SUPPLEMENTARY METHODS

Stimuli. Sixty Scene-shuffle Videos (SV, approximately 30 seconds each) were used. To create these, we filmed thirty 30-second continuous videos with a camcorder (Sony HandyCam DCR-HC211 NTSC, 640 480 pixels, MPEG-1), set either: immobile on a tripod, to pan at a constant speed (6°/second ranging 120° back and forth horizontally), or to follow particular people at a university campus, a beach, a shopping district, a ski resort, and a desert (*filmed videos*). We also recorded ten 30-second continuous videos from television and video games (*recorded videos*). The 30 filmed videos were randomly cut to clip snippets (2 – 4 seconds uniformly distributed), yielding a total of 291 snippets, and reassembled to thirty 30-second SVs. Each SV (approximately 30 seconds) was made from 9 to 11 snippets without any temporal gap in between, and there were no more than one snippet included from the same original video. One set of 30 SVs were made from the filmed videos only. Another set of 30 SVs were made from the 10 recorded videos alone in the same way, but contained snippets of different lengths (0.5-2.5, 1-3, or 2-4 seconds). Within this set, the 10 recorded videos were cut to snippets whose length uniformly distributed from 0.5 – 2.5 seconds (200 snippets) and reassembled to create the first group of 10 SVs. A second group of 10 SVs was made with the same 10 recorded videos, but they were cut to snippets whose length uniformly distributed from 1 – 3 seconds (139 snippets). Similarly, a third group of 10 SVs had snippet lengths varied uniformly from 2 – 4 seconds (93 snippets). Our choice of snippet length (2 - 4 s) was within the range of our daily exposure to television and it enabled us to convey the relative quickness of new, novel and dynamic stimuli (television commercials have 1-2 seconds shot length (in 15- and 30-seconds commercials) in the United States [1], and Hollywood films have an average of 5 seconds shot length [2]).

This snippet length was further motivated by our previous studies of saliency effects during free viewing in normal volunteers [3]. The study investigated the role of memory in guiding attention when

watching continuous, uncut videos versus watching video snippets of 1-3 seconds in length. Perceptual memory is critical in guiding attention, and the authors suggested that perceptual memory of a scene could be quickly replaced by a new scene when the scene changes. Immediately after the scene changes, observers rely more on their bottom-up attention guidance to deploy their attention because it is faster than top-down attention [4,5]. Then, the top-down control is expected to gradually take over after the scene is recognized. Therefore, with longer snippet length, one may expect to observe stronger components of perceptual memory and top-down attention control in guiding attention, and we explicitly wanted to extract bottom-up measures.

Participants. This paper describes data collected from 21 children with ADHD, 13 children with FASD, 14 elderly PD participants, 18 control children, 18 young controls, and 24 elderly controls (see Diagnostic Criteria and Supplementary Table S1 for demographic data). More participants were recruited (3 patient groups: 32 ADHD, 22 FASD, 15 PD; 3 healthy control groups: 24 children, 18 young adults, 25 elderly) than those entered in the final analysis; an individual participant was excluded if he/she had too few (<10) valid eye traces (see Data Acquisition) or had received medication on the day of the experiment in the child participant groups (PD patients were not required to withhold medication). A few of the youngest control children and children with ADHD were excluded to match age across the 3 child groups. All participants had normal or corrected-to-normal vision, were compensated, and were naive to the purpose of the experiment. Young adult controls were not directly involved in classification. They were recruited to (1) provide an independent gaze distribution to compute group-based, young-observer similarities and (2) perform saccade selection (see Saccades Deviated from Norm).

Diagnostic Criteria. Diagnosis of PD was established by one of the co-authors, Dr. Giovanni Pari, a neurologist specializing in movement disorders, based upon the Unified Parkinson's Disease Rating

Scale. Children with ADHD were diagnosed from licensed practitioner across Ontario, Canada. Diagnosis of ADHD was confirmed and co-morbidity assessed using DSM-IV criteria and the Conner Parent's Rating Scales (CPRS) for children. Inclusion criteria for the ADHD pool included meeting DSM-IV criteria and criteria established from the CPRS. ADHD subjects were excluded if we identified the following co-morbid signs: learning disabilities resulting in delayed advancement in school, Tourette's syndrome, or bipolar disease. Children with FASD recruited in this study were previously assessed at diagnostic clinics in Ontario and in accordance with the Canadian Diagnostic guidelines [6]. The FASD group in the study contained children with one of three diagnoses (Fetal Alcohol Syndrome (FAS), partial FAS or Alcohol Related Neurodevelopmental Disorder (ARND) that fall under the FASD umbrella term. FAS requires (i) the presence of a characteristic pattern of craniofacial dysmorphologies (short palpebral fissures, smooth philtrum, thin upper lip, flattened midface); (ii) pre- and/or postnatal growth restriction; and (iii) structural and/or functional abnormalities of the CNS. A diagnosis of FAS can be made in the absence of confirmed maternal alcohol consumption. Partial FAS is the diagnosis when the presentation of the child includes some of the craniofacial and physical features, and structural and/or functional abnormalities of the CNS not explained by other causes, and confirmed maternal alcohol consumption during pregnancy. ARND is the diagnosis used when the child presents with structural and/or functional abnormalities of the CNS not explained by other causes, confirmed maternal alcohol consumption during pregnancy, but few or no physical features.

Data Acquisition. Forty SVs (30 from the *filmed videos* and 10 from the *recorded videos*) were played in a random order. Participants were allowed to rest after every 10 clips (about 5 minutes). A nine-point calibration was performed at the beginning of each session. At the beginning of each SV, participants were required to fixate on a grey cross displayed at the center of the screen, but they could then look anywhere on the screen at the beginning of a snippet. Instantaneous gaze position was

tracked by a head-mounted EyeLink II (SR Research Ltd.) from the participants' right eye, and 5,066 eye traces (137 participants × ≤40 SVs) were obtained. Eye-movement traces from the SVs made of recorded videos were not used in the present study because of their different snippet lengths, and these clips were used to explore other aspects of eye-movements unrelated to this study. The remaining 3,763 eye traces (137 participants × ≤30 SVs from filmed videos) were further analyzed, and each gaze position was classified as fixation, saccade, blink/artifact, saccade during blink, smooth pursuit, and eye tracker drift/misclassification. After excluding participants and discarding invalid eye traces (validity is described in Data Acquisition of Supplementary Methods), there were 108 participants with 2,654 valid eye traces, and 147,489 saccades were obtained from 79,574 seconds of eye-movement recording (see Supplementary Table S1).

An eye trace needed to meet the following criteria to be considered valid: (1) calibration error less than 0.75°, (2) drift correction error less than 5° at the end of an eye trace, (3) gaze position outside a 1° inside border of the screen occurred less than 20% during a trace, (4) loss of gaze tracking occurred less than 10% during a trace, and (5) maximum fixation length of a single fixation was less than 6 seconds; otherwise, the eye trace was removed (bad tracking quality).

The major causes for excluding participants were high calibration and drift error rates (17 out of 20 excluded participants) (Supplementary Table S1). For the 3 child groups, it was harder to obtain accurate calibration, which resulted in higher exclusion rate compared to the older participants. The child groups were less likely to complete the entire 20-minute-long experiment; however, as long as they had 5-minutes (10 eye traces) of valid data, they were not excluded. Although children with FASD had lower completion rate compared to ADHD and control children, children with FASD were engaged with the task. Interestingly, children with ADHD were the most difficult as far as sitting still. The lower completion rate for children with FAS (a subtype of FASD) might result from microcephaly,

which in our experiment using a head-mounted eye-tracker (EYELINK II) posed problems, because it was too big and heavy for some children. We believe that our method would perform even better by minimizing eye-tracking calibration issues, which experience has taught is possible when using newer eye-tracker models with children (e.g. EYELINK 1000 with remote setup instead of head-mounted EYELINK II). Moreover, reducing the length of the experiment would improve the completion rate as well. Finally, we were strict in our criteria for inclusion, and we rejected all clips that followed a bad calibration (10 eye traces). One could also decrease rejection rate by performing calibration more often.

Computing Saliency Maps from Stimuli. The Itti and Koch saliency model [7–9] is a biologically-inspired computation model based on feature integration theory [10] and the primate visual system. The model successfully explains human performance in visual search [8], and has been widely used in predicting human gaze in both artificial and natural scene stimuli. The model computes saliency maps for several low-level features, or combination of features, for every video frame. These saliency maps are topographic maps of conspicuity which highlight locations that may attract attention in a stimulus-driven manner. It has been shown that several regions in the brain resemble a saliency map, such as the superior colliculus [11], frontal eye fields [12], posterior parietal cortex [13], and pulvinar [14].

The saliency model first applied a linear filter of a feature on a video frame at several scales to generate multi-scale (i.e., fine to coarse) filtered maps. Fine filtered maps were then subtracted from coarse filtered maps to simulate center-surround operations in human vision and to produce feature maps. Next, these multi-scale feature maps were normalized and combined together to generate conspicuity maps in a manner that favors feature maps with sparse peak responses. Finally, conspicuity maps of different features were linearly summed together to form a saliency map [7]. A brief description of

each feature and the corresponding citation are provided as follows:

Color contrast [7]

Color red (r_n), green (g_n), and blue (b_n) from a video frame were normalized by intensity I_n (see *Intensity contrast*) to decouple hue from intensity. Next, four filters of different color were created: red ($R_n = r_n - (g_n + b_n)/2$), green ($G_n = g_n - (r_n + b_n)/2$), blue ($B_n = b_n - (r_n + g_n)/2$), and yellow ($Y_n = r_n + g_n - 2(|r_n - g_n|/2 + b_n)$).

Intensity contrast [7]

An intensity filter was calculated as $I_n = (r_n + g_n + b_n)/3$.

Oriented edges [7]

Gabor filters of 4 different orientations ($\theta = \{0°, 45°, 90°, 135°\}$) were generated to filter video frames.

Temporal flicker [9]

Flicker (F_n) was the absolute difference in Intensity between the current frame and the previous frame ($F_n = I_n - I_{n-1}$).

Motion contrast [9]

Motion contrast was computed by the Reichardt model [15]. For each scale, the motion filter $M_n(\theta) = |O_n(\theta) \bullet S_{n-1}(\theta) - O_{n-1}(\theta) \bullet S_n(\theta)|$, where $\bullet$ is the point product, $O_n(\theta)$ is the Gabor filter of orientation θ, and $S_n(\theta)$ is a spatially-shifted difference between two frames that is orthogonal to $O_n(\theta)$.

Line junctions including corners and edge crossings [16]

Junction filters were built on top of oriented Gabor filters. Four types of junction filters were created: L-junction, T-junction, X-junction, and E-junction. The L-junction filter responds to edge corners; the T-junction filter is sensitive to two perpendicular edges that only one edge ends at the intersection; similarly, the X-junction filter responds to two perpendicular edges that cross each other; the E-junction filter responds at the endpoint of an edge.

Intensity variance [16]

Local intensity variance was computed over 16×16 image patches as the following equation:

$Var(i,j) = \sqrt{\frac{\sum_i \sum_j [I(i,j) - \mu]^2}{S_p - 1}}$, where S_p is the size of the patch, and μ is the average intensity value of the image patch.

Texture contrast [16]

Texture contrast was computed as the spatial correlation between a 16×16 image patch and its neighboring patches within a specified radius (16 pixels, where 19 pixels equaled 1 degree visual angle).

$Txt = 1 - \frac{\sum_i^{N_p} \rho_{X,Y_i}}{N_p}$, where N_p is the number of patches within the radius, and

$\rho_{X,Y} = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - E^2(X)}\sqrt{E(Y^2) - E^2(Y)}}$, where X and Y are two image patches and E is the expected value. Therefore, high correlation indicates low texture contrast; low correlation indicates high texture contrast.

The choice of which features were included in the overall saliency maps was mostly due to history of model development. In the Itti and Koch saliency model developed in 1998 [7], there were only features C, I, and O in the model to analyze static images, but later on F and M [9] were added to analyze videos. Since then we have been using CIOFM as the standard measure of "salience." Recently, we added line junctions as a new feature [16]. That is why we here used both CIOFM and CIOFMJ. Features C, I, O, F, M, and J were implemented with an attempt to mimic visual processing of primates. However, texture and variance were image statistics commonly used by computer vision researchers to describe images. As C, I, O, F, M, and J were derived from a different route than variance and texture, we have not attempted to put all of them together into a single map.

Saccades after Scene Onset. Hypothesizing that some differences between clinical populations are more governed by "stimulus-driven" processes, we attempted to find the saccades that were more likely driven by bottom-up processes. We designed videos that changed scenes every 2 – 4 seconds because we assumed that the onset of a new scene would interrupt observers' top-down expectations of what might occur, and the observers may rely more on their bottom-up attention mechanism to allocate their attention [17]. Therefore, we considered the first 3 saccades individually as well as all the saccades combined to compare populations.

Saccades deviated from Norm. To reveal differences between groups, we selected saccades that were deviated from the gaze of "young adult controls", the 'norm'. Examining these deviated saccades may be more revealing in whether one group is more easily attracted to salient features than the others. Because if all saccades were included, differences between the groups may be diluted, which may not only decrease the performance of classification, but also hinder the identification of any underlying differences in attention mechanisms across groups. Hence, a spatiotemporal map of gaze distribution was first computed from the scanpaths of young controls. Next, for each participant (except young

adult controls), we discarded half of his/her saccades that were better predicted by the gaze distribution of young adult controls (see next paragraph for predictability index), and analyzed only the other half of the saccades that were more deviated. Note that this selection did not involve any of the saliency-based features, saccade dynamics, nor comparing the clinical groups against each other, but was based solely on eliminating those saccades which were most predictable given the young adult controls. Then we computed values of saliency-based and group-based features.

The predictability index was determined in the following manner. First, a spatiotemporal map of gaze distribution was computed from the scanpaths of young adult controls. Second, we extracted a map value at the saccade endpoint when a saccade was initiated and compared that to 100 randomly values sampled in the same map. Then, the predictability of this saccade was determined by the rank of the map value at saccade endpoint among the 100 randomly sampled map values. If the saccade endpoint ranked No. 1 (had the highest map value), it meant that this saccade can be easily predicted from the young adult sample. On the contrary, if the saccade endpoint ranked No. 101 (had the lowest map value), then this saccade was the most difficult to predict. This procedure was done on all saccades. Finally, we sorted all saccades of an observer by their ease of prediction (from the easiest to the most difficult), discarded the first half of saccades that were easier to predict, and analyzed only the other half that was more difficult to predict.

Computing Features. Saliency-based and group-based features were titrated by ordinal dominance analysis [18]. Both types of features reflected the correlations/similarities between gaze and maps (saliency maps for saliency-based features; gaze distribution maps of control young adults for group-based features). To compute the correlation, a map value at saccade endpoint (max value in a 2.5° circular window) was obtained when an observer initiated a saccade (Fig. 1c). One hundred map values were also randomly sampled from the same map as a baseline for comparison. These map

values were normalized to values from 0 to 1 relative to the minimum and maximum values of the map. With all the saccades, an observer histogram and a random histogram (bin size was 0.1) were generated from the normalized map values (Fig. 1d). The differences between the two histograms were summarized as an ordinal dominance curve, which is the correlation between the map and gaze. To create an ordinal dominance curve, we incremented a threshold from 0 to 1 and calculated the percentage of sampled map values in each histogram that were above the threshold ("hits"). The vertical axis of the rotated ordinal dominance curve was the percentage of "observer hits", and the horizontal axis was the percentage of "random hits". Thus, calculation of the area under the curve (AUC) shows the predictability of the maps based on observers' saccade endpoints. AUC values obtained from different feature maps typically ranged between chance (0.5) and an upper bound computed from inter-observer gaze similarity (here, around 0.88 among young adults), depending on how predictive the feature of interest was of the observer's gaze. In addition to the measure of AUC, the histograms provided the frequency that an observer looked at locations of low-/medium-/high map values.

An AUC of 0.5 means the maps predicted saccade endpoint no better than random. An AUC above 0.5 indicates the maps predict saccade endpoints above chance. If we assume a model cannot predict a human's gaze better than another group of human, then we can use inter-observer similarity as the upper bound of AUC. To do this, we took one person out of the group of control young adults, and utilized the spatiotemporal gaze distribution of the remaining young adults to predict his gaze. Next, an AUC value was obtained showing the similarity between his gaze and the rest of the group. We did this for every individual in the group of control young adults, and the averaged AUC, 0.88, was obtained.

The group-based features were derived in the same fashion as the saliency-based features. The only

difference was how maps were generated. For saliency-based features, maps were computed by several saliency models of different features, but maps for group-based features were generated by the spatiotemporal gaze distribution of the control young adults. To generate the gaze distributions, the instantaneous eye position of each young adult control was represented as a Gaussian blob (standard deviation = 2°), and combined into a single probability density map across all young adult controls. Because it took roughly 80 ms from planning to initiating a saccade, the timing of instantaneous eye positions was shifted earlier so that the gaze distribution was predictive to eye positions of other participants. Once the gaze distribution maps were generated, the group-based features were derived in the same way as the saliency-based features. The low, medium, and high salience bins for the saliency-based features correspond to low, medium, and high inter-observer similarity bins for the group-based features.

Classification and Feature Selection. The classification and feature selection workflow was inspired by microarray analysis, which has two goals: (1) classifying patients from controls by gene expression in the microarray, and (2) identifying genes relevant to the disease. One gene (similarly to one of our core features, e.g., color or motion) can express several different mRNA/proteins (similar to our sub-features, e.g., color contrast at the target of the first saccade) by alternative splicing. While microarray analysis usually has thousands of genes to be examined, we had 224 oculomotor-based, saliency-based and group-based sub-features.

For feature selection of multiple classes, we used MSVM-RFE [19], an extension of SVM-RFE [20] for feature selection of multiple classes. Note that MSVM-RFE ranked features that were most useful in differentiating all populations, rather than a pair of populations. Therefore, we looked at the weights of the features when the overall classifier reached maximum classification accuracy. However, even with using the same set of features, some features might be more important to classify a pair of clinical

groups, e.g. ADHD vs. FASD, but other features might be more important to classify another pair of populations, e.g. controls vs. children with ADHD. Hence, features with larger squared weight were considered more important in differentiating a pair of clinical groups. In any case, this ranking of features ranked features that were most useful in differentiating all groups, rather than any pair of groups.

Thirty iterations of a repeated leave-one-out bootstrap [21] were used to test the performance of each classifier that used a particular selected subset of features. The average accuracy is reported (with significant deviations quantified using one-tailed paired t-tests, df=29). Chance level was computed by training a classifier with the same bootstrap structure but with randomly permuted class labels. Because classification accuracy varied with the number of features in the process of RFE, we tested the performance of classifiers by comparing the maximum accuracy obtained by the classifier trained with true labels, to that obtained by the classifier trained with randomly permuted labels (chance level), regardless of how many features each classifier used to obtain maximum accuracy. In addition, to test whether a core feature was different between populations, all its sub-features were recruited to build a classifier, and we tested whether it performed better than the one with permuted class labels, as above.

Before training the classifier, features were normalized, and outliers of each feature were identified before feature values were standardized. If a feature value was outside the upper or lower quartile by greater than 1.5 times the inter-quartile difference, it was considered as an outlier. Then, the mean and standard deviation of that feature was calculated excluding the outliers. All the feature values were subtracted from the mean and divided by the standard deviation. Finally, standardized feature values were filtered by an arctangent function to diminish the influence of the outliers.

Direction of Effect. Once a feature (e.g. a core feature, or a feature type) was found discriminative by the classifiers, we examined the direction of differences between patients and controls by comparing

between populations the averages of the sub-features of the discriminative feature (e.g., averages of the 12 saccade amplitude sub-features). Note that these sub-features can be averaged because they were standardized into the same range. Next, the averages between two populations were tested (two-sample t-test, $p<0.01$). If they were significantly different, then the direction was reported.

Properties of Support Vector Machine – Recursive Feature Elimination (SVM-RFE). One characteristic of SVM-RFE is that it does not select redundant features. Mutual information between features is taken into account by the nature of SVM. Consider, for instance, that saccade duration, amplitude and peak velocity are highly correlated features (following the main sequence) that all exhibit significant population differences (e.g. by t-test). If features were selected based on individual significance values (e.g. p-value), then all 3 features would be selected. However, because they are highly correlated, which means highly redundant, having all features does not add much in differentiating populations. Consequently, highly correlated features will both receive smaller weights in SVM, with the one that is least noisy receiving slightly higher weight (correlated features might exhibit different noise levels as the correlation between saccade duration, amplitude and peak velocity may not be perfect because of, say, curved saccade trajectories or measurement noise). Therefore, the noisier ones will be eliminated first during RFE, and the other ones will remain in subsequent steps with other complementary features to maximize the performance of the classifier. (For more details on how SVM works, interested readers are referred to…)

SUPPLEMENTARY RESULTS

Classifying PD and Controls. Classification accuracy reached 89.6% with only 5 features in the process of feature selection (SVM-RFE). The classification accuracy along the process is plotted in Fig. S2a. While using all 224 features, the classifier performed significantly but slightly better than chance. This was probably due to over-fitting because there were more features than participants; therefore, the classifier performed well in training, but poorly in testing. Nevertheless, as the process of feature elimination went on, the classification accuracy started to increase, achieving peak performance (89.6% accuracy) with 5 features (out of 224). Subsequently, when more features were eliminated, the accuracy decreased due to too few features.

Classifying ADHD, FASD and Control Children. The classification accuracy throughout MSVM-RFE is shown in Fig. S2b. The overall classifier (ADHD vs FASD vs control children) reached the highest accuracy (77.3%) with 19 features. With these 19 features, the average 2-way classification accuracy for ADHD vs. control children was 83.3% (chance 53.8%); FASD vs. control children was 79.2% (chance 58.1%); and ADHD vs. FASD was 90.4% (chance 61.8%). When using all the features, children with ADHD or FASD were harder to differentiate in comparison to ADHD vs control and FASD vs control. Nevertheless, as the feature selection processes went on and critical features were identified, all three classifiers' performance improved, especially the classifier for children with ADHD vs FASD.

SUPPLEMENTARY DISCUSSION OF NEUROLOGICAL IMPLICATIONS

Hypotheses and Choice of Feature Categories. We posited that the commonly observed deficits in attention allocation and oculomotor function are based on pathological neural substrates that also govern how a person naturally directs attention in an everyday environment. We predicted PD patients would show deficient oculomotor control (oculomotor-based features), weakened top-down control (group-based features), and stronger bottom-up guidance (saliency-based features) in natural viewing. Because PD patients have been shown to exhibit shorter and slower saccades to pre-defined targets [22], we expected to observe similar oculomotor deficits over our video stimuli even when there is no pre-defined target, and hence our classifiers would reveal significant differences in oculomotor-based features. PD also affects the frontal lobe and other parts of the attention network, as mentioned above, so that they appear more stimulus-driven due to weakened top-down control; hence we expected to see increased guidance of gaze towards salient stimuli (saliency-based features) as well as lowered similarities of gaze distributions between PD and controls (group-based features). With respect to ADHD, as children with ADHD are primarily deficient in frontal cortical processing, we expected that their weakened top-down control would give rise to more saliency-driven saccades (saliency-based features), and thus they would attend to different locations from controls (group-based features). Although ADHD affects the basal ganglia as well, we did not predict that oculomotor-based features would be different because previous literature shows inconsistent results in motor deficits. Finally, as FASD influences the brain globally, we anticipated impairments in both oculomotor function and attention control. First, we predicted slower saccades and longer inter-saccade intervals (oculomotor-based features). Second, due to their frontal impairments possibly affecting top-down control, we expected lower spatiotemporal correlation with young adult controls (group-based features). We also

predicted differences, though initially not in any particular direction, in bottom-up attention (saliency-based features): while weakened top-down attention control could give rise to higher reliance on bottom-up stimuli, deficits in visual sensory processing could decrease bottom-up attention process.

We constructed classifiers with each feature type to test our hypotheses. Furthermore, to derive more precise conclusions than the previous studies which had motivated our fairly coarse initial set of hypotheses, we quantified differences in individual component features for each feature type. For example, saccade amplitude and peak velocity features were considered within the oculomotor-based feature type, and color contrast, oriented edges, and motion features were considered within the saliency-based feature type. In summary, the rich information exhibited in natural scenes and the corresponding eye traces enabled us to quantify the differences in several aspects of oculomotor functions and attention allocation in one simple paradigm, and we utilized these differences between patients and controls to build a classifier that can reliably identify individual participants in different clinical groups.

Parkinson's Disease (PD). As predicted, PD patients demonstrated deficits in saccade dynamics (oculomotor-based features: e.g., shorter duration and smaller amplitude), which suggests disruptions in cortical-subcortical pathways as well as brainstem nuclei (as described in the following paragraphs). Attention allocation (group-based features) of PD patients was different from elderly controls (mixed directions), implying impairment in attention networks involving the frontal lobe, parietal cortex and basal ganglia. However, counter to our expectation that lower top-down control may give rise to higher reliance on salience, bottom-up attention of PD patients seemed to be unaffected as saliency-based features showed no differences (except texture contrast) between the two populations.

During natural viewing, PD patients demonstrated motor deficits as their saccades were of shorter amplitude and duration. Previous studies have shown that voluntary saccades of PD patients are

smaller in amplitude and slower toward pre-determined targets [22–30], while the impairment is less pronounced when their saccades are visually guided [24,26,31,32]. This motor deficit has been primarily attributed to dysfunction in the basal ganglia [26,33–35], which is heavily involved in voluntary saccade control [36]. Smaller saccade amplitude data of the PD patients was not caused by 'square wave jerks' [37–39] because only a few saccades were categorized as square wave jerks during free viewing. In addition to motor deficits, other factors could contribute to the smaller saccades in PD patients. PD patients have been shown to have reduced 'useful fields of view' (i.e., attentional processing range) [40] so that they may not have processed peripheral information as well as control subjects, and therefore their range of the next saccade location is limited. We also observed slightly longer inter-saccade intervals (two-sample t-test, $t(36)=1.72$, $p=0.09$) in PD patients, which could be due to slower visual information processing [41,42] and/or slower voluntary saccade initiation [22,43], which is shown to correlate with reduced frontal cortical activation in PD [44].

Our data suggests that top-down attentional control, but not bottom-up responding, is selectively impaired in PD patients, because the classifiers showed differences (mixed directions) in group-based features, but not in saliency-based features. This finding is consistent with earlier studies that PD patients are impaired in executive functions including top-down attention, and these deficits have been attributed to pathological changes in the striatum, and later, in frontal cortex [45]. When these deficits in top-down attention are considered in the context of task performance, PD and control participants differed behaviorally in tasks that measure response times to visual targets under manipulation of attention direction; for instance, PD patients can show enhanced attentional 'capture' effects to a visual stimulus, resulting in faster processing of subsequent targets that appear at the same location [46,47]. This behavior points to intact bottom-up processing. At the same time, PD patients can show impaired maintenance of attention to specific locations [48], and stronger 'inhibition of return' (an initial cue slows responding to targets that appear later at the same location) than controls for advanced PD

patients [49]. When one considers voluntary saccade deficits in PD, combined with increased capture of attention to transient cues, but decreased maintenance of attention to those locations, a common trend of "hyper-reflexivity" in PD in visual tasks emerges, predicting that we should have observed differences in saliency-based features. However, hyper-reflexivity (measured in many cases by pro-saccade reaction times) was not consistently reported across studies in a recent meta-analysis [50]. This study revealed that the disparities may relate, in part, to differences in target eccentricity: PD patients being faster than controls for small saccades (< 7 degrees) but slower for larger ones, which might be caused by interactions between retinal center-surround inhibition and inhibition on the superior colliculus by the basal ganglia (see Chambers & Prescott [50] for details). One might expect that hyper-reflexivity to abrupt salient scene onset might alter the overall saccade dynamics in PD in the current experiment (independent of specific saliency-based features), but PD patients were not any faster at initiating a saccade following scene change even though the average amplitude was less than 7 degrees. Therefore, it is difficult to take a hyper-reflexivity interpretation derived from cue-target tasks and apply it to the current dynamic and free viewing environment, especially when one considers that basic prosaccades are not always produced at faster latencies.

Nevertheless, our results are consistent with studies that show PD patients have little impairment in reflexive prosaccades, and it may be interpreted as unimpaired bottom-up processing. Reflexive prosaccades can be generated when incoming sensory information directly inputs to saccade motor cells in the superior colliculus (thus bypassing frontal cortex and basal ganglia circuits) [51]. Clinically, it is also interesting that "freezing symptoms" (i.e, hypokinetic movements) can be ameliorated by providing visual cues, suggesting that bottom-up processing may not only be less impaired, but can be useful to guide voluntary behavior [52]. The neuronal implications of these findings together point to a pathology in the basal ganglia that affects circuits important to voluntary movement and voluntary attentional orienting: those that include the premotor, prefrontal and motor cortices and basal ganglia

circuits. In contrast, bottom-up signals may be utilized by neural networks that are less dependent on these brain regions.

In summary, contrary to our initial prediction, bottom-up attention of PD patients seemed to be unaffected because saliency-based features showed no difference (except intensity variance and texture contrast) compared to elderly controls. While we expected PD patients to show higher correlation between gaze and visual salience (as described above), PD also impairs visual salience computation by damaging retinal, LGN and V1 processing [53,54] and reducing pattern contrast and flicker sensitivity [55,56]. Therefore, the effect of high correlation to visual salience may have been offset by low contrast sensitivity. We must also acknowledge the potential effect of medication on our results. First, Bodis-Wollner and colleagues [56] showed that PD patients displayed better contrast and flicker sensitivity when they were 'on' dopamine medication compared to 'off' medication, suggesting that PD patients will regain the sensitivity to visual salience to a certain degree on-meds. Second, L-DOPA medication (taken by most PD patients in our study) decreases error rates in the anti-saccade task and increases reaction time in the pro-saccade task [57], implying better top-down control. Moreover, L-DOPA medication might also improve flexibility in attention shifting, but have had less of an effect on maintaining attention [58], which perhaps impacts classifier performance when new scenes demanded shifting attention. The exact effects of medication on attentional selection during free viewing are unclear without performing an on/off medication experiment. Nevertheless, our results show that even for patients who were on medication, oculomotor-based and group-based features can differentiate PD from control behavior – here quantified for the first time in a natural viewing setting – and thus potential biomarkers may be revealed in free-viewing conditions.

ADHD. As predicted, no motor deficit was observed in children with ADHD (oculomotor-based features). Children with ADHD showed slightly lower similarity to young adults in gaze distribution

than control children (group-based features), but it did not approach significance (two-sample t-test, $t(37)=1.04$, $p=0.31$).

Children with ADHD also showed differences in (mixed directions) correlation to salience as their bottom-up attention was affected. The best feature in differentiating children with ADHD from controls was texture processing. This is interesting because children with ADHD have deficits in sensory processing and modulation [59–63], which influences their early development and task performance in daily life [64]. For example, they are responsive to certain food and tactile textures but they will try to avoid them (sensory avoidance) [61]. However, they can be very unresponsive to some sensory stimuli as well, so that they keep seeking more sensory experience (sensory seeking) [64]. Our results suggest that this sensitivity to textures might not be limited to the tactile domain. These children appear to be sensitive to texture salience (contrast in texture) because they made saccades more often to locations of high or low texture salience values, but not to medium values. In spite of this overall sensitivity, the actual behavior could result from both sensory seeking and sensory avoidance. For example, high texture stimuli might initially capture their attention, but later cognitive processes may initiate either salience-seeking or salience-avoiding behavior, similar to what is seen in the tactile domain. The converse explanation (e.g., initial seeking or avoidance behavior, followed by attentional capture) could be applied as well, when one considers locations of low texture salience, however this is purely speculative. Nevertheless, we have shown that children with ADHD appear to be sensitive to texture salience, because they made saccades to locations of medium salience significantly less often than control children.

Some other saliency-based features were found discriminative between children with ADHD and control children with mixed directions of the effects. We believe that it was the profile of these features, rather than the direction, that differentiated the groups, indicating the disorder impacts natural

viewing behavior in a more complicated manner than expected. For example, ADHD impairs response inhibition [65] so that children with ADHD may be more stimulus-driven; however, ADHD also delays development [66] and weakens early visual processing (e.g., color contrast sensitivity [67,68]), and thus discounts their stimulus-driven behavior during natural viewing. Nevertheless, the classifier still revealed several saliency-based features that were discriminative between children with ADHD and controls, such as color contrast and oriented edges. Oriented edges are important to the perceptual construction of the contour and shape of objects. Edge detection is achieved in the retina [69] and the visual cortex, and is related to visual acuity. However, children with ADHD can show reduced visual acuity. This deficit can be improved by psycho-stimulants that increase dopamine-levels [70], and it has been shown with fMRI that dopaminergic network activity is different in the brains of children with ADHD and control children [71] indicating that it is possible that dopaminergic changes in ADHD may impact edge detection, just as retinal dopamine depletion has been suggested to impact contrast sensitivity in PD [72]. However, these links drawn between these two disorders, dopamine, and retinal processing are speculative. We also point out that the classifier showed that edge detection was discriminative between controls and children with ADHD: the direction of effect did not reveal whether children with ADHD were more or less sensitive to edges, so it is difficult to speculate on the underlying mechanisms. To our knowledge, there are no previous studies investigating how ADHD might affect the processing of oriented edges, and thus the discovery and selection of the oriented edge feature by the classifier points to a novel finding for future investigation.

FASD. As expected, because children with FASD suffer from a global impact of alcohol on brain development, they showed significant differences ($p<0.01$) in correlation to salience (mixed directions, saliency-based features) as well as similarity in gaze distribution to young controls (slightly lower similarity ($p = 0.15$), group-based features). Deficits in visual processing are consistently observed in subjects with FASD [73,74]. These deficits could reflect problems with attention and/or processing of

sensory information (they do not appear to be due to problems with motor control), which may result in different correlation to salience observed in this study. Furthermore, deficits in top-down control can further interrupt bottom-up process as showed in pro-saccade tasks that children with FASD make more direction errors than controls in automatic reflexive saccade [75,76]. Among saliency-based features, line junction, overall salience, and texture contrast were found discriminative between children with FASD and control children, and children with FASD also showed higher correlation to texture contrast ($p<0.01$). We are not aware of any research that has attempted to break down the deficits in visual processing in children with FASD based on different domains of salient features in a visual display, but structural injury to the brain in children with FASD may lead to some of the observed deficits. Children with FASD also exhibited lower similarity to young adults' gaze distribution, as predicted, which is consistent with previous literature showing deficits in the frontal lobe for children with FASD [77].

Microcephaly is common in individuals with full FAS [78], and structural abnormalities have been observed in FASD that correlate to behavioral deficits (see review [79]). Because these characteristics are more prominently associated with FASD than ADHD, it is possible that there could be structurally based impacts on visual and perceptual processes that could explain some of the differences between the groups. In particular, structural dysmorphology and decreased brain size in FASD has been observed in the parietal and temporal lobes [80,81], which should be expected to influence bottom-up attention to visual objects. However, it has also been observed that despite smaller overall brain size, there can be increased grey matter density (with decreased white matter density) in FASD in superior-posterior temporal and inferior parietal regions, though mostly laterally in the perisylvian region [82]. A recent study also suggests that there is reduced inter-hemispheric connectivity through the corpus callosum between parietal cortices in FASD (measured by diffusion tensor imaging, and inter-hemispheric correlations in fMRI BOLD signal) [83]. In any case, these studies suggest that both dorsal

and ventral stream processes guiding visuomotor behavior might be affected in some way if these structural abnormalities do impact visual processing. Functionally, however, little research examining dysfunction in visual processing in FASD in these regions has been done, as most studies of dysfunction in fronto-parietal and fronto-temporal networks are related to primarily cognitive tasks (e.g., working memory [79], arithmetic operations [84]).

ADHD versus FASD. It is important to understand the differences between ADHD and FASD because they can present with similar symptoms clinically, but have different underlying pathologies and treatments. However, this study did not reveal differences between ADHD and FASD in each feature type or core feature alone, but we will summarize a few studies that have directly compared patients with ADHD or FASD in the context of attentional control.

Coles and colleagues [85] measured the performance of children with ADHD or FASD on four factors of attentional control: sustaining, focusing, encoding (sequential memory and learning), and shifting (flexibility and executive function), and each factor was associated with different brain regions (e.g., sustain: the prefrontal cortex, parietal cortex; shift: the frontal eye field, posterior parietal cortex). They found that children with ADHD suffered more than children with FASD in sustaining and focusing attention, and children with FASD performed worse than children with ADHD in encoding and shifting attention. We mainly discuss the shifting attention factor, because our saliency-based and group-based features were measured only when observers make an overt attention shift.

In shifting of attention, consistent with Coles et al. [85], children with FASD or ADHD exhibit different patterns of deficits in structured pro/anti-saccade tasks, suggesting different pathological effects on oculomotor behavior [65,75,76]. In the pro-saccade task, children with FASD had longer reaction times and made more directional errors than both children with ADHD and controls, implying that children with FASD are more deficient in orienting than children with ADHD. This would suggest

that alcohol damaged the posterior parietal cortex, frontal cortex, and basal ganglia [76]. In the anti-saccade task, children with ADHD or FASD made more directional errors compared to controls, implying similar level of difficulty in inhibiting automatic response for both groups. Interestingly, the ability to inhibit impulsive responses may depend on different sub-types and event rate. In a Go/No-Go task, Kooristra et al. [86] reported that children of the ADHD-combined type made more mistakes in the slow-paced condition, and children with FASD or ADHD-inattentive type performed worse in the fast-paced condition.

However, other studies that compared ADHD and FASD directly have arrived at different conclusions, especially when patients of different subtypes or severity were recruited, or different outcome measures were used. Taking sustaining attention for example, while Coles et al. [85] reported children with ADHD were more deficient than children with FASD, other studies found no differences between ADHD and FASD [86,87]. Moreover, Kooistra et al. [88] looked into the ADHD inattentive-type, and found that their performance in a flanker task was similar to that of controls, implying no deficits in their sustaining attention.

Taken together, studies have revealed through attentional processing and response control that FASD and ADHD are difficult to segregate based on an individual parameter. Both can point to deficits in executive control over responding, as well as attentional maintenance. Thus, it fits with why, we found null results with regards to the differentiability of each feature type, or core feature alone, between ADHD and FASD. Nevertheless, this could potentially be due to the heterogeneity of the patients recruited in this study. Future studies could investigate a greater number of people with each subtype to better identify unique attention profiles.

Implications of Mixed Directions of Effects. Lack of uniformity in saliency and group-based features may reflect greater variability of responses that are typical for patient populations. Indeed, PD, ADHD,

and FASD (for anti-saccades only), have shown greater variability in saccade reaction times [22,65,75]. Other laboratories have also characterized response variability across other behaviors in ADHD using "sigma" and "tau" parameters, which relate to increased variability in ADHD in the normally distributed populations of reaction times (sigma) as well as increased skew towards a higher occurrence of longer-latency reaction times (tau) [89,90]. This tau parameter may be particularly useful as a primary characteristic of a neurological disorder, because it indicates lapses in attention.

However, it is still unknown what causes the general variability across responses, and in particular, what are the specific neuronal reasons for why this variability might be increased in a neurological condition. For example, is it that sensory information is sometimes not input into the system optimally on some trials? Or, is it that sensory information is sometimes not accessed efficiently by motor networks important for response generation? We can only speculate that because ADHD and PD impact frontal cortex and basal ganglia in particular, that it is related to this latter case (sensory information not accessed by motor networks). For FASD, it could represent a more widespread dysfunction in the brain.

Future Directions and Study Limitations. Because ADHD and FASD (and to a lesser extent PD) have very broad clinical spectra (e.g., subtypes, disease severity), our limited number of participants might not have covered the entire extent of the neurobehavioral profile. Thus, the usefulness of this approach for screening of early stage and yet undiagnosed patients remains to be proven. This likely will require a large-scale, multi-year screening study with many more participants. Our goal is to eventually run such study. Nevertheless, we believe that we have achieved an exciting proof of concept for our methodology here. Second, different medications taken by individuals further complicated the variability in our samples. We attempted to minimize this confound in the child populations (children in the ADHD and FASD groups did not take stimulant medication on the day of

experiment); however, whether the effects of medication taken in earlier days were completely 'washed out' is unclear. PD patients were not required to withhold their medications, as we, and others, have previously shown that medicated PD patients continue to display deficits in visual processing and saccade generation [22,43,91]. Nevertheless, a stricter off-medication condition should be considered in the future. Third, while our short video set was sufficient to reliably classify patient and control groups, the limited content may not identify all the features that are discriminative between the groups. It is possible that other features are discriminative with a different stimulus set or with a given viewing instruction. Finally, the comorbidity of the disorders made classification accuracy a challenge. Children with ADHD or FASD were often co-morbid with other disorders (see Supplementary Table S1). A majority of FASD children included in our analysis had a co-morbid diagnosis of ADHD, although they were still shown to be different from children with ADHD only. This is potentially very important, since one of the best practical applications of a successful classifier to different patient groups is one that can perform above chance when the data from two populations display significant overlapping symptoms. Therefore, it is apparent that the classification method utilized in this study demonstrates the robustness of its application in differentiating disorders that may have overlapping behavioral phenotypes, but that nonetheless affect visual processing differently.

Machine learning techniques could further support a better understanding of subtypes in two ways: supervised and unsupervised learning. For supervised learning, the method can give each subtype a class label and train a classifier. When the classifier attempts to differentiate each subtype, the features that are important to identify each subtype will emerge, and thus the oculomotor and attention profile of each subtype is revealed. Similarly, if there are no categorical labels, but a continuous scale such as a severity spectrum, we could learn how the profile changes across the severity through regression. For unsupervised learning, clustering methods can be applied to see whether the discovered clusters correspond to the known sub-types. If there are more clusters than number of subtypes, then

researchers can investigate whether a known subtype needs to be further divided. If there are fewer clusters than the number of subtypes, then investigators can see which known subtypes show no differences in their oculomotor and attention profiles.

Our method could assist confirmatory diagnoses. Currently, confirmatory diagnoses rarely depend on a single behavioral measure, especially for differential diagnoses. Nevertheless, our approach has the potential to be used as a large-scale screening tool, and then trigger further medical examinations and diagnosis. Furthermore, our objective, quantitative results could be integrated into expert systems that assist doctors to evaluate participants and to diagnose disorders.

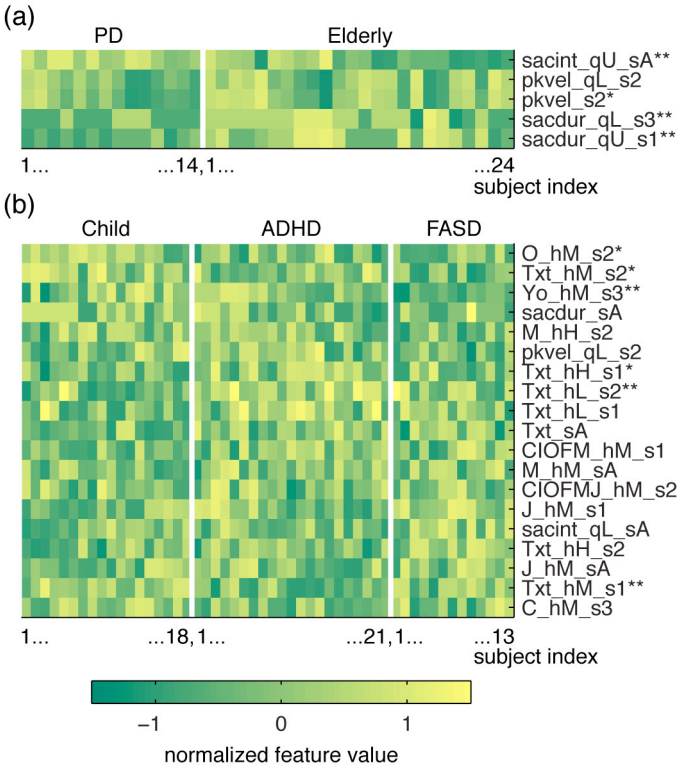
Supplementary Table S1. Demographric data of participants in analysis (after removing ineligible participants). ODD, Oppositional Defiant Disorder; DD, Developmental Delay; LD, Learning Disability; NS, Noonan's Syndrome; MMR, Mild Mental Retardation. Notes: 1 Participants were not required to finish the entire 20-minute-long experiment in order to be included. 2For the 2 child clinical populations, 'None' meant the child had never taken medicine for the disorder. If they took medicine regularly but not on the day of the experiment, they were listed in the table. If they took medicine on the day of experiment, they were removed from any analysis. For the PD population, even they took medicine on the day of experiment they were listed in the table and were included in analysis.

Category	Ctrl. Elderly	PD	Ctrl. Young	Ctrl. Child	ADHD	FASD
recruited n	25	15	18	24	32	23
unfinished 20-minute-long experiment n[1]	0	0	0	5	7	8
excluded n (<10 valid eye movement traces)	1	1	0	4	6	8
excluded n (medication on the day of experiment)	0	0	0	0	2	2
excluded n (to balance age)	0	0	0	2	3	0
n (analysis)	24	14	18	18	21	13
number of valid eye traces (time)	638 (19125.6 sec.)	357 (10720.7 sec.)	516 (15461.5 sec.)	436 (13054.1 sec.)	450 (13486.6 sec.)	257 (7725.8 sec.)
number of valid eye traces / participant	26.6	25.5	28.7	24.2	21.4	19.8
saccade frequency (number per second)	2.18	1.90	1.97	1.62	1.66	1.48
number of saccades	41774	20338	30530	21098	22347	11402
age±SD (year)	70.33±7.53	67.43±6.62	23.17±2.60	10.67±1.82	11.19±1.83	12.31±2.10
male:female	11:13	10:4	8:10	10:8	16:5	7:6
subtype					inattentive: 4 hyperactive: 0 combined: 19	FAS: 4 pFAS: 2 ARND: 7
medication[2]		None: 0 Amantadine: 1 Clonazepam: 1 Entacapone: 1 Ldopa/carbidopa: 8 Ldopa/carbidopa-CR: 1 Pramipexole: 2			None: 5 Non-stimulant: 2 Stimulant-LA: 7 Stimulant-SA: 7	None: 3 Antianxiety: 1 Anticonvulsant: 1 Antidepressant: 2 Antipsychotic: 7 Antihypertensive: 2 Non-stimulant: 1 Stimulant-LA: 6

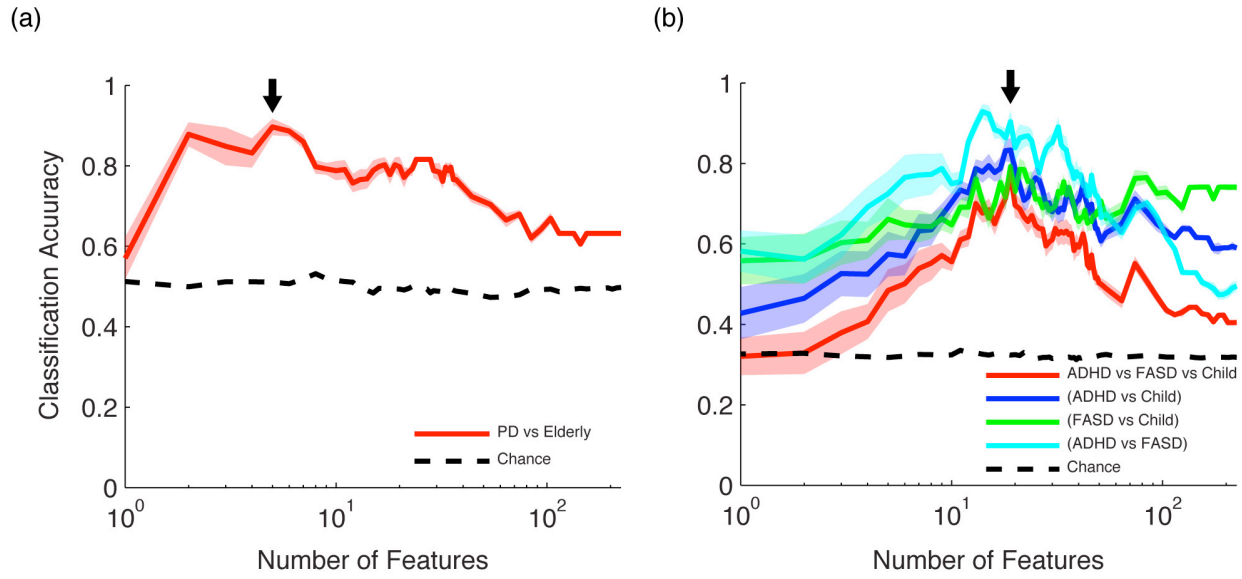
		Ropinirole HCl: 9				Stimulant-SA: 2
comorbidity					LD: 3 ODD: 1 MMR: 1	ADHD: 10 Anxiety: 5 Bipolar: 1 Conduct: 1 DD: 2 Depression: 1 ODD: 3 NS: 1 LD: 2
Clinical rating scores		UPDRS motor: 25.17±7.22 Hoehn and Yahr stage: 2.42±0.36				

Supplementary Table S2. Look up table for feature abbreviation, whose format is [feature]_[measure]_[saccade]. *after onset of snippets

category	feature	measure	saccade
Oculomotor-based	pkvel (peak velocity) sacamp (saccade amplitude) sacdur (saccade duration) sacint (saccade interval)	qL (lower quartile) none (median) qU (upper quartile)	
Saliency-based	C (color contrast) I (intensity contrast) O (oriented edges) F (temporal flicker) M (motion contrast) J (line junction) Txt (texture contrast) Var (intensity variance) CIOFM (overall salience) CIOFMJ (overall salience and line junction)	hL (low salience bin) hM (medium salience bin) hH (high salience bin) None (AUC)	sA (all saccades) s1 (1st saccade*) s2 (2nd saccade*) s3 (3rd saccade*)
Group-based	Yo (similarity to young-observer)	hL (low similarity bin) hM (medium similarity bin) hH (high similarity bin) None (AUC)	



Supplementary Fig. S1. A supplementary figure related to Fig. 4 showing selected sub-feature names. (a) Normalized values for sub-features selected by SVM-RFE while classifying PD and elderly controls. **, ANOVA $p<0.01$; *, ANOVA $p<0.05$. (b) Normalized feature values for features selected by MSVM-RFE while classifying ADHD, FASD, and control children. **, ANOVA $p<0.01$; *, ANOVA $p<0.05$.



Supplementary Fig. S2. (a) Classification accuracy for differentiating PD and elderly controls, plotted as a function of the number of selected features during SVM-RFE. Maximum classification accuracy (89.6%) was obtained with 5 features (black arrow). Shaded region indicates mean ±1 standard deviation over the repeated leave one out bootstrap validation. Chance performance (classification accuracy with permuted class labels, ~52.0%) is indicated by the dashed curve. (b) Classification accuracy for differentiating ADHD, FASD, and control children is plotted as a function of the number of selected features during MSVM-RFE. The red line indicates the overall classification accuracy for differentiating 3 populations, and the classifier reaches peak performance (77.3%) with 19 features (black arrow). The blue, green, and cyan lines are classification accuracies for classifying each pair of populations. The black dashed line is the chance level for the overall classifier. Shaded regions indicate mean ± 1 standard deviation. Chance performance (~32.4%) is indicated by the dashed curve.

SUPPLEMENTARY REFERENCES

1. Bordwell D: Intensified Continuity Visual Style in Contemporary American Film. Film Quarterly 2002 Mar;55:16-28.

2. Cutting JE, DeLong JE, Brunick KL: Visual activity in Hollywood film: 1935 to 2005 and beyond. Psychology of Aesthetics, Creativity, and the Arts 2011;5:115-125.

3. Carmi R, Itti L: The role of memory in guiding attention during natural vision. Journal of vision 2006 Jan;6:898-914.

4. Henderson J: Human gaze control during real-world scene perception. Trends in Cognitive Sciences 2003 Nov;7:498-504.

5. Wolfe JM, Alvarez GA, Horowitz TS: Attention is fast but volition is slow. Nature 2000 Aug 17;406:691.

6. Chudley AE, Conry J, Cook JL, Loock C, Rosales T, LeBlanc N, et al.: Fetal alcohol spectrum disorder: Canadian guidelines for diagnosis. Canadian Medical Association journal 2005 Mar 1;172:S1-S21.

7. Itti L, Koch C, Niebur E: A model of saliency-based visual attention for rapid scene analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence 1998 Nov 12;20:1254-1259.

8. Itti L, Koch C: Computational modelling of visual attention. Nature reviews Neuroscience 2001 Mar;2:194-203.

9. Itti L, Dhavale N, Pighin F: Realistic Avatar Eye and Head Animation Using a Neurobiological Model of Visual Attention; in Bosacchi B, Fogel DB, Bezdek JC (eds): Proc. SPIE 48th Annual International Symposium on Optical Science and Technology. Bellingham, WA, SPIE Press, 2003, pp 64-78.

10. Treisman AM, Gelade G: A feature-integration theory of attention. Cognitive psychology 1980;12:97–136.

11. Kustov AA, Robinson DL: Shared neural control of attentional shifts and eye movements. Nature 1996 Nov 7;384:74-7.

12. Thompson KG, Schall JD: Antecedents and correlates of visual detection and awareness in macaque prefrontal cortex. Vision research 2000 Jan;40:1523-38.

13. Colby CL, Goldberg ME: Space and attention in parietal cortex. Annual review of neuroscience 1999 Jan;22:319-49.

14. Robinson DL, Petersen SE: The pulvinar and visual salience. Trends in neurosciences 1992 Apr;15:127-32.

15. Reichardt W: Evaluation of optical motion information by movement detectors. J Comp Physiol A 1987 Sep;161:533-547.

16. Li Z, Itti L: Saliency and Gist Features for Target Detection in Satellite Images. IEEE transactions on image processing : a publication of the IEEE Signal Processing Society 2010 Dec 13;

17. Carmi R, Itti L: Visual causes versus correlates of attentional selection in dynamic scenes. Vision research 2006 Dec;46:4333-4345.

18. Bamber D: The area above the ordinal dominance graph and the area below the receiver operating characteristic graph. Journal of mathematical psychology 1975;12:387–415.

19. Zhou X, Tuck DP: MSVM-RFE: extensions of SVM-RFE for multiclass gene selection on DNA microarray data. Bioinformatics 2007 May;23:1106-1114.

20. Guyon I, Weston J, Barnhill S, Vapnik V: Gene selection for cancer classification using support vector machines. Machine Learning 2002;46:389-422.

21. Jiang W, Simon R: A comparison of bootstrap methods and an adjusted bootstrap approach for estimating the prediction error in microarray classification. Statistics in medicine 2007;26:5320–5334.

22. Chan F, Armstrong IT, Pari G, Riopelle RJ, Munoz DP: Deficits in saccadic eye-movement control in Parkinson's disease. Neuropsychologia 2005;43:784-796.

23. Shibasaki H, Tsuji S, Kuroiwa Y: Oculomotor abnormalities in Parkinson's disease. Arch Neurol 1979 Jun;36:360-364.

24. White OB, Saint-Cyr JA, Tomlinson RD, Sharpe JA: Ocular motor deficits in Parkinson's disease. II. Control of the saccadic and smooth pursuit systems. Brain : a journal of neurology 1983 Sep;106 (Pt 3):571-87.

25. Crawford TJ, Henderson L, Kennard C: Abnormalities of nonvisually-guided eye movements in Parkinson's disease. Brain 1989 Dec;112 Pt 6:1573-1586.

26. Briand KA, Strallow D, Hening W, Poizner H, Sereno AB: Control of voluntary and reflexive saccades in Parkinson's disease. Experimental brain research 1999 Nov;129:38-48.

27. Nakamura T, Bronstein AM, Lueck C, Marsden CD, Rudge P: Vestibular, cervical and visual remembered saccades in Parkinson's disease. Brain 1994 Dec;117 Pt 6:1423-1432.

28. Vermersch a I, Rivaud S, Vidailhet M, Bonnet a M, Gaymard B, Agid Y, et al.: Sequences of memory-guided saccades in Parkinson's disease. Annals of neurology 1994 Apr;35:487-90.

29. Kitagawa M, Fukushima J, Tashiro K: Relationship between antisaccades and the clinical symptoms in Parkinson's disease. Neurology 1994 Dec;44:2285-2289.

30. Kimmig H, Haussmann K, Mergner T, Lucking CH: What is pathological with gaze shift fragmentation in Parkinson's disease? Journal of neurology 2002 Jun;249:683-692.

31. Rascol O, Clanet M, Montastruc JL, Simonetta M, Soulier-Esteve MJ, Doyon B, et al.: Abnormal ocular movements in Parkinson's disease. Evidence for involvement of dopaminergic systems. Brain : a journal of neurology 1989 Oct;112 (Pt 5:1193-214.

32. Vidailhet M, Rivaud S, Gouider-Khouja N, Pillon B, Bonnet AM, Gaymard B, et al.: Eye movements in parkinsonian syndromes. Annals of neurology 1994;35:420–426.

33. Kori A, Miyashita N, Kato M, Hikosaka O, Usui S, Matsumura M: Eye movements in monkeys with local dopamine depletion in the caudate nucleus. II. Deficits in voluntary saccades. The Journal of neuroscience : the official journal of the Society for Neuroscience 1995 Jan;15:928-941.

34. Fukushima J, Fukushima K, Miyasaka K, Yamashita I: Voluntary control of saccadic eye movement in patients with frontal cortical lesions and parkinsonian patients in comparison with that in schizophrenics. Biological psychiatry 1994 Jul 1;36:21-30.

35. Pierrot-Deseilligny C, Milea D, Müri RM: Eye movement control by the cerebral cortex. Current opinion in neurology 2004 Feb;17:17-25.

36. Munoz DP, Everling S: Look away: the anti-saccade task and the voluntary control of eye movement. Nature reviews Neuroscience 2004 Mar;5:218-28.

37. Rascol O, Sabatini U, Simonetta-Moreau M, Montastruc JL, Rascol A, Clanet M: Square wave jerks in parkinsonian syndromes. Journal of neurology, neurosurgery, and psychiatry 1991 Jul;54:599-602.

38. Averbuch-Heller L, Stahl JS, Hlavin ML, Leigh RJ: Square-wave jerks induced by pallidotomy in parkinsonian patients. Neurology 1999 Jan 1;52:185-188.

39. Shaikh AG, Xu-Wilson M, Grill S, Zee DS: "Staircase" square-wave jerks in early Parkinson's disease. The British journal of ophthalmology 2011 May;95:705-709.

40. Uc EY, Rizzo M, Anderson SW, Qian S, Rodnitzky RL, Dawson JD: Visual dysfunction in Parkinson disease without dementia. Neurology 2005 Dec;65:1907-1913.

41. Lieb K, Brucker S, Bach M, Els T, Lucking CH, Greenlee MW: Impairment in preattentive visual processing in patients with Parkinson's disease. Brain : a journal of neurology 1999 Feb;122 Pt 2:303-313.

42. Johnson AM, Almeida QJ, Stough C, Thompson JC, Singarayer R, Jog MS: Visual inspection time in Parkinson's disease: deficits in early stages of cognitive processing. Neuropsychologia 2004;42:577-583.

43. Amador SC, Hood AJ, Schiess MC, Sereno RI, B A: Dissociating cognitive deficits involved in voluntary eye movement dysfunctions in Parkinson's disease patients. Neuropsychologia 2006;44:1475-1482.

44. Rieger JW, Kim A, Argyelan M, Farber M, Glazman S, Liebeskind M, et al.: Cortical functional anatomy of voluntary saccades in Parkinson disease. Clinical EEG and neuroscience : official journal of the EEG and Clinical Neuroscience Society (ENCS) 2008 Oct;39:169-74.

45. Leh SE, Petrides M, Strafella AP: The neural circuitry of executive functions in healthy subjects and Parkinson's disease. Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology 2010 Jan;35:70-85.

46. Fielding J, Georgiou-Karistianis N, White O: The role of the basal ganglia in the control of automatic visuospatial attention. Journal of the International Neuropsychological Society : JINS 2006 Sep;12:657-67.

47. Pollux PM, Robertson C: Voluntary and automatic visual spatial shifts of attention in Parkinson's disease: an analysis of costs and benefits. Journal of clinical and experimental neuropsychology 2001 Oct;23:662-670.

48. Wright MJ, Burns RJ, Geffen GM, Geffen LB: Covert orientation of visual attention in Parkinson's disease: an impairment in the maintenance of attention. Neuropsychologia 1990;28:151-159.

49. Briand KA, Hening W, Poizner H, Sereno AB: Automatic orienting of visuospatial attention in Parkinson's disease. Neuropsychologia 2001;39:1240-1249.

50. Chambers JM, Prescott TJ: Response times for visually guided saccades in persons with Parkinson's disease: a meta-analytic review. Neuropsychologia 2010 Mar;48:887-899.

51. Munoz DP: Commentary: saccadic eye movements: overview of neural circuitry. Progress in brain research 2002 Jan;140:89-96.

52. Oliveira RM, Gurd JM, Nixon P, Marshall JC, Passingham RE: Micrographia in Parkinson's disease: the effect of providing external cues. Journal of Neurology, Neurosurgery & Psychiatry 1997 Oct 1;63:429-433.

53. Bodis-Wollner I: Neuropsychological and perceptual defects in Parkinson's disease. Parkinsonism & related disorders 2003 Aug; 9 Suppl 2:S83-9.

54. Castelo-Branco M, Mendes M, Silva F, Massano J, Januario G, Januario C, et al.: Motion integration deficits are independent of magnocellular impairment in Parkinson's disease. Neuropsychologia 2009 Jan;47:314-320.

55. Regan D, Maxner C: Orientation-selective visual loss in patients with Parkinson's disease. Brain : a journal of neurology 1987 Apr;110 Pt 2:415-432.

56. Bodis-Wollner I, Marx MS, Mitra S, Bobak P, Mylin L, Yahr M: Visual dysfunction in Parkinson's disease. Loss in spatiotemporal contrast sensitivity. Brain : a journal of neurology 1987 Dec;110 Pt 6:1675-98.

57. Hood AJ, Amador SC, Cain AE, Briand KA, Al-Refai AH, Schiess MC, et al.: Levodopa slows prosaccades and improves antisaccades: an eye movement study in Parkinson's disease. Journal of neurology, neurosurgery, and psychiatry 2007 Jun;78:565-570.

58. Cools R: Dopaminergic modulation of cognitive function-implications for L-DOPA treatment in Parkinson's disease. Neurosci Biobehav Rev 2006;30:1-23.

59. Ermer J, Dunn W: The sensory profile: a discriminant analysis of children with and without disabilities. The American journal of occupational therapy : official publication of the American Occupational Therapy Association 1998 Apr;52:283-90.

60. Mangeot SD, Miller LJ, McIntosh DN, McGrath-Clarke J, Simon J, Hagerman RJ, et al.: Sensory modulation dysfunction in children with attention-deficit–hyperactivity disorder. Developmental Medicine and Child Neurology 2001 Jun 12;43:399.

61. Dunn W, Bennett D: Patterns of sensory processing in children with attention deficit hyperactivity disorder. Occupational Therapy Journal of Research 2002;4-15.

62. Parush S, Sohmer H, Steinberg A, Kaitz M: Somatosensory functioning in children with attention deficit hyperactivity disorder. Developmental medicine and child neurology 1997 Jul;39:464-8.

63. Parush S, Sohmer H, Steinberg A, Kaitz M: Somatosensory function in boys with ADHD and tactile defensiveness. Physiology & behavior 2007 Mar 16;90:553-8.

64. Dunn W: The impact of sensory processing abilities on the daily lives of young children and their families: A conceptual model. Infants and Young Children 1997;9:23-35.

65. Munoz DP, Armstrong IT, Hampton K a, Moore KD: Altered control of visual fixation and saccadic eye movements in attention-deficit hyperactivity disorder. Journal of neurophysiology 2003 Jul;90:503-14.

66. Karatekin C: Eye tracking studies of normative and atypical development. Developmental Review 2007 Sep;27:283-348.

67. van der Stelt O, van der Molen M, Boudewijn Gunning W, Kok A: Neuroelectrical signs of selective attention to color in boys with attention-deficit hyperactivity disorder. Brain research Cognitive brain research 2001 Oct;12:245-64.

68. Jonkman LM, Kenemans JL, Kemner C, Verbaten MN, van Engeland H: Dipole source localization of event-related brain activity indicative of an early visual selective attention deficit in ADHD children. Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology 2004 Jul;115:1537-1549.

69. Werblin FS: The retinal hypercircuit: a repeating synaptic interactive motif underlying visual function. The Journal of physiology 2011 Aug 1;589:3691-702.

70. Martin L, Aring E, Landgren M, Hellström A, Andersson Grönlund M: Visual fields in children with attention-deficit / hyperactivity disorder before and after treatment with stimulants. Acta ophthalmologica 2008 May;86:259-64.

71. Vaidya CJ, Austin G, Kirkorian G, Ridlehuber HW, Desmond JE, Glover GH, et al.: Selective effects of methylphenidate in attention deficit hyperactivity disorder: a functional magnetic resonance study. Proceedings of the National Academy of Sciences of the United States of America 1998 Nov 24;95:14494-9.

72. Bodis-Wollner I: Retinopathy in Parkinson Disease. Journal of neural transmission (Vienna, Austria : 1996) 2009 Nov;116:1493-501.

73. Simmons RW, Thomas JD, Levy SS, Riley EP: Motor response selection in children with fetal alcohol spectrum disorders. Neurotoxicology and teratology 2006;28:278-85.

74. Burden MJ, Westerlund A, Muckle G, Dodge N, Dewailly E, Nelson CA, et al.: The effects of maternal binge drinking during pregnancy on neural correlates of response inhibition and memory in childhood. Alcoholism, Clinical and Experimental Research 2011 Jan;35:69-82.

75. Green CR, Munoz DP, Nikkel SM, Reynolds JN: Deficits in eye movement control in children with fetal alcohol spectrum disorders. Alcoholism, clinical and experimental research 2007 Mar;31:500-11.

76. Green CR, Mihic AM, Brien DC, Armstrong IT, Nikkel SM, Stade BC, et al.: Oculomotor control in children with fetal alcohol spectrum disorders assessed using a mobile eye-tracking laboratory. The European journal of neuroscience 2009 Mar;29:1302-9.

77. Riley EP, McGee CL: Fetal alcohol spectrum disorders: an overview with emphasis on changes in brain and behavior. Experimental biology and medicine (Maywood, NJ) 2005 Jun;230:357-65.

78. Jones KL, Smith DW: Recognition of the fetal alcohol syndrome in early infancy. Lancet 1973 Nov 3;302:999-1001.

79. O'Hare ED, Lu LH, Houston SM, Bookheimer SY, Mattson SN, O'Connor MJ, et al.: Altered frontal-parietal functioning during verbal working memory in children and adolescents with heavy prenatal alcohol exposure. Human brain mapping 2009 Oct;30:3200-8.

80. Sowell ER, Thompson PM, Mattson SN, Tessner KD, Jernigan TL, Riley EP, et al.: Regional brain shape abnormalities persist into adolescence after heavy prenatal alcohol exposure. Cerebral cortex (New York, NY : 1991) 2002 Aug;12:856-65.

81. Archibald SL, Fennema-Notestine C, Gamst A, Riley EP, Mattson SN, Jernigan TL: Brain dysmorphology in individuals with severe prenatal alcohol exposure. Developmental medicine and child neurology 2001 Mar;43:148-54.

82. Sowell ER, Thompson PM, Mattson SN, Tessner KD, Jernigan TL, Riley EP, et al.: Voxel-based morphometric analyses of the brain in children and adolescents prenatally exposed to alcohol. Neuroreport 2001 Mar 5;12:515-23.

83. Wozniak JR, Mueller BA, Muetzel RL, Bell CJ, Hoecker HL, Nelson ML, et al.: Inter-Hemispheric Functional Connectivity Disruption in Children With Prenatal Alcohol Exposure. Alcoholism, Clinical and Experimental Research 2011 Feb 8;

84. Meintjes EM, Jacobson JL, Molteno CD, Gatenby JC, Warton C, Cannistraci CJ, et al.: An FMRI study of number processing in children with fetal alcohol syndrome. Alcoholism, Clinical and Experimental Research 2010 Aug;34:1450-1464.

85. Coles CD, Platzman KA, Raskind-Hood CL, Brown RT, Falek A, Smith IE: A comparison of children affected by prenatal alcohol exposure and attention deficit, hyperactivity disorder. Alcoholism, clinical and experimental research 1997 Feb;21:150-61.

86. Kooistra L, Crawford S, Gibbard B, Ramage B, Kaplan BJ: Differentiating attention deficits in children with fetal alcohol spectrum disorder or attention-deficit-hyperactivity disorder. Developmental medicine and child neurology 2010 Feb;52:205-11.

87. Nanson JL, Hiscock M: Attention Deficits in Children Exposed to Alcohol Prenatally. Alcoholism: Clinical and Experimental Research 1990 Oct;14:656-661.

88. Kooistra L, Crawford S, Gibbard B, Kaplan BJ, Fan J: Comparing Attentional Networks in fetal alcohol spectrum disorder and the inattentive and combined subtypes of attention deficit hyperactivity disorder. Developmental neuropsychology 2011 Jan;36:566-77.

89. Leth-Steensen C, King Elbaz Z, Douglas VI: Mean response times, variability, and skew in the responding of ADHD children: a response time distributional approach. Acta Psychologica 2000 May;104:167-190.

90. Vaurio RG, Simmonds DJ, Mostofsky SH: Increased intra-individual reaction time variability in attention-deficit/hyperactivity disorder across response inhibition tasks with different cognitive demands. Neuropsychologia 2009 Oct;47:2389-96.

91. Cameron IGM, Watanabe M, Pari G, Munoz DP: Executive impairment in Parkinson's disease: response automaticity and task switching. Neuropsychologia 2010 Jun;48:1948-1957.