaining attention, particularly in the presence of distractors. Likewise, children and adolescents with ADHD show worse abilities in suppressing inappropriate behavioral responses. Consequently, the domain of inhibitory control is also of particular interest in ADHD.[4]

Although research has tried to address these domains with functional magnetic resonance imaging[7] and electroencephalography,[8] these methods come with drawbacks; magnetic resonance imaging entails large financial costs, and consequently its suitability is limited and it is not widely available for clinical diagnosis.[9] Electroencephalography requires time-consuming preparation and thus is less likely to be applicable in clinics on a daily basis.[10]

By contrast, eye movement tracking provides an unobtrusive and easily accessible means for studying higher-order processes such as attentional and inhibitory control.[11,12] Eye tracking is used to identify shifts in overt attention through reallocation of gaze.[13] Previous investigations of eye movements in ADHD mainly used tasks specially designed for probing attention and inhibition.[14] Under typical task instruction, inattention manifests in the form of an inability to maintain fixation (ie, periods in which the gaze remains relatively still) on given stimuli.[15] Previous studies have confirmed poorer sustained attention in children and adolescents with ADHD failing to maintain fixation during a fixation task.[16–19] In comparison, impairments in inhibitory control appear as reflexive saccades (ie, fast ballistic eye movements to reallocate the fovea on an area of interest) toward task-irrelevant stimuli.[15] Several studies have identified decreased inhibitory performance in children and adolescents with ADHD in an antisaccade task.[18,20–23] Poorer performance in ADHD has also been shown across other eye-tracking tasks targeting visual attention and inhibition.[14,21,24] In summary, previous conventional task-based eye-tracking studies have confirmed that children and adolescents with ADHD have difficulties in sustained attention and inhibition evident through higher occurrences of task-inappropriate eye movements.

Although such task-based studies are key to guiding future investigations of ADHD, performance under task instruction bears the risk of masking differences in eye movements that are present in natural and spontaneous gaze behavior.[25] Task-free viewing studies have a higher practicability and ecological validity and better resemble real-world scenarios, in which symptoms of ADHD mostly occur because they allow spontaneous unbiased eye movements.[12,25–27] Additionally, natural, dynamic video sequences that are complex and rich in semantic and salient content are most suitable in eliciting eye movements closer to gaze behavior in a natural context and clinical symptoms.[13,26–28]

Theoretical models of visual attention allocation vary depending on their emphasis on bottom-up or top-down processing, which are considered 2 opposite ends of a spectrum.[13,29,30] Bottom-up attention is driven by visually salient features, such as bright colors, high contrast, and sudden movements. Stimuli possessing these features are processed more quickly and efficiently than stimuli lacking them. Top-down attention, also known as endogenous attention, is goal driven and recruits prior knowledge and expectations like the individual's goals, task demands, motivation, and cognitive load to voluntarily guide attention.[31,32] This type of attention relies on executive functions, specifically inhibitory control and working memory, which help to suppress distractions and maintain focus on relevant information.[4–6,31]

Both bottom-up and top-down processes contribute to the allocation of visual attention in a flexible and dynamic manner, depending on the task demands and environmental context. For example, when searching for a red bike in a cluttered scene, bottom-up attention initially captures the gaze to visually salient red objects, while top-down attention guides the gaze to specific regions where a heavy object such as a bike is likely to be found (ie, on the ground). Top-down attention not only guides visual attention, but also modulates the perception of visually salient stimuli depending on whether they match the observer's expectations.[32,33] In ADHD, deficits in executive functions can lead to difficulties with top-down attention and may contribute to symptoms of inattention and distractibility. Individuals with ADHD may encounter difficulties in filtering out irrelevant stimuli, shifting attention between tasks, and sustaining attention over time.[4–6]

Eye movements and visual attention recruit overlapping networks of frontostriatal circuits, which are suggested to be involved in the pathophysiology of ADHD.[11,12,28,34–37] Several brain structures that are part of the frontostriatal network (FSN) are also involved in the mechanisms of top-down and bottom-up visual attention. Top-down signals, generated in dorsolateral prefrontal regions, frontal eye fields, and the lateral intraparietal area, reach the superior colliculus (SC) through the basal ganglia and the thalamus.[38,39] Likewise, bottom-up signals from primary visual cortical areas are projected to the SC. Both signals are integrated in a winner-takes-all mechanism in the SC, which triggers the respective eye movements.[37,38]

Saliency models are a powerful means to quantify the visual relevance of each pixel within a frame of a video and can be used to compare a person's gaze pattern with relevant regions within a frame.[40] Low-level saliency models are a suitable proxy for quantifying the bottom-up visual relevance of video content by evaluating basic image features such as color, contrast, and orientation.[41] In task-free eye movement studies, top-down instructions are not explicit. However, semantically relevant regions can be recognized with higher-level saliency models, which are implemented as deep neural networks trained on the eye movement patterns obtained from a large number of neurotypical viewers.[42] Such saliency maps of various degrees of semantic content were previously used to model visual attention[13] and have been applied to identify individuals with autism spectrum disorder.[12] Yet this method has rarely been applied to ADHD. Tseng *et al.*[28] were the first to investigate bottom-up and top-down visual attention during task-free natural scene viewing in ADHD. Their study design involved a disorder classification based on discriminative features from the gaze on low-level saliency maps and a map generated from the gaze of neurotypical controls (NCs). They obtained a classification accuracy of 83.3% for distinguishing individuals with ADHD (n = 21) from an age-matched control group (n = 18) by using saliency-based features.[28] However, top-down–relevant areas were extracted from a small control group, whose gaze pattern might not generalize to the true top-down visual attractiveness of the image content. Although the classification approach is promising for future clinical screening applications, it does not identify sufficiently whether increased bottom-up or decreased top-down attention is the driving factor and how these relate to deficits in attention or inhibitory control.

In this study, we investigated the influence of bottom-up and top-down mechanisms in a very large sample (N = 815) including children and adolescents with ADHD and NCs with eye tracking in a task-free naturalistic viewing experiment. For all participants, we used the same video trailer for a full-length movie that is suitable in content for the wide age range. We quantified the bottom-up and top-down content of attentional relevance in the video with a low-semantic (Finegrained) and a high-semantic (ViNet) saliency map (see "Method" and Montabone and Soto[41] and Jain *et al.*[42]). Correspondence between a person's gaze pattern and the saliency maps was quantified with the normalized scanpath saliency (NSS) measure[31] to contrast increased bottom-up visual attention with its decreased top-down counterpart in ADHD.

First, we excluded differences in basic oculomotor characteristics as the cause of any observed effects by investigating the total number of fixations, mean fixation duration, and skewness of fixation duration distribution. Second, we used the NSS to test whether children and adolescents with ADHD diverge from NCs in their coherence to the saliency maps.

Using this study design, we were able to differentiate between the increased bottom-up component and the decreased top-down component to disentangle the role of both in the visual attention mechanism that is impaired in ADHD. With this approach we aimed at further quantifying and characterizing observable behaviors of ADHD in alignment with the National Institute of Mental Health Research Domain Criteria framework to enhance the understanding of this disorder.[43,44]

METHOD

Participants

Participants were included from the Healthy Brain Network initiative by the Child Mind Institute[45] from the first to the sixth release and from the seventh release until and including the spring acquisition season in 2019. All participants were administered the Schedule for Affective Disorders and Schizophrenia for School-Age Children (K-SADS),[46,47] and further psychological assessment was done upon indication by licensed clinicians. The K-SADS is considered the standard for diagnosing child and adolescent psychiatric conditions within the age range of 6 to 18 years old.[46,47] The computerized version of K-SADS (K-SADS-COMP) is a semistructured interview that follows the *DSM-5* guidelines and involves both the child and their parent being interviewed by a clinician. The interview uses a standardized 5-point rating scale to determine clinical diagnoses. The K-SADS-COMP has displayed a high level of interrater reliability, with 8 raters scoring 94% of the items identically. The K-SADS-COMP has also shown good convergent validity against established clinical rating scales and dimensional diagnostic-specific ratings.[46]

In addition, various supplementary tests were performed for the purpose of establishing a comprehensive ADHD assessment. These included the Conners 3–Self-Report, Conners' Adult ADHD Rating Scales (CAARS), and the Strengths and Weaknesses of Attention-Deficit/Hyperactivity Disorder Symptoms and Normal Behavior (SWAN) rating scale.[48–50] All diagnoses were derived on consensus by multiple licensed psychologists and social workers, with psychopharmacological consultation support provided by psychiatrists, based on the review of the psychological assessment and all materials collected during study participation. All diagnoses were assigned based on *DSM-5*.[1,45] Participants with a clinical ADHD diagnosis in accordance with *DSM-5* were included in the ADHD group. They were subsequently grouped by their specific ADHD diagnosis, ADHD-IN and ADHD-C, based on the *DSM-5* classification. All participants who did not meet criteria for any clinical diagnosis were included in the NC group. Further exclusion criteria for both groups are detailed in Supplement 1, available online.

Written informed consent was provided before participation by legal guardians or participants of legal age. Study approval was given by the Chesapeake Institutional Review Board.

Eye tracking was performed in all participants while they completed a naturalistic and task-free viewing

TABLE 1 Demographic and Clinical Characteristics of Attention-Deficit/Hyperactivity Disorder (ADHD) Groups and Neurotypical Control (NC) Group

	Group						
	ADHD-C (n = 303)		ADHD-IN (n = 318)		NC (n = 194)		
	Mean	**(SD)**	**Mean**	**(SD)**	**Mean**	**(SD)**	**Statistical Tests**
Age, y[a]	9.34	(2.77)	11.39	(3.25)	10.09	(3.53)	ADHD-C: $\beta = -.75$, $p = .009$**, 95% CI = [−1.323, −0.185] ADHD-IN: $\beta = 1.30$, $p < .001$***, 95% CI = [0.733, 1.860]
IQ[a]	99.80	(14.46)	98.07	(14.99)	105.94	(15.02)	ADHD-C: $\beta = -6.15$, $p < .001$***, 95% CI = [−8.933, −3.360] ADHD-IN: $\beta = -7.87$, $p < .001$***, 95% CI = [−10.610, −5.136]
	n	**(%)**	**n**	**(%)**	**n**	**(%)**	
Sex, female	79	(26.07)	117	(36.79)	98	(50.52)	$\chi^2 = 30.76$, $p < .001$***, $V = 0.19$
Handedness							
Left-handed	30	(9.90)	17	(5.35)	14	(7.22)	$\chi^2 = 7.59$, $p = .022$*, $V = 0.10$
Ambidextrous	37	(12.21)	35	(11.01)	31	(15.98)	$\chi^2 = 2.12$, $p = .346$, $V = 0.05$
Right-handed	226	(74.59)	261	(82.08)	147	(75.77)	$\chi^2 = 4.70$, $p = .095$, $V = 0.08$
Comorbidities							
Specific learning disorder	62	(20.46)	67	(21.07)		—	$\chi^2 < 0.01$, $p = .930$, $V < 0.01$
Oppositional defiant disorder	85	(28.05)	25	(7.86)		—	$\chi^2 = 42.02$, $p < .001$***, $V = 0.26$
Conduct disorder	93	(30.69)	29	(9.12)		—	$\chi^2 = 44.39$, $p < .001$***, $V = 0.27$
Anxiety disorder	85	(28.05)	107	(33.65)		—	$\chi^2 = 2.02$, $p = .155$, $V = 0.06$
Major depressive disorder	9	(2.97)	23	(7.23)		—	$\chi^2 = 4.93$, $p = .026$*, $V = 0.10$
Other mental disorder	114	(37.62)	149	(46.86)		—	$\chi^2 = 5.04$, $p = .025$*, $V = 0.09$
Race and ethnicity							
Asian	7	(2.31)	5	(1.57)	12	(6.19)	$\chi^2 = 3.13$, $p = .209$, $V = 0.06$
Black or African American	55	(18.15)	38	(11.95)	18	(9.28)	$\chi^2 = 9.15$, $p = .010$*, $V = 0.11$
Hispanic	30	(9.90)	29	(9.12)	19	(9.79)	$\chi^2 = 0.12$, $p = .940$, $V = 0.01$
Two or more races	41	(13.53)	55	(17.30)	29	(14.95)	$\chi^2 = 1.72$, $p = .423$, $V = 0.05$
Unknown/some other race	22	(7.26)	23	(7.23)	25	(12.89)	$\chi^2 = 12.74$, $p = .002$**, $V = 0.13$
White	148	(48.84)	168	(52.83)	91	(46.91)	$\chi^2 = 1.92$, $p = .383$, $V = 0.05$
Sampling rate, 60 Hz	206	(67.99)	205	(64.47)	158	(81.44)	$\chi^2 = 17.25$, $p < .001$***, $V = 0.15$

Note: Clinical diagnoses were assigned based on DSM-5. χ^2 refers to Pearson χ^2 test with Yates continuity correction; V refers to Cramer's V. ADHD = attention-deficit/hyperactivity disorder; ADHD-C = ADHD combined presentation; ADHD-IN = ADHD predominantly inattentive presentation.
[a]Statistical test was performed with linear regression in reference to neurotypical controls.
**p < .05; **p < .01; ***p < .001.*

paradigm watching a short video clip. Participants in whom eye tracking was performed under incompatible acquisition setups (n = 551) and whose data did not pass quality control (n = 180) were excluded. Full details of exclusions are provided in Supplement 1, available online. A final sample of 815 participants aged 5 to 21 years (294 female;

mean [SD] age = 10.32 [3.28]) was analyzed. Demographics and clinical characteristics are presented in Table 1. For further details and variable grouping, see Supplement 1, Table S1 and Figures S1 and S2, available online.

Naturalistic Stimuli Paradigm

Participants were shown a 117-second video trailer of a feature-length film titled *Diary of a Wimpy Kid*. This video was shown as part of the naturalistic stimuli paradigm of a larger test battery (for details on the test battery see Alexander *et al.*[45]). Before and after the video, one or more other paradigms from the test battery were performed, which are not analyzed here. The order of the video within the battery was randomized for each participant. The video resolution was set to 800 × 600 pixels. Importantly, besides watching the video, no task was given. The full clip of the trailer for *Diary of a Wimpy Kid* is available at https://www.youtube.com/watch?v=7ZVEIgPeDCE.

Data Acquisition

Data were recorded at 2 sites using the same setups (Rutgers University Brain Imaging Center and CitiGroup Cornell Brain Imaging Center). Participants were seated in a sound-shielded room at a distance of approximately 65 cm from the screen. Gaze position was recorded using an infrared video-based eye tracker (iView X RED-m; SensoMotoric Instruments GmbH, Teltow, Germany) (spatial resolution 0.1°, position accuracy 0.5°) at a sampling rate of 60 Hz (70% of sample) or 120 Hz (30% of sample). Before each video, a calibration procedure on a 5-point grid was performed until a spatial error of less than 2° for any point and an average spatial error of less than 1° over all points was obtained.[45]

Fixations, saccades, and blinks were detected from the raw gaze stream of the right eye using a dispersion-based algorithm employing a dispersion-threshold identification with a duration threshold of 75 ms and a visual angle threshold of 1.6° to detect fixations.[51] Specifically, samples that did not surpass a spatial dispersion of more than 1.6° of visual angle and that lasted for at least 75 ms were defined as fixations. Samples between any 2 fixations were classified as saccade. Samples with missing measurements were classified as blinks or tracker loss. To ensure high-quality data from eye tracking in a pediatric clinical sample, a quality control procedure was performed before extracting subsequent oculomotor measures. For details on the quality control procedure, see Supplement 1, Figure S1, available online.

Saliency Maps

To quantify visual attention with eye-tracking data, we analyzed the correspondence between the viewer's gaze and the salience values of the stimulus video computed by the 2 different saliency models. These models estimate the visual relevance of each region of an image and thus predict the spatial distribution of the participants' gaze when viewing the respective image.[40] We used 2 orthogonal saliency models: Finegrained[41] and ViNet.[42] Finegrained (Figure 1C) is an optimized version of the seminal model proposed by Itti *et al.*,[40] in which low-level features of the image such as color, intensity, and orientation are extracted and integrated. ViNet (Figure 1B) is a state-of-the-art model based on deep neural networks trained for saliency prediction in dynamic scenes. In contrast to Finegrained, ViNet incorporates a notion of temporality by modeling the dynamic effects that occur in a video. Moreover, it uses a network pretrained on an action-recognition dataset; subsequently, the network was trained on a video saliency dataset. Consequently, ViNet captures high-level features such as objects or faces and encodes action and movement information. It is one of the top-performing models according to a benchmark evaluation hosted at https://mmcheng.net/videosal/.[52] The evaluation was conducted based on a saliency prediction for 300 videos that were not used in model development and with a holdout training dataset. The models' ranks are set by 5 classical metrics that were computed based on the accuracy of the model predictions.[42] ViNet has been shown to perform well in predicting human eye fixations on images. The model has been shown to outperform several state-of-the-art saliency models on benchmark datasets, indicating its superior performance.[41,42,52]

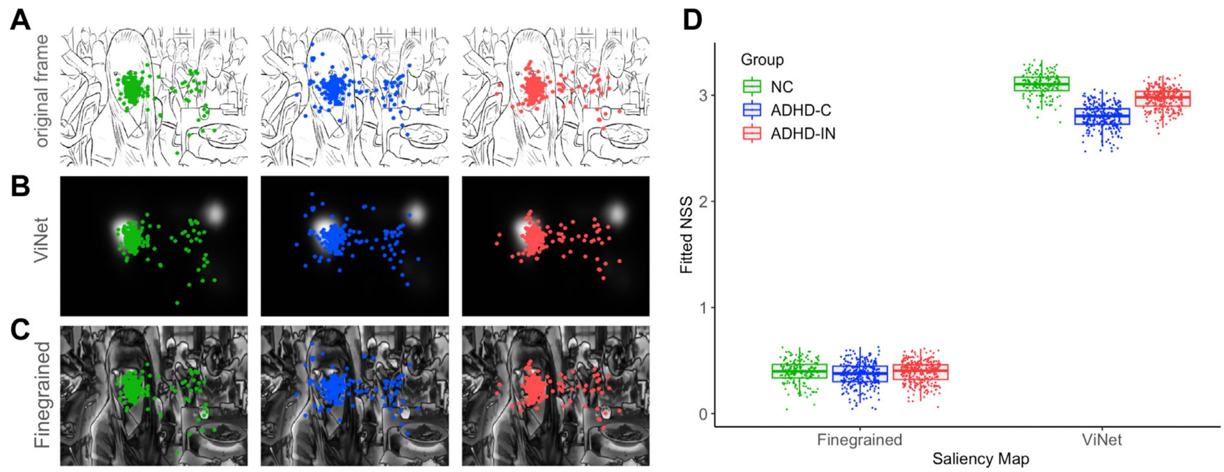
In this study, Finegrained was used as a bottom-up representation, and ViNet was used as a top-down representation of the video content. To calculate Finegrained, we used the implementation provided by OpenCV 4.5.1 for Python 3.7 (https://docs.opencv.org/4.5.1/d8/d65/group__saliency.html), and to calculate ViNet, we used the original implementation provided in Jain *et al.*[42]

Eye-Tracking Measures

Sensitivity Analysis. To exclude any chance that the main analysis results are driven by oculomotor differences between the groups, we calculated the total number of fixations, mean fixation duration, and skewness of fixation duration distribution. Full details of these measures and statistical analyses are provided in Supplement 2, Figures S3 to S5 and Tables S2 to S4, available online.

Main Analysis. A number of metrics exist to compare a person's gaze to a saliency map.[53–55] Here, we chose the NSS because it makes the fewest assumptions about the saliency map and has been recommended for saliency

FIGURE 1 Fixation Distribution of Study Populations on an Example Frame



Note: *Fixation distribution of ADHD-C and ADHD-IN groups compared with NC group on an example frame (frame 898) of the video trailer for Diary of a Wimpy Kid. (A) Fixations over a sketch of the original image. (B) Fixations over the ViNet (top-down representing) saliency map. (C) Fixations over the Finegrained (bottom-up representing) saliency map. (D) Box plot contrasting fitted mean normalized scanpath saliency values from the linear mixed-effects model of the NC, ADHD-IN, and ADHD-C groups for the Finegrained and ViNet saliency maps. In panels (B) and (C), bright areas denote highly salient regions according to the respective saliency map. Illustrated fixations in panels (A), (B), and (C): NC, n = 193; ADHD-C, n = 193; ADHD-IN, n = 193. For the purpose of comparison in panels (A), (B), and (C), illustrated fixations of the ADHD-C and ADHD-IN groups were randomly drawn to match the number of fixations of the control group in this frame. A figure including all fixations of these groups is provided in Supplement 2, Figure S7, available online. ADHD = attention-deficit/hyperactivity disorder; ADHD-C = ADHD combined presentation; ADHD-IN = ADHD predominantly inattentive presentation; NC = neurotypical control.*

evaluation.[31,54] The NSS measures the correspondence between a participant's fixation location as measured by the eye tracker and the saliency value of the fixated location of the corresponding video frame. For a formal definition of the NSS measure, see Supplement 3, available online. Mean NSS values for each participant were calculated over the entire video by averaging NSS scores for each fixation and over each scene by averaging NSS scores for fixations occurring within the scene. Scenes were identified by visual inspection; corresponding start and end frames are provided in Supplement 3, Table S5, available online.

Statistical Analysis. All statistical analyses were conducted in R version 4.1.1. Formulas are expressed in Wilkinson notation. Effects from sensitivity analysis and mean NSS analysis were evaluated at a significance level of .05. Effects from scene-wise analysis were evaluated at a corrected level of 0.00071 to account for multiple comparisons using the method proposed by Nyholt.[56] For a full description and calculation of the corrected significance level, see Supplement 3, available online.

Main Analysis. To investigate correspondence with the low- and high-level saliency map, we analyzed the effect of ADHD on mean NSS and scene-wise NSS of both saliency maps using a mixed-effects linear model. We used the NSS (continuous) as the dependent variable; used age (continuous), sex (categorical: male [reference category], female), IQ (continuous), handedness (categorical: right-handed [reference category], ambidextrous, left-handed), race and ethnicity (categorical: Asian, Black or African American, Hispanic, unknown/some other race [reference category], two or more races, White), and sampling rate (categorical: 60 Hz [reference category], 120Hz) as confounding covariates; used the saliency map (categorical: Finegrained [reference category], ViNet) and ADHD presentation (categorical: NC [reference category], ADHD-IN, ADHD-C) as fixed effects of interest; and added an interaction effect of saliency map and ADHD presentation. The model was defined as follows:

$$\begin{aligned}&\text{NSS} \sim \text{Age} + \text{Sex} + \text{IQ} + \text{Handedness}\\&+ \text{Race and Ethnicity} + \text{Sampling Rate}\\&+ \text{Saliency Map} \times \text{ADHD Presentation} + (1|\text{Participant})\end{aligned}$$

Because the random effect of participant—ie, (1|Participant)—is partially captured by the saliency map, supplementary results of a parsimonious model definition without a random effect of participant are provided in Supplement 3, Table S6, available online.

The results were validated using the median instead of the mean, which is more robust to outliers in the data. To inspect potential ADHD severity effects, we performed an additional sensitivity analysis with the same model specification of the mean NSS but with a dimensional representation of ADHD that did not yield significant results. The results on supplementary analyses and an auxiliary analysis with the ADHD-C group as reference category are provided in Supplement 3, Tables S7 to S11, available online. For the scene-wise analysis, separate models with the same model definitions were calculated for each scene with the mean NSS values for the corresponding scene.

RESULTS

Sensitivity Analysis

The sensitivity analysis yielded no significant effect for either ADHD-IN or ADHD-C on total number of fixations, mean fixation duration, and skewness of fixation duration distribution (all $p > .05$). For a detailed report on these results and the calculation of the measures, see Supplement 2, Tables S2 to S4, available online.

Main Analysis

The analysis on mean NSS yielded a main effect of age ($\beta = .0185$, $p = .002$, 95% CI = [0.00702, 0.02994]), indicating that older participants have higher NSS scores. We observed a main effect of the saliency map on NSS ($\beta = 2.7050$, $p < .001$, 95% CI = [2.57069, 2.83969]) showing higher NSS values in the ViNet (top-down representing) saliency map. Most importantly, the analysis yielded a significant interaction effect of saliency map × ADHD-C on mean NSS ($\beta = -.2802$, $p = .001$, 95% CI = [−0.45196, −0.10847]), indicating that individuals with ADHD-C scored lower NSS corresponding to the ViNet saliency map than NCs, but not corresponding to the Finegrained (bottom-up representing) saliency map (Figure 1D). Importantly, there was no significant main effect of ADHD-C or ADHD-IN or any significant interaction effect of saliency map × ADHD-IN. Results are summarized in Table 2, and an example of groupwise fixation patterns in 1 frame is shown in Figure 1A-C.

Scene-Wise NSS Analysis

The scene-wise analysis yielded 39 out of 92 scenes in which the interaction effect of saliency map × ADHD-C on scene-wise mean NSS was significant. After correction for multiple comparisons at a corrected significance level of .00071, the interaction effect remained significant in 8 scenes (scene numbers 14, 23, 25, 31, 43, 62, 84, and 89). Scene-wise fitted NSS values by group for the ViNet saliency map including highlighted scenes in which the interaction effect was significant are provided in Figure 2. Detailed results for each scene are provided in Supplement 3, Table S12, available online.

Post Hoc Exploration of Scene Characteristics

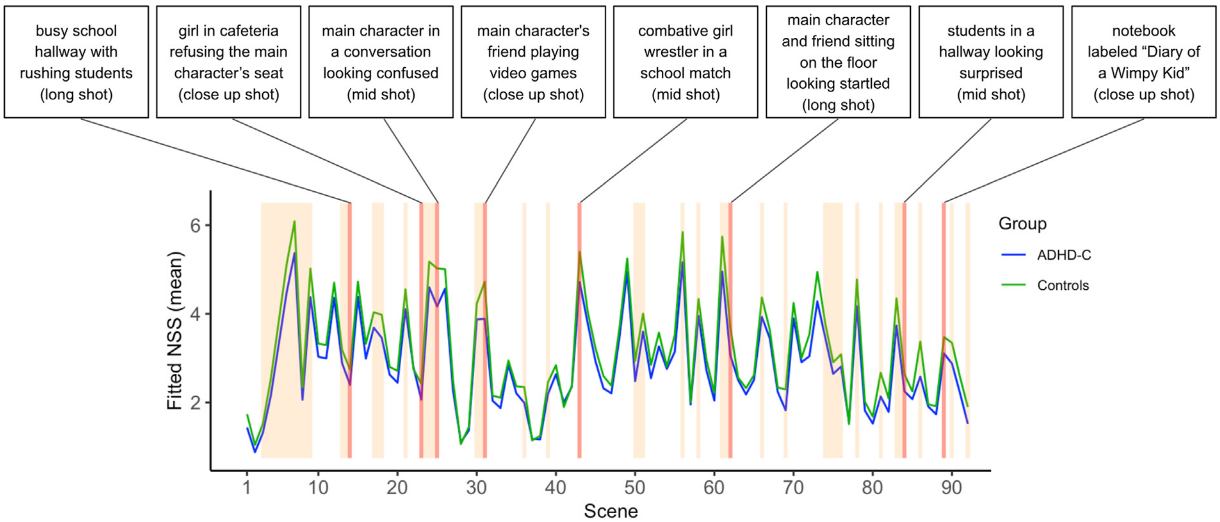
To explore potential factors contributing to the interaction effect in certain scenes, we sorted the scenes by magnitude

TABLE 2 Results of Mixed-Effects Linear Model for Mean Normalized Scanpath Saliency Values

Variable	β	SE	t	p	95% CI
Age	.0185	0.0059	3.14	.002**	[0.00702, 0.02994]
Sex	.0219	0.0375	0.58	.560	[−0.05101, 0.09480]
IQ	.0015	0.0013	1.23	.219	[−0.00090, 0.00398]
Ambidextrous	.0220	0.0558	0.40	.693	[−0.08638, 0.13045]
Left-handed	.0304	0.0681	0.45	.655	[−0.10187, 0.16273]
Asian	.1119	0.1179	0.95	.343	[−0.11728, 0.34113]
Black or African American	.1143	0.0791	1.45	.149	[−0.03944, 0.26803]
Hispanic	.1002	0.0850	1.18	.238	[−0.06493, 0.26539]
Two or more races	.1404	0.0765	1.84	.067	[−0.00831, 0.28920]
White	.2229	0.0664	3.36	<.001***	[0.09386, 0.35192]
Sampling rate	−.0512	0.0389	−1.32	.188	[−0.12669, 0.02438]
ADHD-C	.0059	0.0657	0.09	.929	[−0.12213, 0.13383]
ADHD-IN	−0.0093	0.0644	−0.14	.886	[−0.13482, 0.11629]
Saliency map	2.7050	0.0687	39.39	<.001***	[2.57069, 2.83969]
Saliency map × ADHD-C	−0.2802	0.0877	−3.20	.001**	[−0.45196, −0.10847]
Saliency map × ADHD-IN	−0.1302	0.0859	−1.52	0.130	[−0.29834, 0.03798]

Note: *ADHD = attention-deficit/hyperactivity disorder; ADHD-C = ADHD combined presentation; ADHD-IN = ADHD predominantly inattentive presentation; SE = standard error.*
$^{**}p < .01$; $^{***}p < .001$.

FIGURE 2 Scene-Wise Comparison of Fitted Normalized Scanpath Saliency Values

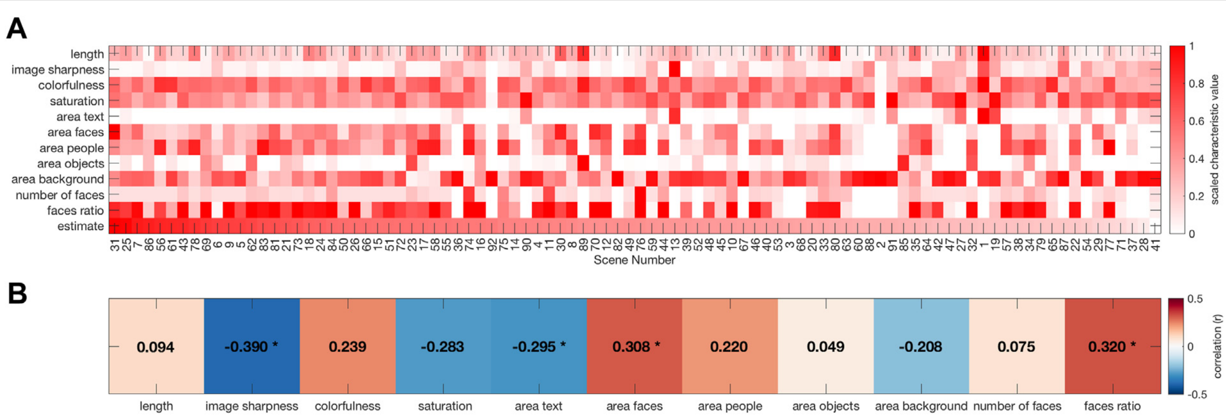


Note: Scene-wise fitted NSS values (mean per scene) for the top-down representing ViNet saliency map of neurotypical controls and ADHD-C group. Highlighted areas indicate scenes in which the interaction effect of saliency map × ADHD-C was significant at an uncorrected level (orange) and at a level corrected for multiple comparisons (red) including a description of the primary visual content of that scene. ADHD = attention-deficit/hyperactivity disorder; ADHD-C = ADHD combined presentation; NSS = normalized scanpath saliency.

of the interaction effect estimate from smallest (negative) to largest and illustrated corresponding scene length; image sharpness (combining clarity of focus and contrast), colorfulness, and saturation; proportion of area covered by text, faces, people, objects, and background; average number of faces; and proportion of frames with a face over all frames, termed the faces ratio. Details of the extraction of those features are provided in other publications[57,58] and in Supplement 4, available online.

Pairwise correlations of each scene characteristic with the interaction effect magnitude were evaluated at a corrected significance level of .0054 according to the correction proposed by Nyholt[56] (see Supplement 4, available online, and Nyholt[56]). Figure 3 shows the value distribution (scaled between 0 and 1) of each feature, ordered by the magnitude of the interaction effect in the scene (Figure 3A) and pairwise correlation of scene characteristic with the magnitude of the interaction effect (Figure 3B).

FIGURE 3 Comparison of Scene Characteristics With Magnitude of Interaction Effect



Note: (A) Value distribution of scene characteristic features (scaled between 0 and 1) ordered by magnitude of interaction effect of saliency map × attention-deficit/hyperactivity disorder combined presentation in each scene. (B) Correlation between scene characteristic features (scaled between 0 and 1) and magnitude of interaction effect. Correlation coefficients are provided for each scene characteristic; asterisks denote significant effects at a level corrected for multiple comparisons of 0.0054.

The post hoc analysis yielded significant negative correlations with the image sharpness ($r = -0.390$, $p = .0001$) and the area of text ($r = -0.295$, $p = .0043$). Significant positive correlations were obtained for the area of faces ($r = 0.308$, $p = .0028$) and faces ratio ($r = 0.320$, $p = .0019$).

DISCUSSION

Eye movement studies in ADHD have gained popularity[14] due to the widespread overlap between brain areas involved in visual attention and the frontostriatal circuits that are thought to be anomalous in ADHD.[11,12,28,34–37] Previous studies using task-based paradigms such as antisaccade and fixation tasks showed higher occurrences of task-inappropriate eye movements in children and adolescents with ADHD.[14] Yet apart from one study,[28] past research has scarcely investigated eye movements in ADHD during task-free viewing of complex, natural, and dynamic content. Such content would simulate real-life visual stimuli more closely and elicit gaze behavior similar to natural gaze in daily life, thus offering greater ecological validity for findings.[12,13,25,26,28] Furthermore, past studies did not differentiate objectively between top-down and bottom-up mechanisms of visual attention to naturalistic dynamic contents in ADHD. To this end, the main objective of this study was to contrast top-down with bottom-up visual attention in a very large sample of children and adolescents with ADHD in a task-free, dynamic and naturalistic viewing paradigm with objective quantifications for content saliency.

The main aim of this study was to identify the cause of differences in gaze allocation in ADHD. We used 2 opposing saliency maps to quantify top-down and bottom-up visual attractiveness of all areas in a video and compared participants' gaze correspondence with both saliency maps using the NSS measure.

The main analysis revealed a significant interaction effect of saliency map × ADHD-C on overall mean NSS. Participants with ADHD-C scored lower NSS values on the ViNet saliency map, which represents areas visually and semantically relevant to a top-down attention mechanism. However, participants with ADHD-C did not score higher on the Finegrained saliency map, showing no increased tendency to allocate their gaze to bright, high-contrast features visually salient to a bottom-up attention mechanism. Because we mainly observed a tendency to look away from areas with high semantic value in the ADHD-C group compared with the NC group, we suggest poorer top-down visual attention, rather than increased bottom-up visual attention, to be the driving factor in both competing mechanisms in ADHD-C.[13,28,29] Furthermore, our results provide no evidence that bottom-up–salient features would have worked as distractors for ADHD-C because we did not observe a compensatory effect of increased NSS on the Finegrained saliency map in this group.

We did not observe a similar interaction effect for the ADHD-IN group. The categorical distinction between the ADHD-IN and the ADHD-C groups is that in addition to inattention, individuals with ADHD-C also show symptoms of increased and inappropriate motor activity and impulsive behaviors. A recent meta-analysis by Saad *et al.*[59] identified distinct neurobiological signatures for both presentations. The authors concluded the frontostriatal-thalamic regions that include motor networks to be aberrant in ADHD-C, while in ADHD-IN frontoparietal regions are the main areas deviating from NCs.[59] Hence, observing an effect solely for the ADHD-C group might provide further support for distinct neurocognitive pathophysiology in the 2 presentations. Attributing the observed effect to reduced top-down integration with consecutive generation of inappropriate eye movement signals might be analogous to the inappropriate motor activity observed in ADHD-C, but not ADHD-IN.

Furthermore, the analysis did not yield a main effect of either ADHD-C or ADHD-IN; consequently, a simplistic contrast between ADHD and NCs regardless of the semantic level does not adequately describe the impairments of bottom-up and top-down visual attention mechanisms in ADHD. In turn, differences in gaze patterns are specific to the semantic level of information in the video. This finding further supports the view that visual attention in ADHD-C is a more complex mechanism related to aberrant control of top-down visual attention.

In addition, we observed a significant main effect of age, as NSS scores increased with age. This may lead to the interpretation that with increasing age, gaze patterns appear to exhibit improved integration of top-down and bottom-up visual attention, leading to a pattern of mature visual attention. This finding raises the question whether children and adolescents with ADHD-C show a mere delay in top-down attentional control or whether such deficits constitute an endophenotype of ADHD-C.[60] To answer this question, longitudinal designs would be needed to test multiple time points across development from childhood through adolescence and into adulthood, including simultaneous assessment of symptom severity and remission in an individual.

In the scene-wise analysis, we observed a large proportion of scenes in which the interaction effect of saliency map × ADHD-C was significant and 6 scenes in which it survived correction for multiple comparisons. A post hoc analysis revealed significant positive correlations for the area covered by faces and the proportion of frames with at least 1 face. Due to the plot of the trailer, which focuses on social interactions, faces constitute a highly semantic feature. In line with our argument about poorer top-down control, this finding could indicate that children and adolescents with ADHD-C have reduced ability to hold their gaze on such content in particular in scenes with a high occurrence of faces.

We also observed a significant negative correlation with image sharpness. Image sharpness depends on image focus and contrast and increases with the appearance of sharp edges. This finding might suggest that with lower image sharpness, top-down visually relevant areas are much less visually distinguishable from irrelevant content for individuals with ADHD-C.

Lastly, we observed a negative correlation for the area covered by text. This finding is not surprising because it shows that the interaction effect diminishes in the presence of text. As the majority of participants in this study are of school attendance age (5-21 years), it is expected that the appearance of text triggers its reading, hence the similar gaze between both groups.

In recent years, FSN dysfunction has emerged as a promising causal theory for explaining symptoms in ADHD.[36,61–63] Several brain structures that are part of the FSN are also involved in the mechanisms of top-down and bottom-up visual attention. Top-down signals are predominantly generated in the dorsolateral prefrontal regions.[38] Together with regions in the frontal eye fields, which play a substantial role in target selection,[39] and the lateral intraparietal area, which is involved in covert spatial attention, projections of these areas to the basal ganglia and thalamus provide the main input of visually relevant information to the SC.[38] Bottom-up signals emerge from primary visual cortical areas with direct projections to the SC. The SC is responsible for the integration of both bottom-up and top-down signals through a winner-takes-all mechanism and sends corresponding motor commands to the brainstem.[37,38]

Whereas the output of the SC determines the next gaze allocation, several of the structures presented here can play a role in the aberrant gaze patterns observed in ADHD-C. The deviant fixation patterns we observed were with video content relevant to top-down visual attention, and a large body of literature stresses deviation in frontal regions in ADHD.[36,61–63] Consequently, areas generating bottom-up signals are unlikely to play a role in explaining these aberrant gaze patterns.

The symptoms of ADHD-C include increased and inappropriate motor activity[1]; thus, another line of investigation might focus on the role of the basal ganglia. The basal ganglia control the release and inhibition of body and eye movements via the direct and indirect pathways, respectively.[64] An increased tendency to look away from semantically relevant stimuli, as observed in the ADHD-C group in our study, might indicate disturbances in these pathways. A poor balance between the 2 pathways might release eye movements that are not beneficial in following the content of the video, which we observed through lower NSS values on the higher-level saliency map. This effect was exclusively observed in the ADHD-C subgroup in our study; the ADHD-C presentation is characterized by increased hyperactivity and impulsivity manifesting in increased and inappropriate motor activity.[1] In line with the clinical definition of ADHD-C, we argue that eye movements may constitute another facet of these symptoms because similar structures are responsible for releasing both body and eye movements. To validate this argument, further studies are needed that link aberrant gaze allocation to top-down–relevant information with deviations in neuroanatomy and neurophysiology in the basal ganglia.

Although fixations reveal shifts in overt attention,[13] it is also known that covert attention precedes overt attention in determining the next gaze position.[29] Our findings can be interpreted only within the scope of overt attention. To investigate bottom-up and top-down attentional processes of covert attention, our study would need to go beyond the current analysis incorporating content in the periphery of the visual field.

Furthermore, in this study, saliency models were used to quantify top-down and bottom-up visual attractiveness. However, these models have limitations in accurately representing top-down and bottom-up visual attractiveness due to initial assumptions and the quality and quantity of training data. We used the state-of-the-art ViNet saliency model, a deep learning–based model trained on fixations from neurotypical individuals, which makes it generalizable to a wide range of visual content.[41,42] Although these training data from neurotypical individuals boosts the prediction accuracy of the model, it may not be sufficiently representative of neurotypical top-down attention, constituting a potential limitation on its ability to contrast individuals with ADHD from NCs. However, the high overall NSS values of the saliency map (mean NSS $= 2.93 >$ benchmark for saliency maps),[42] bolsters our confidence that this saliency map reflects top-down visual attractiveness to the best extent currently available. Moreover, evaluating attention through the coherence to 2 orthogonal saliency models of visual attention may not capture all aspects of attentional (dys)function. While the competition between bottom-up and top-down attention is crucial to consider, attentional control and sustained attention may also be of interest. Nonetheless, our investigation of the interaction between ADHD presentation and coherence to bottom-up and top-down visual attractiveness models contributes to understanding attentional dysfunction in ADHD.

The lack of an ADHD predominantly hyperactive-impulsive presentation subgroup in this study, which was excluded due to a small sample size, is a further limitation. Contrasting the differences between the exclusively hyperactive and impulsive presentation with other presentations of ADHD might have provided more clearly observable effects. However, the ADHD predominantly hyperactive-impulsive presentation is very rare in the general population.[65]

Finally, although a chin rest was used to ensure stable head position and participants were instructed to remain still, minor head movements can occur, particularly in pediatric and psychiatric populations.[66] An increased tendency for movement in ADHD might have posed a further issue in the accuracy and precision of eye movement data obtained.[1,66,67] However, because we did not observe any group effects in the sensitivity analysis and in the Fine-grained NSS analysis, it is very unlikely that potential head motion caused our observed results of reduced top-down visual attention in ADHD-C.

In light of the unknown contribution of competing mechanisms of bottom-up and top-down visual attention in ADHD, we evaluated quantitative measures of gaze allocation, which yielded support for the notion of decreased top-down visual attention in ADHD-C. We did not observe similar effects in the ADHD-IN group, suggesting distinct pathophysiology in this presentation. The high overlap of areas in the FSN and areas responsible for visual attention renders FSN dysfunction highly plausible as the cause of inappropriate eye movements in ADHD-C. Yet, additional investigations are needed to link reduced top-down visual attention to disturbance in the direct and indirect pathways in the basal ganglia to further validate this theory as a potential etiology of ADHD.

Future research would also need to deepen the analysis of gaze coherence with attentional maps by manipulating attentional demands and by comparison with other measures of attention (eg, reaction times or accuracy on visual search tasks). Further research may combine this approach with other neuroimaging techniques such as electroencephalography or functional magnetic resonance imaging to gather insights into the neural mechanisms underlying attention and contribute to the discovery of reliable biomarkers of ADHD. Future research in this area could also benefit from investigating how visual attention differs between individuals with ADHD and individuals with other types of neurological and psychiatric dysfunction, thereby further advancing our understanding of the underlying mechanisms of visual attention in psychopathology. The knowledge gained from such studies might have the potential to be translated into clinical settings as ADHD screening and into educational settings to improve learning and scholastic achievements.

CRediT authorship contribution statement

Sabine Dziemian: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Gaston Bujia:** Writing – review & editing, Writing – original draft, Software, Methodology, Formal analysis, Data curation. **Paul Prasse:** Writing – review & editing, Software, Data curation. **Zofia Barańczuk-Turska:** Writing – review & editing, Visualization, Methodology, Formal analysis. **Lena A. Jäger:** Writing – review & editing, Software, Resources, Funding acquisition. **Juan E. Kamienkowski:** Writing – review & editing, Software, Resources, Methodology, Funding acquisition. **Nicolas Langer:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization.

Accepted March 27, 2024.

Sabine Dziemian, Zofia Barańczuk-Turska, Lena A. Jäger, and Nicolas Langer are with University of Zurich, Zurich, Switzerland. Sabine Dziemian and Nicolas Langer are also with the Neuroscience Center Zurich (ZNZ), Zurich, Switzerland. Gaston Bujia and Juan E. Kamienkowski are with the University of Buenos Aires, Buenos Aires, Argentina. Lena A. Jäger and Paul Prasse are with the University of Potsdam, Potsdam, Germany.

This work was partially funded by the Swiss National Science Foundation under grant number 100014_175875, by the Neuroscience Center Zurich under a PhD grant, by the German Federal Ministry of Education and Research under grant 01|S20043, and by the CONICET, Argentina, and the University of Buenos Aires, Argentina.

The research was performed with permission from the Chesapeake Institutional Review Board.

Consent has been provided for descriptions of specific patient information.

Zofia Barańczuk-Turska served as the statistical expert for this research.

The authors thank all the children and their families for their participation.

Disclosure: Sabine Dziemian, Gaston Bujia, Paul Prasse, Zofia Barańczuk-Turska, Lena A. Jäger, Juan E. Kamienkowski, and Nicolas Langer have reported no biomedical financial interests or potential conflicts of interest.

Correspondence to Nicolas Langer, PhD, Department of Psychology, Methods of Plasticity Research, University of Zurich, Binzmühlestrasse 14, Zurich CH-8050, Switzerland; e-mail: n.langer@psychologie.uzh.ch

https://doi.org/10.1016/j.jaacop.2024.03.001

REFERENCES

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 5th ed. Arlington, VA: American Psychiatric Association; 2013.
2. Hechtman L, Swanson JM, Sibley MH, *et al*. Functional adult outcomes 16 years after childhood diagnosis of attention-deficit/hyperactivity disorder: MTA results. J Am Acad Child Adolesc Psychiatry. 2016;55(11):945-952. https://doi.org/10.1016/j.jaac.2016.07.774
3. Faraone SV, Banaschewski T, Coghill D, *et al*. The World Federation of ADHD International Consensus Statement: 208 evidence-based conclusions about the disorder. Neurosci Biobehav Rev. 2021;128:789-818. https://doi.org/10.1016/j.neubiorev.2021.01.022
4. Barkley RA. Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. Psychol Bull. 1997;121(1):65-94.
5. Swanson JM, Sergeant JA, Taylor E, Sonuga-Barke EJ, Jensen PS, Cantwell DP. Attention-deficit hyperactivity disorder and hyperkinetic disorder. Lancet. 1998; 351(9100):429-433. https://doi.org/10.1016/S0140-6736(97)11450-7
6. Swanson JM. Role of executive function in ADHD. J Clin Psychiatry. 2003;64(Suppl 14):35-39.
7. McCarthy H, Skokauskas N, Frodl T. Identifying a consistent pattern of neural function in attention deficit hyperactivity disorder: a meta-analysis. Psychol Med. 2014;44(4): 869-880. https://doi.org/10.1017/S0033291713001037
8. Kaiser A, Aggensteiner PM, Baumeister S, Holz NE, Banaschewski T, Brandeis D. Earlier versus later cognitive event-related potentials (ERPs) in attention-deficit/hyperactivity disorder (ADHD): a meta-analysis. Neurosci Biobehav Rev. 2020;112: 117-134. https://doi.org/10.1016/j.neubiorev.2020.01.019
9. Papanicolas I, Woskie LR, Jha AK. Health care spending in the United States and other high-income countries. JAMA. 2018;319(10):1024-1039. https://doi.org/10.1001/jama.2018.1150
10. Mihajlović V, Grundlehner B, Vullers R, Penders J. Wearable, wireless EEG solutions in daily life applications: what are we missing? IEEE J Biomed Health Inform. 2014;19(1): 6-21. https://doi.org/10.1109/JBHI.2014.2328317
11. Itti L. New eye-tracking techniques may revolutionize mental health screening. Neuron. 2015;88(3):442-444. https://doi.org/10.1016/j.neuron.2015.10.033
12. Wang S, Jiang M, Duchesne XM, *et al*. Atypical visual saliency in autism spectrum disorder quantified through model-based eye tracking. Neuron. 2015;88(3):604-616. https://doi.org/10.1016/j.neuron.2015.09.042
13. Borji A, Itti L. State-of-the-art in visual attention modeling. IEEE Trans Pattern Anal Mach Intell. 2013;35(1):185-207. https://doi.org/10.1109/TPAMI.2012.89
14. Chamorro Y, Betz LT, Philipsen A, Kambeitz J, Ettinger U. The eyes have it: a meta-analysis of oculomotor inhibition in attention-deficit/hyperactivity disorder. Biol Psychiatry Cogn Neurosci Neuroimaging. 2022;7(11):1090-1102. https://doi.org/10.1016/j.bpsc.2021.05.004
15. Galgani F, Sun Y, Lanzi PL, Leigh J. Automatic analysis of eye tracking data for medical diagnosis2009 IEEE Symposium on Computational Intelligence and Data Mining, Nashville, Tennessee, USA. New York, NY: IEEE; 2009:195-202. https://doi.org/10.1109/cidm.2009.4938649
16. Gould TD, Bastain TM, Israel ME, Hommer DW, Castellanos FX. Altered performance on an ocular fixation task in attention-deficit/hyperactivity disorder. Biol Psychiatry. 2001;50(8):633-635. https://doi.org/10.1016/S0006-3223(01)01095-2
17. Loe IM, Feldman HM, Yasui E, Luna B. Oculomotor performance identifies underlying cognitive deficits in attention-deficit/hyperactivity disorder. J Am Acad Child Adolesc Psychiatry. 2009;48(4):431-440. https://doi.org/10.1097/CHI.0b013e31819996da
18. Bucci MP, Stordeur C, Septier M, Acquaviva E, Peyre H, Delorme R. Oculomotor abnormalities in children with attention-deficit/hyperactivity disorder are improved by methylphenidate. J Child Adolesc Psychopharmacol. 2017;27(3):274-280. https://doi.org/10.1089/cap.2016.0162
19. Falck-Ytter T, Pettersson E, Bölte S, D'Onofrio B, Lichtenstein P, Kennedy DP. Difficulties maintaining prolonged fixation and attention-deficit/hyperactivity symptoms share genetic influences in childhood. Psychiatry Res. 2020;293:113384. https://doi.org/10.1016/j.psychres.2020.113384
20. Munoz DP, Armstrong IT, Hampton KA, Moore KD. Altered control of visual fixation and saccadic eye movements in attention-deficit hyperactivity disorder. J Neurophysiol. 2003;90(1):503-514. https://doi.org/10.1152/jn.00192.2003
21. Mahone EM, Mostofsky SH, Lasker AG, Zee D, Denckla MB. Oculomotor anomalies in attention-deficit/hyperactivity disorder: evidence for deficits in response preparation and inhibition. J Am Acad Child Adolesc Psychiatry. 2009;48(7):749-756. https://doi.org/10.1097/CHI.0b013e3181a565f1
22. Goto Y, Hatakeyama K, Kitama T, *et al*. Saccade eye movements as a quantitative measure of frontostriatal network in children with ADHD. Brain Dev. 2010;32(5): 347-355. https://doi.org/10.1016/j.braindev.2009.04.017
23. Siqueiros Sanchez M, Falck-Ytter T, Kennedy DP, *et al*. Volitional eye movement control and ADHD traits: a twin study. J Child Psychol Psychiatry. 2020;61(12):1309-1316. https://doi.org/10.1111/jcpp.13210
24. Hanisch C, Radach R, Holtkamp K, Herpertz-Dahlmann B, Konrad K. Oculomotor inhibition in children with and without attention-deficit hyperactivity disorder (ADHD). J Neural Transm. 2006;113(5):671-684. https://doi.org/10.1007/s00702-005-0344-y
25. Kulke L, Pasqualette L. Emotional content influences eye-movements under natural but not under instructed conditions. Cogn Emot. 2022;36(2):332-344. https://doi.org/10.1080/02699931.2021.2009446
26. Salmi J, Metwaly M, Tohka J, *et al*. ADHD desynchronizes brain activity during watching a distracted multi-talker conversation. Neuroimage. 2020;216:116352. https://doi.org/10.1016/j.neuroimage.2019.116352
27. Langer N, Ho EJ, Alexander LM, *et al*. A resource for assessing information processing in the developing brain using EEG and eye tracking. Sci Data. 2017;4:170040. https://doi.org/10.1038/sdata.2017.40
28. Tseng PH, Cameron IGM, Pari G, Reynolds JN, Munoz DP, Itti L. High-throughput classification of clinical populations from natural viewing eye movements. J Neurol. 2013;260(1):275-284. https://doi.org/10.1007/s00415-012-6631-2
29. Peterson MS, Kramer AF, Irwin DE. Covert shifts of attention precede involuntary eye movements. Percept Psychophys. 2004;66(3):398-405. https://doi.org/10.3758/BF03194888
30. Navalpakkam V, Itti L. Modeling the influence of task on attention. Vision Res. 2005; 45(2):205-231. https://doi.org/10.1016/j.visres.2004.07.042
31. Peters RJ, Iyer A, Itti L, Koch C. Components of bottom-up gaze allocation in natural images. Vision Res. 2005;45(18):2397-2416. https://doi.org/10.1016/j.visres.2005.03.019
32. Theeuwes J. Top-down and bottom-up control of visual selection. Acta Psychol. 2010; 135(2):77-99. https://doi.org/10.1016/j.actpsy.2010.02.006
33. Koch C, Ullman S. Shifts in selective visual attention: towards the underlying neural circuitry. Hum Neurobiol. 1985;4:219-227.
34. Corbetta M, Shulman GL. Control of goal-directed and stimulus-driven attention in the brain. Nat Rev Neurosci. 2002;3(3):201-215. https://doi.org/10.1038/nrn755
35. Miller EK, Buschman TJ. Cortical circuits for the control of attention. Curr Opin Neurobiol. 2013;23(2):216-222. https://doi.org/10.1016/j.conb.2012.11.011
36. Dickstein SG, Bannon K, Castellanos FX, Milham MP. The neural correlates of attention deficit hyperactivity disorder: an ALE meta-analysis. J Child Psychol Psychiatry. 2006;47(10):1051-1062. https://doi.org/10.1111/j.1469-7610.2006.01671.x
37. Leigh JR, Zee DS. The Neurology of Eye Movements. Oxford: Oxford University Press; 2015.
38. White BJ, Munoz DP. The superior colliculus. In: Liversedge SP, Gilchrist I, Everling S, eds. The Oxford Handbook of Eye Movements. Oxford: Oxford University Press; 2011: 196-213.
39. Schall JD, Cohen JY. The neural basis of saccade target selection. In: Liversedge SP, Gilchrist I, Everling S, eds. The Oxford Handbook of Eye Movements. Oxford: Oxford University Press; 2011:358-381.
40. Itti L, Koch C, Niebur E. A model of saliency-based visual attention for rapid scene analysis. IEEE Trans Pattern Anal Mach Intell. 1998;20(11):1254-1259. https://doi.org/10.1109/34.730558
41. Montabone S, Soto A. Human detection using a mobile platform and novel features derived from a visual saliency mechanism. Image Vis Comput. 2010;28(3):391-402. https://doi.org/10.1016/j.imavis.2009.06.006
42. Jain S, Yarlagadda P, Jyoti S, Karthik S, Subramanian R, Gandhi V. ViNet: pushing the limits of visual modality for audio-visual saliency prediction. In: 2021 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Prague, Czech Republic, 2021. New York, NY: IEEE; 2021:3520-3527. https://doi.org/10.1109/iros51168.2021.9635989
43. Clark LA, Cuthbert B, Lewis-Fernández R, Narrow WE, Reed GM. Three approaches to understanding and classifying mental disorder: ICD-11, DSM-5, and the National Institute of Mental Health's Research Domain Criteria (RDoC). Psychol Sci Public Interest. 2017;18(2):72-145. https://doi.org/10.1177/15291006177272
44. Insel T, Cuthbert B, Garvey M, *et al*. Research Domain Criteria (RDoC): toward a new classification framework for research on mental disorders. Am J Psychiatry. 2010;167(7): 748-751. https://doi.org/10.1176/appi.ajp.2010.09091379
45. Alexander LM, Escalera J, Ai L, *et al*. An open resource for transdiagnostic research in pediatric mental health and learning disorders. Sci Data. 2017;4:170181. https://doi.org/10.1038/sdata.2017.181
46. Townsend L, Kobak K, Kearney C, *et al*. Development of three web-based computerized versions of the kiddie schedule for affective disorders and schizophrenia child psychiatric diagnostic interview: preliminary validity data. J Am Acad Child Adolesc Psychiatry. 2020;59(2):309-325. https://doi.org/10.1016/j.jaac.2019.05.009
47. Kaufman J, Birmaher B, Brent D, *et al*. Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (K-SADS-PL): initial reliability and validity data. J Am Acad Child Adolesc Psychiatry. 1997;1;36(7):980-988. https://doi.org/10.1097/00004583-199707000-00021
48. Conners CK. Conners. 3rd Edition. Torrance, CA: WPS; 2008.

49. Conners CK, Erhardt D, Sparrow E. Conners' Adult ADHD Rating Scales–Self-Report: Long Version (CAARS–S: L). Toronto: Multi-Health Systems; 2002.
50. Swanson JM, Schuck S, Porter MM, *et al*. Categorical and dimensional definitions and evaluations of symptoms of ADHD: history of the SNAP and the SWAN rating scales. Int J Educ Psychol Assess. 2012;10(1):51-70.
51. Salvucci DD, Goldberg JH. Identifying fixations and saccades in eye-tracking protocols. In:Proceedings of the Eye Tracking Research & Application Symposium, ETRA 2000, Palm Beach Gardens, Florida, USA, November 6-8, 2000. New York, NY: ACM Press; 2000.
52. Wang W, Shen J, Guo F, Cheng MM, Borji A. Revisiting video saliency: a large-scale benchmark and a new model. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, Utah, USA. New York, NY: IEEE; 2018: 4894-4903.
53. Le Meur O, Baccino T. Methods for comparing scanpaths and saliency maps: strengths and weaknesses. Behav Res Methods. 2013;45(1):251-266. https://doi.org/10.3758/s13428-012-0226-9
54. Bylinskii Z, Judd T, Oliva A, Torralba A, Durand F. What do different evaluation metrics tell us about saliency models? IEEE Trans Pattern Anal Mach Intell. 2019;41(3): 740-757. https://doi.org/10.1109/TPAMI.2018.2815601
55. Kümmerer M, Wallis TS and Bethge M, Saliency benchmarking made easy: separating models, maps and metrics, In: Proceedings of the European Conference on Computer Vision (ECCV). 2018:770–787.
56. Nyholt DR. A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. Am J Hum Genet. 2004;74(4):765-769. https://doi.org/10.1086/383251
57. Schwenzow J, Hartmann J, Schikowsky A, Heitmann M. Understanding videos at scale: How to extract insights for business research. J Bus Res. 2021;123:367-379. https://doi.org/10.1016/j.jbusres.2020.09.059
58. Sekachev B, Manovich N, Zhiltsov M, *et al*. Opencv/cvat: v1.1.0. Zenodo; 2020. Accessed June 8, 2022. https://zenodo.org/records/4009388
59. Saad JF, Griffiths KR, Korgaonkar MS. A systematic review of imaging studies in the combined and inattentive subtypes of attention deficit hyperactivity disorder. Front Integr Neurosci. 2020;14:31. https://doi.org/10.3389/fnint.2020.00031
60. McAuley T, Crosbie J, Charach A, Schachar R. The persistence of cognitive deficits in remitted and unremitted ADHD: a case for the state-independence of response inhibition. J Child Psychol Psychiatry. 2014;55(3):292-300. https://doi.org/10.1111/jcpp.12160
61. Baroni A, Castellanos FX. Neuroanatomic and cognitive abnormalities in attention-deficit/hyperactivity disorder in the era of "high definition" neuroimaging. Curr Opin Neurobiol. 2015;30:1-8. https://doi.org/10.1016/j.conb.2014.08.005
62. Durston S, van Belle J, de Zeeuw P. Differentiating frontostriatal and fronto-cerebellar circuits in attention-deficit/hyperactivity disorder. Biol Psychiatry. 2011;69(12):1178-1184. https://doi.org/10.1016/j.biopsych.2010.07.037
63. Krain AL, Castellanos FX. Brain development and ADHD. Clin Psychol Rev. 2006; 26(4):433-444. https://doi.org/10.1016/j.cpr.2006.01.005
64. Riva D, Taddei M, Bulgheroni S. The neuropsychology of basal ganglia. Eur J Paediatr Neurol. 2018;22(2):321-326. https://doi.org/10.1016/j.ejpn.2018.01.009
65. Ramtekkar UP, Reiersen AM, Todorov AA, Todd RD. Sex and age differences in attention-deficit/hyperactivity disorder symptoms and diagnoses: implications for DSM-V and ICD-11. J Am Acad Child Adolesc Psychiatry. 2010;49(3):217-228.e1-e3. https://doi.org/10.1016/j.jaac.2009.11.011
66. Hessels RS, Hooge ITC. Eye tracking in developmental cognitive neuroscience—the good, the bad and the ugly. Dev Cogn Neurosci. 2019;40:100710. https://doi.org/10.1016/j.dcn.2019.100710
67. Kong XZ, Zhen Z, Li X, *et al*. Individual differences in impulsivity predict head motion during magnetic resonance imaging. PLoS One. 2014;9(8):e104989. https://doi.org/10.1371/journal.pone.0104989