

Girsa: Direction-Based Pro- and Antisaccade Assessment for Scalable ADHD Research Using Web and Phone Cameras

Kris Jamsa*

Ph. D, MBA, USA.

***Corresponding author:** Kris Jamsa, MBA, USA.

ARTICLE INFO

Received: 📅 October 06, 2026

Published: 📅 October 16, 2026

Citation: Kris Jamsa. Girsa: Direction-Based Pro- and Antisaccade Assessment for Scalable ADHD Research Using Web and Phone Cameras. Biomed J Sci & Tech Res 67(1)-2026. BJSTR. MS.ID.010444.

ABSTRACT

Background: Laboratory eye tracking provides detailed spatial and temporal measurements, but access requirements can restrict participation. GIRSA addresses a simpler behavioral question: did the participant look toward a target or away from it? This distinction creates an opportunity to investigate attention-deficit/hyperactivity disorder (ADHD) using cameras available in compatible computers and mobile devices.

Objective: To explain the implemented GIRSA direction-assessment workflow, examine the potential for observing and timing antisaccade corrections, and describe how larger samples could support associations with independently established ADHD diagnoses.

Methods: This methodological report combines a documented browser-based protocol with worked sampling and statistical examples. It defines participant-level directional outcomes, explains standard error in clinical terms, distinguishes recorded corrections from laboratory saccade timing, and proposes a staged association study.

Analytical findings: A gaze direction sustained for 700 milliseconds offers approximately 21 sampling opportunities at 30 frames per second and 42 at 60 frames per second. Visible look-toward responses followed by look-away responses may also support an approximate correction interval when both states are adequately sampled. For an average participant error percentage, standard error decreases with the square root of participant count. In a hypothetical comparison, doubling the spread of participant measurements requires four times as many independent participants to recover the same precision. These are analytical illustrations, not observed GIRSA performance results.

Conclusion: GIRSA makes an important research proposition testable: sufficiently informative directional observations from a larger population may provide more precise and more broadly applicable evidence than a smaller, higher-precision study. The proposed clinical connection is to evaluate associations with ADHD diagnosis and symptoms, then determine whether GIRSA adds useful information for clinician consideration.

Keywords: ADHD; Prosaccade; Antisaccade; Webcam Eye Tracking; Directional Errors; Correction Latency; Standard Error; Remote Research

Introduction: A Different Measurement Question

GIRSA is not designed primarily to determine the exact point on a screen at which a person is looking. Its principal question is whether the eyes move toward a displayed target or in the opposite direction. A calibrated research eye tracker can estimate gaze location and movement dynamics much more finely, but that additional detail is not required for every question about behavior. Distinguishing a large leftward response from a large rightward response is a different

measurement problem from reconstructing a precise gaze trajectory. Laboratory and webcam methods can therefore contribute at different levels of detail. [1,2] Consider a target that appears on the right. In a look-toward trial, a rightward response follows the instruction. In a look-away trial, a leftward response follows the instruction. For GIRSA's directional endpoint, the essential distinction is right versus left relative to the starting gaze position and target side, not whether the participant fixates an exact coordinate or achieves a precise mirror-image amplitude. The camera must resolve the eye sufficiently

to distinguish those responses; it need not locate every fixation with laboratory precision. This choice of measurement is central to GIRSA's purpose. A participant may be able to complete a short browser-based session but be unable or unwilling to travel to a specialized laboratory. Remote participation could extend recruitment, allow repeated observation, and include communities with limited local research infrastructure.

The relevant comparison is sometimes a consumer-camera observation versus no structured observation, rather than two equally accessible instruments. The resulting research question is specific: can widely available cameras preserve enough information about toward-versus-away responses, and sometimes their visible corrections, that larger and more varied samples improve what researchers can learn about ADHD? The present report explains how to test that proposition. It separates implemented direction scoring, proposed correction endpoints, original analytical examples, and future clinical association studies. It does not report a new participant cohort.

ADHD and Clinical Assessment

The Scale of the Problem and Common Symptoms

ADHD affects a substantial population. The US Centers for Disease Control and Prevention reports that approximately 7 million US children aged 3-17 years, or 11.7%, had a current ADHD diagnosis in the 2024 National Survey of Children's Health, based on parent reports. A separate US survey conducted in October–November 2023 estimated that 15.5 million adults, or 6.0%, had a current diagnosis based on self-report. These are estimates of reported diagnoses in specified US populations and years; they are not global prevalence estimates or measurements of everyone with undiagnosed ADHD. [3,4] The main symptom domains are inattention and hyperactivity-impulsivity. Everyday examples include difficulty sustaining attention, organizing or completing tasks, following instructions, keeping track of belongings, remaining still, waiting a turn, and resisting an immediate response. Individuals may have predominantly inattentive symptoms, predominantly hyperactive-impulsive symptoms, or both. Symptoms and their expression can change with age, and clinically important difficulties affect every day functioning rather than representing an occasional lapse [5].

How Clinicians Establish the Diagnosis

ADHD diagnosis integrates history, symptoms, functional impairment, and information about behavior in different settings. For children and adolescents, the American Academy of Pediatrics guideline calls for documentation that diagnostic criteria are met, symptoms and impairment occur in more than one major setting, and alternative explanations are considered. Information commonly comes from the patient, parents or guardians, teachers, and other involved professionals. Clinicians also consider developmental history and coexisting conditions. [6] No single computerized or eye-movement test

supplies this clinical history. Sleep difficulties, anxiety, depression, learning problems, and other conditions can resemble or coexist with ADHD. Hearing and vision assessment can also contribute when excluding other explanations. This is particularly relevant to an ophthalmic reader: difficulty seeing a target, understanding the task, or producing a usable eye recording must be distinguished from difficulty following an inhibitory instruction [7].

Where Computerized Assessments Can Contribute

Computerized tasks can add standardized observations of what a person does during a defined activity. Depending on the test, these observations may concern responses, omissions, response timing, or motor activity. They complement, rather than replace, the broader clinical account. GIRSA's proposed contribution is a different, eye-directed source of information: how often the participant follows a look-toward or look-away instruction, whether a visible initial error is corrected, and how long the observed correction takes when timing is supportable. These could become additional data points for clinician consideration if studies establish their reliability, relationship to clinically assessed ADHD, and added usefulness. The immediate research goal is to establish those relationships, not to assign an ADHD diagnosis from the camera recording.

What GIRSA Measures and How the Implementation Works

Prosaccades, Antisaccades, and Visible Error Correction

A saccade rapidly redirects the eyes. A prosaccade task asks the participant to look toward a target. An antisaccade task asks the participant to suppress that response and look in the opposite direction. The antisaccade instruction therefore adds a requirement for voluntary response control. A participant may initially look toward the target and then move away, producing a corrected directional error. This sequence allows the initial response and subsequent recovery to be studied separately. [2] There is a clinical research rationale for doing so. A systematic review and meta-analysis of pediatric studies found more antisaccade direction errors in ADHD groups than in comparison groups, while noting that the evidence across other eye-movement measures was heterogeneous and insufficient for broad conclusions. This supports a focused hypothesis about direction errors; it does not establish that every person with ADHD has abnormal eye movements or that a direction error is specific to ADHD. [8] GIRSA translates that hypothesis into an accessible task. For an antisaccade target on the right, the recording may show an initial leftward response, a rightward response followed by a leftward correction, or a rightward response without an observed correction during the available window. These patterns are potentially more informative than a single final-position label. In particular, a final look-away response does not by itself establish that the participant avoided an initial error.

Documented Browser Workflow

The implementation described in the supplied technical documentation uses a browser and the device’s front-facing camera. Its sequence includes participant positioning, nine calibration targets, five separate validation targets, baseline viewing, instructions, and 48 pro- and antisaccade trials. Green targets indicate look toward; red targets indicate look away. With the documented timings and an illustrative mean fixation interval of 1.4 seconds, the nominal sequence lasts about three minutes. Actual duration varies with participant interaction and task timing. The documented GIRSA capture path requests 60 frames per second (fps) and records the frame rate reported by the active video track. A request is not proof that every frame arrives at that rate: browser constraints express requested conditions, while track settings describe the selected configuration. Decoded frame timestamps, missing or repeated frames, and usable eye images establish the sampling actually available for analysis. A continuous recording is transferred in approximately 30-second segments for ordered reconstruction; segment length is independent of camera frame rate. [9,10] Person-specific calibration relates eye-region measurements to displayed locations, and separate validation targets provide a check using observations not used to fit the mapping. The documented mapping uses polynomial features with ridge regularization, a fitting method that discourages unstable coefficients.

For direction scoring, the important signal is simpler: a normalized iris-position measure compared with the participant’s own pre-target baseline. Changes relative to eye width help reduce dependence on how large the face appears in the image. The analyzer applies short-window smoothing, estimates a baseline from approximately 350 milliseconds before target onset, and accepts a horizontal direction only after a threshold-crossing shift persists across two consecutive frames. This is a sustained-direction rule, not a requirement to reconstruct the complete movement between positions. Calibration can assist positioning and quality assessment without making exact screen location the primary research outcome. These implementation descriptions follow the documented web client dated 23 September 2026 and direction-analysis module dated 30 August 2026.

The documented composite averages the prosaccade and antisaccade direction-accuracy percentages with equal weight. For example, 90% and 70% produce a composite of 80%. Within that existing score, a scorable trial with no qualifying response is incorrect; an unscorable trial is handled separately. The research measures below retain more detail than this composite.

Directional Outcomes and Their Denominators

A useful primary research endpoint is the observed antisaccade direction-error percentage: the percentage of antisaccade trials with a classifiable directional response in which the first detected qualifying direction is toward the target. The phrase first detected matters. It identifies what the recording and algorithm show, rather than claiming that no briefer movement occurred between frames.

Participant direction-error percentage = 100 × (observed look-toward errors ÷ trials with a classifiable first direction)

For example, six observed look-toward errors among 24 classifiable antisaccade responses produce 100 × 6 ÷ 24 = 25%. The numerator counts the observed wrong-direction responses. The denominator counts the trials on which an initial qualifying direction can be classified. The resulting 25% describes this participant’s task performance. It is neither a 25% camera error rate nor a 25% probability that the participant has ADHD. No detected response and inadequate recording quality are separate outcomes. A trial with adequate video but no qualifying shift is recorded as no detected response; a trial with inadequate observations is unscorable. Neither is silently relabeled as a correct look-away response. Their counts accompany the directional percentage, and a participant with no classifiable directional responses has an undefined percentage, not zero errors. Consequently, this research percentage is not always the complement of the existing antisaccade-accuracy component. Direction scoring is implemented. The correction definitions here specify candidate research endpoints; their exact event rules and timestamp conventions must be matched to the study software release and evaluated. A visible correction is a recorded sequence, not evidence of conscious error awareness (Table 1).

Table 1: Candidate GIRSA research endpoints and their interpretation.

Endpoint	What the number of desribers	Connection to an ADHD research question
Observed antisaccade direction-error percentage	First detected look-toward responses ÷ classifiable initial directions, expressed as a percentage.	Are these observed errors more frequent in an independently diagnosed ADHD group?
Prosaccade performance	Correct look-toward responses, with its denominator and no-response count reported.	Is an observed difficulty more pronounced for look-away than look-toward instructions?
Visible correction proportion	Detected errors followed by a qualifying look-away response ÷ detected errors with adequate follow-up.	Do groups differ in the proportion of visible errors that are corrected within the task window?
Observed correction interval	Elapsed milliseconds from the first qualifying look-toward state to the subsequent qualifying look-away state.	Among measurable corrections, do timing distributions relate to diagnosis or symptom severity?
Recording and response completeness	Unscorable trials, no detected responses, and usable follow-up counts, with explicit denominators.	Does an apparent association persist when differences in observability are accounted for?

Sampling and Correction Timing: what the Camera Can Observe

A Fast Movement Can Leave a Sustained, Observable Response

Frames per second means how many camera images are sampled in one second. A millisecond (ms) is one-thousandth of a second. At a regularly delivered 30 fps, frames are approximately 33.3 ms apart; at 60 fps, they are approximately 16.7 ms apart. The calculation simply divides the 1,000 milliseconds in a second by the number of frames.

Time between frames in milliseconds = 1,000 ÷ frames per second

The distinction that favors GIRSA is between seeing the eye during the entire movement and seeing the direction in which it subsequently rests. Imagine that a target is displayed for 950 ms, the par-

ticipant changes gaze after 250 ms, and that direction persists for the remaining 700 ms. The transition could fall between frames, but the following leftward or rightward state can still appear in many frames.

Expected sampling opportunities = frames per second × duration in seconds

For the 700-ms example, 700 ÷ 1,000 = 0.7 seconds. At 30 fps, 30 × 0.7 = 21 sampling opportunities are available; at 60 fps, 60 × 0.7 = 42. This is why a camera can potentially answer toward or away without capturing the full trajectory. These counts describe regular sampling, not a guarantee that every frame contains an interpretable eye image. Original sampling arithmetic. All durations are documented task timings or explicitly hypothetical examples. Fractional counts express average opportunities over different alignments between an event and the sampling schedule; an actual recording contains a whole number of frames (Table 2).

Table 2: What regular 30-fps and 60-fps sampling provides.

Sampling quantity	30 fps	60 fps
Time between frames	33.3 ms	16.7 ms
Opportunities during a 950-ms target window	28.5, approximately 28–29	57
Opportunities during a sustained 700-ms response	21	42
Opportunities during a sustained 150-ms wrong-direction state	4.5, approximately 4–5	9

Antisaccade Correction Adds a Second Observable Behavior

An antisaccade correction can be described informally as, ‘I looked at the target when I was supposed to look away, then redirected my eyes. The recording cannot establish that the participant consciously noticed the mistake. It can, however, potentially show a target-directed state followed by an opposite-direction state. GIRSA could therefore investigate both whether an error was visible and whether the eyes subsequently moved in the instructed direction. Here, observed correction interval means the elapsed time from the first qualifying look-toward state to the later qualifying look-away state. This is an operational, camera-derived form of correction timing. It is different from target-onset-to-correction latency, and it is not automatically the laboratory intersaccadic interval or the duration of the corrective saccade itself. Those measurements use different start and end events.

Observed correction interval = time of qualifying look-away state – time of qualifying look-toward state.

Suppose a 30-fps recording first shows a qualifying rightward error at 200 ms and a qualifying leftward correction at 433.3 ms. The observed correction interval is 433.3-200 = approximately 233 ms. The timestamps in this example are assigned to the first frame in each

direction’s accepted two-frame pair, rather than to the later confirmation frame. A study should use the same timestamp convention for both states. The record might therefore say: target on the right; first detected response toward the target; subsequent response away from the target; observed correction interval approximately 233 ms. That is a potentially useful behavioral description even though the camera does not reveal every millisecond of the two movements. Because both timestamps come from the same video, a constant offset between target presentation and the video clock cancels from their difference. Irregular frame timing and direction-detection delays still require evaluation.

How Frame Rate Affects the Chance of Seeing a Brief Error

A correction can be measured only if the preceding error is observed. To explain that dependence, consider a simplified camera that samples perfectly regularly, never drops a frame, and can classify direction whenever it samples the relevant state. Assume the start of a wrong-direction state is randomly positioned between camera frames. Let D be how many milliseconds that state remains visibly classifiable. D is not the duration of the saccade; it is the time during which the eyes provide the wrong-direction evidence the algorithm needs. At 30 fps, frames are 33.3 ms apart. A wrong-direction state lasting 20 ms occupies about 20 ÷ 33.3 = 60% of one sampling inter-

val. Depending on when it starts, it either contains one sampled frame or falls entirely between frames. Its chance of containing at least one sample is therefore approximately 60% in this idealized model. At 60 fps, the 16.7-ms gap is shorter than the 20-ms state, so at least one sampling opportunity is always present under the same assumptions.

Chance of at least one sample = the smaller of 100% and $[100 \times D \div \text{frame interval}]$

The implemented direction rule requires two consecutive frames, so one sample is not enough. With regularly spaced samples, a state shorter than one frame interval cannot contain two samples. Once it lasts between one and two intervals, the chance rises from zero to 100%. Once it lasts at least two intervals, two sampling opportunities are assured in this idealized model.

Chance of at least two samples = $[100 \times (D \div \text{frame interval} - 1)]$, limited to the range 0%–100%

For example, at 30 fps a 50-ms wrong-direction state lasts about 1.5 frame intervals. Subtracting one leaves 0.5, or a 50% chance of two samples. At 60 fps the same state lasts three intervals, so the two-sample opportunity is 100%. The calculation explains why higher frame rate can reveal more brief, corrected errors even when both cameras can observe a longer final look-away response. These are derived probabilities of sampling a sustained, classifiable state, not measured GIRSA detection probabilities. They assume a uniformly random event start relative to regular instantaneous samples, no dropped frames, and no image or classification failures. Smoothing, exposure, thresholding, and the need to observe both directions can reduce complete correction detection (Table 3).

Table 3: Ideal sampling opportunities for a visible wrong-direction state.

Dura- tion D	30fps: at least 1 sample	30fps: at least 2 sample	60fps: at least 1 sample	60fps: at least 2 sample
20 ms	60%	0%	100%	20%
50 ms	100%	50%	100%	100%
100 ms	100%	100%	100%	100%

What Approximate Correction Timing Means

Frame spacing also limits how tightly a directional transition can be located in time. In the previous 30-fps example, suppose the frame at 166.7 ms is still central and the first qualifying toward frame is at 200 ms. The toward transition lies somewhere between those frames. If the frame at 400 ms is still toward and the first qualifying away frame is at 433.3 ms, the away transition lies in the second interval. The difference between the two Underlying directional-transition times could then be approximately 200-267 ms, rather than exactly 233 ms. This approximately ± 33.3 -ms range is a sampling-only bracket under the stated ideal conditions, not a measured confidence in-

terval or a validated GIRSA timing accuracy. At 60 fps, the analogous bracket is approximately ± 16.7 ms. Real uncertainty can be wider because gaze transitions are not instantaneous, frames have exposure time, and smoothing or threshold rules affect when a direction is accepted. The practical opportunity is to evaluate whether intervals at this level of resolution preserve informative group or within-person differences. Each correction analysis should retain the actual time-stamps, frame gaps, accepted direction states, and follow-up window. A detected error without a correction before the window ends means no correction observed within that window, not that the participant never corrected. A blink or recording gap that obscures follow-up makes correction status uncertain. Participants with no visible corrected errors have no measured correction interval; assigning them zero milliseconds would confuse absence of a measurable event with instantaneous recovery.

Why Larger Samples May Make Noisier Measurements Valuable

Standard Error, Stated in GIRSA Terms

For GIRSA, the relevant measurement need not be an eye coordinate. It can be each participant's observed antisaccade direction-error percentage, such as six look-toward errors among 24 classifiable look-away trials, or 25%. Researchers then average those participant percentages within a defined group. Standard error describes how much that group average would typically vary if the study were repeated with different, comparable participants from the same population. A smaller standard error means a more stable estimate of the group's average observed error percentage. It does not mean that every person's recording is more accurate. [11] For example, suppose a study estimates a group's average observed direction-error percentage as 25%. A second, comparable sample might average 23%, another 26%, and another 24%. Standard error summarizes the expected study-to-study variation of those averages; it is not the percentage of trials GIRSA misclassifies. Camera classification error and participant task errors are different quantities, even though the former can affect the latter's measured value. Standard deviation answers a different question: how spread out are the individual participant percentages in this sample? Some participants may have few direction errors and others many. Standard deviation describes that person-to-person spread, including measurement variability. Standard error describes uncertainty in the average calculated from those people [11].

Calculating Uncertainty in the Group Average

For independent, comparably sampled participants, the estimated standard error of the group average is the standard deviation of their recorded percentages divided by the square root of the number of participants. The square root of 100 is 10, because $10 \times 10 = 100$. The square root of 400 is 20. Thus, the denominator gets larger as more participants are included, reducing uncertainty in the average [12].

Standard error of the group average = standard deviation of participant percentages ÷ $\sqrt{N}$

N is the number of independent participants, not the number of video frames or the total number of trials. If the standard deviation of participant error percentages is 20 percentage points and N = 100, the standard error is 20 ÷ 10 = 2 percentage points. With the same spread and 400 participants, it is 20 ÷ 20 = 1 percentage point. A change from 25% to 27% is two percentage points, which is the unit used here. The calculation works because averaging reduces random fluctuation. Participants with an unusually high observed percentage are balanced, on average, by participants with an unusually low one. For this reasoning to describe an underlying behavioral average, measurement deviations must not introduce a material systematic shift, or that shift must be characterized and addressed. Even a very precise average of camera labels remains an average of camera labels until its relationship to behavior is established.

A Laboratory-Versus-Webcam Example

Table 4 compares hypothetical studies of the same partici-

pant-level direction-error outcome in comparable populations. Suppose higher-precision measurements produce a standard deviation of 10 percentage points across participant scores, while noisier webcam measurements produce a standard deviation of 20 percentage points. These are assumed total spreads, not measured GIRSA values, recruitment forecasts, or a recommended study size. Both estimates are assumed to target the same underlying mean without material systematic bias. Original hypothetical calculations assuming independent participants, comparable sampling and outcome definitions, and no material unaddressed measurement bias. Standard deviation includes genuine participant differences as well as measurement variability. With 100 participants in each study, the reference method gives the more precise average. With 400 webcam participants, the two approaches have the same standard error. With 2,500 webcam participants, the webcam study's average has the smaller standard error in this example. Nothing improved the individual camera recording; the larger number of independent participants improved the population estimate. This is the statistical argument for GIRSA's potential value.

Table 4: More participants can compensate for noisier participant measurements.

Hypothetical approach	Participants N	Standard deviation	Calculation	Standard error
Higher-precision reference	100	10 percentage points	10 ÷ $\sqrt{100}$	1.00 percentage point
Noisier webcam study	100	20 percentage points	20 ÷ $\sqrt{100}$	2.00 percentage points
Larger webcam study	400	20 percentage points	20 ÷ $\sqrt{400}$	1.00 percentage point
Expanded webcam study	1,000	20 percentage points	20 ÷ $\sqrt{1,000}$	0.63 percentage points
Expanded webcam study	2,500	20 percentage points	20 ÷ $\sqrt{2,500}$	0.40 percentage points

A method does not have to provide the most precise measurement for each person to provide the most precise estimate of a group average. If accessible participation allows a sufficiently larger usable sample, the gain in sample size can outweigh additional random measurement variability. Quadrupling the sample size halves the standard error; multiplying it by 100 reduces the standard error tenfold, provided the spread and sampling conditions remain comparable.

Where Additional Measurement Variation Enters

A simple measurement-error model explains where the additional spread comes from. A participant's observed percentage consists of an underlying participant value plus a measurement deviation. For this simplified model, assume deviations average to zero, are independent across participants, and do not systematically depend on the participant's underlying value. The model can represent a suitably calibrated outcome; it is not an automatic property of raw binary classifications.

Observed participant value = underlying participant value + random measurement deviation

Variance is the square of standard deviation, another way to express spread. Under the assumptions just stated, the variance across observed participant values equals the genuine variance between participants plus the variance added by measurement deviations. The two components add because the model assumes they do not systematically move together.

Total observed variance = between-participant variance + measurement variance

First divide this total variance by N, then take the square root to obtain the standard error. For example, standard deviations of 10 and 20 correspond to variances of 100 and 400: the second variance is four times larger. Four times as many independent participants are therefore needed to offset it. This is the same result shown in Table 4, stated in variance rather than standard-deviation units.

Standard error = $\sqrt{[(\text{between-participant variance} + \text{measurement variance}) \div N]}$

This illustration concerns random variation around the intended outcome. If a device systematically misses short initial errors, adding

more participants does not make those errors visible. Instead, validation can identify what the observed endpoint represents, quantify device-related disagreement, and guide analyses within comparable conditions. Scale and measurement characterization work together; they solve different parts of the problem.

Interpreting an Uncertainty Range

For a sufficiently large, independent sample with an appropriate approximately normal sampling distribution, an approximate 95% confidence interval is the estimated mean plus or minus 1.96 standard errors, often rounded to two. This describes a procedure that would contain the target population mean in about 95% of repeated comparable studies under its assumptions. It is not a range containing 95% of individual patients. [11] Using a hypothetical mean of 25%, a standard error of 2 percentage points gives an approximate interval of 21.1%-28.9%. Reducing the standard error to 0.4 percentage points narrows the interval to approximately 24.2%-25.8%. The larger sample provides a clearer estimate of the group's average observed direction-error percentage. These intervals describe sampling uncertainty; they do not automatically cover unaddressed camera bias or recruitment bias.

Connecting GIRSA Measurements to ADHD Diagnosis and Symptoms

An Independent Clinical Reference

The direct clinical research question is whether GIRSA measurements differ with independently established ADHD diagnosis and whether they track symptom severity. A prospective study would collect both the GIRSA recording and a separate clinical assessment. For the initial association study, clinicians determining ADHD status would be unaware of GIRSA results, and the GIRSA scoring procedure would not use diagnostic labels. This prevents the test result from helping create the diagnosis against which it is then compared. Diagnostic information should distinguish clinician-confirmed ADHD, a completed evaluation not supporting ADHD, self-reported prior diagnosis, and not yet assessed. These categories are not interchangeable. In particular, no previous diagnosis is not equivalent to a clinical assessment ruling out ADHD. In remote settings where diagnostic assessment is not available, validated symptom questionnaires can support symptom-association research, while diagnostic claims are reserved for the clinically assessed subgroup. The study can recruit from the intended clinical population, including people referred for attention-related concerns whose eventual assessment does not support ADHD, rather than comparing only clear ADHD cases with unusually healthy volunteers.

A community comparison group can answer a complementary question. Age groups should be defined in advance, with children, adolescents, and adults evaluated within appropriate developmental and clinical frameworks. The clinical reference follows recognized assessment criteria rather than a GIRSA threshold [6,7].

Comparing Directional Outcomes Between Diagnostic Groups

Suppose, purely as an illustration, the mean observed antisaccade direction-error percentage is 35% in a clinically diagnosed ADHD group and 15% in a clinically assessed non-ADHD group. The estimated group difference is $35 - 15 = 20$ percentage points. In plain language, the average participant in the ADHD group produced more visible look-toward responses when instructed to look away. This would be an association to investigate, not evidence that every ADHD participant performs worse than every comparison participant.

Estimated diagnosis-associated difference = ADHD group mean - comparison-group mean

Each group mean has uncertainty, so the difference also has a standard error. For two independent groups, square each group's standard error, add those squares, then take the square root. This combines uncertainty from both groups instead of treating either average as perfectly known.

Standard error of the difference = $\sqrt{[(\text{ADHD-group SE})^2 + (\text{comparison-group SE})^2]}$

For example, with 100 participants in each group and a standard deviation of 20 percentage points within each, each group mean has a standard error of 2 points. The difference's standard error is $\sqrt{(2^2 + 2^2)} = \sqrt{8} =$ approximately 2.83 points. A 20-point observed difference would then have an approximate 95% confidence interval of 14.5-25.5 points, using the large-sample illustration above. The numerical values are hypothetical, not a GIRSA finding or an expected ADHD effect size.

If each group instead contained 400 comparable participants with the same spread, each group's standard error would fall to 1 point and the difference's standard error to $\sqrt{2}$, or about 1.41 points. Its approximate interval would narrow to 17.2-22.8 points. Thus, the sample-size argument applies directly to the proposed ADHD question: more participants can produce a more precise estimate of the difference in observed antisaccade performance between diagnosed and comparison groups. A narrow interval is useful whether the difference is substantial, small, or near zero. A reproducible difference would support further investigation of the eye-movement association. A well-estimated small difference would clarify where the measure contributes little. The benefit of scale is a more informative answer to the clinical research question, not a guarantee of a favorable result.

Retaining Group Differences Despite Imperfect Direction Labels

A separate example addresses the fact that direction is a category, not a continuous coordinate. Assume a hypothetical classifier reverses the true label on 10% of trials, with the same reversal rate for toward and away responses and for both clinical groups. This is a deliberately simplified model, not an estimate of GIRSA accuracy. A

true wrong-direction percentage of 35% would then be observed as 38%: 90% of the 35% true errors remain labeled errors, and 10% of the other 65% are mistakenly added to them.

$$\text{Observed error percentage} = (\text{true error percentage} \times 0.90) + (\text{true correct percentage} \times 0.10)$$

For the hypothetical ADHD group, $(35 \times 0.90) + (65 \times 0.10) = 31.5 + 6.5 = 38\%$. For the comparison group, $(15 \times 0.90) + (85 \times 0.10) = 13.5 + 8.5 = 22\%$. The true 20-percentage-point difference becomes an observed 16-point difference. The signal is reduced, but it is not erased. More generally, if q is the chance of reversing either label, the fraction of the true group difference retained is $1 - 2q$. Here $q = 0.10$, so $1 - 0.20 = 0.80$: 80% of the difference remains. The subtraction includes two kinds of reversal, because errors can be mistaken for correct responses and correct responses for errors.

$$\text{Observed group difference} = \text{true group difference} \times (1 - 2q)$$

A larger sample can make the retained 16-point difference more precisely estimated. It does not, by itself, turn 16 back into 20 or remove the shift in either group's mean. A benchmark that estimates the actual classification pattern can inform correction or sensitivity analyses, with the benchmark's own uncertainty included. This example explains both the opportunity and the distinction from the zero-mean model in Section 5: symmetric label reversals do not generally produce an unbiased observed percentage. Actual missed-error patterns may differ from this model. A camera may miss brief corrected errors more often than sustained errors, and that pattern may depend on frame rate or participant behavior. Comparing frame-rate strata and using reference measurements helps determine whether a diagnostic association reflects behavior rather than different opportunities to observe it. Once that relationship is characterized, scale can strengthen the estimate of the information that remains.

Symptoms, Correction, and Within-Person Patterns

Diagnosis is one useful outcome, but symptoms vary within and across diagnostic groups. Researchers can compare GIRSA measures with validated, age-appropriate symptom ratings obtained independently of GIRSA. A positive correlation would mean that participants with more severe rated symptoms tend, on average, to have higher measured error percentages. It would not mean that every individual follows that pattern, nor that one measurement causes the other. Inattention and hyperactivity-impulsivity ratings can be examined separately rather than assuming they relate identically to every eye measure. Correction behavior provides a second set of hypotheses. Among visible errors, do participants with ADHD differ in how often an away response follows within the observation window? Among measurable corrected trials, do their correction intervals differ in typical duration or variability? Is the relationship stronger for one symptom domain? The direction of any timing association should remain an empirical question; the present report does not assume that ADHD necessarily produces slower correction. These analyses

distinguish frequent initial errors with frequent visible recovery from frequent errors with little observed recovery. The patterns could describe different aspects of task performance that a final accuracy score obscures.

Correction timing must be interpreted alongside the number of errors and measurable corrections, because a participant contributes a correction interval only after producing an observed, subsequently corrected error. Comparing intervals alone would selectively compare those events, not all participants' inhibitory control. Repeated sessions could examine whether a participant's direction errors or correction patterns vary with symptoms over time. Sleep, fatigue, usual medication timing, practice, and recording conditions would be documented rather than assumed constant. Such observational associations do not establish a medication effect, and participation need not involve changing treatment. Participants' repeated trials and sessions remain linked to that participant in the analysis.

Competing Explanations and Added Clinical Value

A statistical adjustment asks whether an association remains when relevant differences between groups are considered. For GIRSA, a prespecified analysis could account for age, medication status, relevant visual or oculomotor findings, task understanding, site, device category, actual frame rate, and recording quality. The aim is not to add every available variable indiscriminately, but to address plausible alternative explanations identified before examining results. Comparisons within similar age and device groups can make the result easier to interpret. Trials from one participant are not independent people. Statistical models can retain trial-level information while recognizing that observations from the same person belong together. Similarly, family or recruitment-site relationships may require additional handling. In an uncertainty analysis based on resampling, whole participants-not isolated frames-would be resampled so that within-person relationships are preserved. Missing or unusable recordings would be described by clinical group rather than silently omitted. After a reproducible association is established, the next question is whether GIRSA adds useful information beyond data already available in routine assessment. A study can compare predictions of an independent clinical reference using usual questionnaire and history variables alone versus the same variables plus GIRSA features. Performance should be evaluated in people not used to select the features or fit the model, followed by replication at another site or in a different device mix.

Separate studies can then ask whether presenting GIRSA results improves the assessment process or clinical decisions. If individual diagnostic-support estimates are later developed, sensitivity would describe how often the method flags people with the reference diagnosis, and specificity how often it does not flag people without that diagnosis. Predictive value would describe what a flag means in the population where it is used; it depends partly on how common ADHD is there. These would be additional validation outcomes, not values inferred from a significant group difference. A 35% task-error score is

not automatically a 35% probability of ADHD. The clinical pathway is therefore coherent: establish what the camera measures, test whether those measurements relate to independently assessed ADHD, determine whether the information generalizes, and evaluate what it adds for clinicians. GIRSA can be valuable at the research-association stage even before it becomes an individually predictive adjunct. Its accessible delivery may make those association studies possible at a scale and frequency that laboratory-only recruitment cannot readily provide.

Research Opportunity: Accessible Participation and Focused Validation

Existing Consumer-Camera Studies Support the Underlying Approach

Published webcam research provides a credible starting point. Van der Cruyssen and colleagues replicated three established eye-tracking effects online, although the effect sizes were smaller than their reference effects. Kaduk and colleagues compared webcam and laboratory recordings simultaneously across several viewing tasks. Together, these studies support the proposition that selected behavioral information can survive reduced instrumentation; they do not establish GIRSA's performance. [1,13] Mobile pro- and antisaccade work is also relevant. Lai and colleagues used a front-facing-camera application to collect more than 235,000 eye movements from 80 adults aged 20–92 years and analyzed direction errors and latency. Their study demonstrates a practical route to repeated mobile observations while emphasizing variation between and within participants. Its acquisition and timing procedures are specific to that application [14].

A particularly relevant 2026 study by Koerner and colleagues compared simultaneous iPad and laboratory recordings in 12 people with Parkinson's disease and 19 comparison participants. The devices yielded concordant group-level conclusions across the prespecified measures in that sample. The relevance to GIRSA is methodological: a consumer-device system can be tested on whether it preserves a disease-associated group difference, not only on whether every individual measurement matches a laboratory trace. The study concerns Parkinson's disease and another system; it provides a model for evaluation, not ADHD or GIRSA validation [15].

Broader Reach is A Measurable Research Outcome

GIRSA could enable participation from home, local clinics, schools working under approved research arrangements, or assisted community sites. The opportunity is especially important when travel, cost, scheduling, or absence of nearby equipment prevents laboratory attendance. In rural or low-resource communities, a structured recording that supports an answerable question may be more useful than collecting no eye-movement data. There is precedent for adapting eye-movement research to resource-constrained settings. Forssman and colleagues studied the feasibility of eye-tracking-based cogni-

tive assessment in a rural Malawian setting using dedicated infrared equipment. GIRSA asks whether compatible consumer cameras can further reduce the equipment barrier for defined directional tasks. Local partnerships could support language adaptation, task explanation, device access, and appropriate consent [16]. An access study would measure who participates, who completes a usable session, who returns, and who remains excluded. Device ownership, connectivity, vision, digital familiarity, and language affect practical participation. Expanding the participant count and expanding population coverage are related but separate achievements. A large convenience sample is not automatically representative, yet broader and transparently characterized participation can still answer questions that a small local sample cannot.

A Staged Program that Can Begin Before Full Laboratory Benchmarking

Initial work can test calibration completion, successful video reconstruction, direction separability, repeated-session stability, and agreement with blinded reviewers on visible toward-versus-away responses. Deterministically rendered eye sequences or synthetic landmark trajectories with known directions and event times can test algorithm behavior at 30 and 60 fps, including brief errors and corrections. These engineering tests can identify failure modes and support progress before a high-speed-camera partnership is available. Reviewers of the same webcam video share its sampling limitations, and synthetic performance remains distinct from biological validation. A focused reference study can then compare the same participants and trials using GIRSA and a synchronized high-speed tracker. It would separately evaluate first detected direction, reference-defined initial direction, visible correction proportion, and the operational correction interval. The benchmark should include brief corrected errors as well as sustained responses, and a range of devices, head positions, lighting, and visual characteristics that resembles intended participation. Its output would identify conditions under which each endpoint retains useful information. The larger distributed study would connect those endpoints to independent ADHD assessments and symptom ratings using the design in Section 6 [17].

Recruitment, completion, usable observations, cost per usable session, and geographic coverage would be measured alongside association strength and uncertainty. A smaller laboratory component can thus support a larger population study without requiring every participant to attend the laboratory. The appropriate success criterion is not that a webcam reproduces every laboratory measurement. It is that a defined, reproducible GIRSA measurement answers a useful research question, with enough retained information and sufficient participant coverage to justify its use. That question may concern a diagnostic-group difference, symptom association, age-related pattern, or repeated observation. Each result would clarify what accessible eye-movement research can contribute.

Scope and Interpretation

This report presents an implemented direction-assessment approach and original analytical examples, not observed estimates of GIRSA accuracy, ADHD association, or clinical benefit. Exact gaze location, peak saccadic velocity, microsaccade measurement, and fully resolved movement trajectories are not the primary endpoints proposed here. Correction timing is a promising additional endpoint where both directional states and an adequate timebase are available; its performance is to be measured under an explicit event definition. The statistical value of scale depends on what the observations retain and whom the sample represents. Random variability can be reduced in population estimates; unaddressed systematic misclassification or selection bias cannot be cured by participant count alone. Trial completeness, observable-error duration, device differences, and diagnosis-independent scoring therefore belong in the measurement plan. Video is potentially identifiable, so appropriate ethical review, consent and assent where applicable, secure handling, retention policies, and controls on reuse are integral to distributed research.

Conclusion

GIRSA asks a focused question that commonly available cameras may be able to answer: did the participant look toward a target or away from it? Where sampling permits, a second question follows: after a visible look-toward error, did the participant redirect their eyes, and how long was the observed interval? Neither question requires exact screen coordinates or a complete high-speed account of every movement. The sample-size argument gives this approach a concrete scientific rationale. For the average participant direction-error percentage, standard error falls as the number of independent participants increases. A sufficiently larger sample can offset additional random measurement variability and provide a more precise estimate of an ADHD-associated group difference. The potential advantage lies not in making each webcam recording perfect, but in making an answerable population question more precisely and broadly investigated. The proposed clinical connection is to compare GIRSA outcomes with independently established ADHD diagnoses and validated symptom ratings, then evaluate whether those outcomes add useful information for clinicians. Laboratory reference measurements and distributed participation can serve complementary roles. For populations whose practical alternative is no structured eye-movement observation, that combination offers a meaningful research opportunity: accessible, imperfect measurements that retain enough information to improve understanding of ADHD.

References

1. Kaduk T, Goeke C, Finger H, König P (2024) Webcam eye tracking close to laboratory standards: Comparing a new webcam-based system and the EyeLink 1000. *Behav Res Methods* 56: 5002-5022.
2. Munoz DP, Everling S (2004) Look away: the anti-saccade task and the voluntary control of eye movement. *Nat Rev Neurosci* 5(3): 218-228.
3. (2026) Centers for Disease Control and Prevention. Data on ADHD in Children. Accessed.
4. Staley BS, Robinson LR, Claussen AH, Samuel M Katz, Melissa L Danielson, et al. (2024) Attention-Deficit/Hyperactivity Disorder Diagnosis, Treatment, and Telehealth Use in Adults—National Center for Health Statistics Rapid Surveys System, United States, October–November 2023. *MMWR Morb Mortal Wkly Rep* 73(40): 890-895.
5. (2026) Centers for Disease Control and Prevention. Symptoms of ADHD. Accessed.
6. Wolraich ML, Hagan JF, Allan C, Eugenia Chan, Dale Davison, et al. (2019) Clinical practice guideline for the diagnosis, evaluation, and treatment of attention-deficit/hyperactivity disorder in children and adolescents. *Pediatrics* 144(4): e20192528.
7. (2026) Centers for Disease Control and Prevention. Diagnosing ADHD. Accessed.
8. Sherigar SS, Gamsa AH, Srinivasan K (2023) Oculomotor deficits in attention deficit hyperactivity disorder: a systematic review and meta-analysis. *Eye (Lond)* 37: 1975-1981.
9. (2026) MDN contributors. Media Track Constraints: frame Rate property. MDN Web Docs.
10. (2026) MDN contributors. Media Track Settings: get Settings method. MDN Web Docs.
11. (2026) National Institute of Standards and Technology. NIST/SEMATECH e-Handbook of Statistical Methods: Glossary. Entries: standard deviation and standard error. Accessed.
12. (1996) National Institute of Standards and Technology. Standard Deviation of the Mean. *Dataplot Reference Manual*. 3: 2-52.
13. Van der Cruyssen I, Ben Shakhar G, Pertzov Y, Guy N, Cabooter Q, et al. (2024) The validation of online webcam-based eye-tracking: The replication of the cascade effect, the novelty preference, and the visual world paradigm. *Behav Res Methods*. 56: 4836-4849.
14. Lai HY, Saavedra Pena G, Sodini C, Heldt T, Sze V (2022) App-Based Saccade Latency and Directional Error Determination Across the Adult Age Spectrum. *IEEE Trans Biomed Eng*. 69(2): 1029-1039.
15. Koerner J, Zou E, Karl JA, Cynthia Poon, Roneil G Malkani, et al. (2026) iPad Eye Tracking Reproduces Clinical Grade Oculomotor Differences in Parkinson's Disease. *npj Digit Med* 9: 614.
16. Forssman L, Ashorn P, Ashorn U, Maleta K, Matchado A, et al. (2017) Eye-tracking-based assessment of cognitive function in low-resource settings. *Arch Dis Child*. 102(4): 301-302.
17. (2024) National Institute for Health and Care Excellence. Digital technologies for assessing attention deficit hyperactivity disorder (ADHD).

ISSN: 2574-1241

DOI: [10.26717/BJSTR.2026.67.010444](https://doi.org/10.26717/BJSTR.2026.67.010444)

Kris Jamsa . Biomed J Sci & Tech Res



This work is licensed under Creative Commons Attribution 4.0 License

Submission Link: <https://biomedres.us/submit-manuscript.php>



Assets of Publishing with us

- Global archiving of articles
- Immediate, unrestricted online access
- Rigorous Peer Review Process
- Authors Retain Copyrights
- Unique DOI for all articles

<https://biomedres.us/>