Received: 20 March 2023 | Revised: 28 July 2023 | Accepted: 4 October 2023

DOI: 10.1111/psyp.14461

PSYCHOPHYSIOLOGY SPR WILEY

ORIGINAL ARTICLE

Temporal and spectral electrooculographic features in a discrete precision task

Germano Gallicchio[1] | Donghyun Ryu[2,3] | Mudit Krishnani[2] | Guy L. Tasker[2] | Alessandra Pecunioso[4] | Robin C. Jackson[2]

[1]School of Psychology and Sport Science, Bangor University, Bangor, UK

[2]School of Sport, Exercise and Health Sciences, Loughborough University, Loughborough, UK

[3]Department of Physical Education, Seoul National University, Seoul, Republic of Korea

[4]Department of General Psychology, University of Padova, Padova, Italy

Correspondence
Germano Gallicchio, School of Psychology and Sport Science, Bangor University, Bangor, UK.
Email: germano.gallicchio@gmail.com

Donghyun Ryu, School of Sport, Exercise and Health Sciences, Loughborough University, Loughborough, UK.
Email: d.ryu@outlook.com

Abstract

This study aimed to evaluate the utility and applicability of electrooculography (EOG) when studying ocular activity during complex motor behavior. Due to its lower spatial resolution relative to eye tracking (ET), it is unclear whether EOG can provide valid and accurate temporal measurements such as the duration of the Quiet Eye (QE), that is the uninterrupted dwell time on the visual target prior to and during action. However, because of its greater temporal resolution, EOG is better suited for temporal-spectral decomposition, a technique that allows us to distinguish between lower and higher frequency activity as a function of time. Sixteen golfers of varying expertise (novices to experts) putted 60 balls to a 4-m distant target on a flat surface while we recorded EOG, ET, performance accuracy, and putter kinematics. Correlational and discrepancy analyses confirmed that EOG yielded valid and accurate QE measurements, but only when using certain processing parameters. Nested cross-validation indicated that, among a set of ET and EOG temporal and spectral oculomotor features, EOG power was the most useful when predicting performance accuracy through robust regression. Follow-up cross-validation and correlational analyses revealed that more accurate performance was preceded by diminished lower-frequency activity immediately before movement initiation and elevated higher-frequency activity during movement recorded from the horizontal channel. This higher-frequency activity was also found to accompany a smoother movement execution. This study validates EOG algorithms (code provided) for measuring temporal parameters and presents a novel approach to extracting temporal and spectral oculomotor features during complex motor behavior.

KEYWORDS

Electrooculography, Eye Tracking, Movement smoothness, Nested cross-validation, Quiet Eye

1 | INTRODUCTION

With the broad aim of understanding what factors characterize expertise in complex goal-directed actions, extensive interest has been directed towards the Quiet Eye period (QE; Vickers, 1996, 2007). The QE period is operationalized as the uninterrupted dwell time wherein the gaze is maintained within a certain area of interest, typically of 3° of visual angle, centered around an object or spatial location that is critical for the task.

For example, in golf putting, the critical location is the ball until the ball is struck and, after that, it becomes the location previously occupied by the ball. The QE period must start before movement initiation but can extend indefinitely as long as the gaze lies within the area of interest. Despite some contrasting findings (e.g., Harris et al., 2021; Lee, 2015), there is substantial evidence that, for various motor skills, a longer QE is associated with better performance (for the most recent meta-analysis see Lebeau et al., 2016). For example, more skilled golfers tend to maintain their gaze on the ball longer than less skilled golfers when putting (Vickers, 1992). When it comes to assessing oculomotor activity during movement, head-mounted camera-based eye-tracking glasses (ET) are the de facto standard (Vickers, 2007); however, electrooculography (EOG) is emerging as a potentially complementary tool (e.g., Gallicchio et al., 2018; Gallicchio & Ring, 2020; Gonzalez et al., 2017; Mann et al., 2011; Xu et al., 2021; Yoshikawa et al., 2019). Being relatively underdeveloped for mobile applications, the EOG poses some technical challenges as well as yet-unexplored opportunities.

The EOG detects changes to the electric dipole generated by the corneoretinal potential of each eye. As the eyes rotate within their sockets so do their dipoles. Therefore, an array of electrodes applied on the skin around the eyes can detect their rotation (Shackel, 1960; Young & Sheena, 1975). Spatially, the EOG signals provide only basic information in the form of voltage fluctuations indicating whether the eyes have rotated and, if so, in what direction (i.e., towards/away from which electrode) and by how much (i.e., how large the fluctuation). Because EOG-detected oculomotor activity is relative to the eye sockets—thus regardless of head movements—voltage fluctuations can indicate saccades or vestibulo-ocular adjustments. The former is an attention-orienting mechanism that shifts the gaze from one location to another, whereas the latter is a reflex that allows the gaze to be kept on an object or location while the head moves. The absence of prominent fluctuations indicates that the eyes are relatively still; however, it is impractical to determine the exact spatial location of the gaze, and researchers can only assume the gaze location based on task instructions. Accordingly, it is unclear whether EOG-based QE measurements are valid and accurate when compared to those obtained through ET.

Complicating matters further, the EOG confounds eye movements with other biological activity that can be detected through electrodes applied on the skin around the eyes but that is not indicative of oculomotor activity (e.g., facial and neck muscular activity, eyelid movements). The impact of these biological confounders can be lessened through appropriate skin preparation, electrode placement, and signal processing (Marmor et al., 2011). However, each algorithmic choice—for example, how to filter the signals or whether to rely on a horizontal rather than a vertical electrode montage—leads to forks in the processing path that result in many alternative measurements for each index of interest. While the large number of choices is, in principle, an opportunity to finely tune the data to be most sensitive to the phenomenon of interest, it represents, in practice, a challenge to making reliable inference.

A potential advantage of EOG is that it can record data at a much greater sampling rate (e.g., 1000 Hz in Mann et al., 2011; 2048 Hz in Gallicchio & Ring, 2020) than most mobile ET (e.g., 30 Hz in Vickers, 1996; 120 Hz in Harris et al., 2021). Not only can the resulting higher temporal resolution provide more precise temporal measurements but it can also enable spectral decomposition, a signal-processing procedure that extracts the frequency components from a time series. Because the spectral content of saccadic movements is dominated by lower frequencies and that of fixational movements is dominated by higher frequencies (Krauzlis et al., 2017), spectral decomposition may allow to evaluate which eye movements contribute to behavioral performance and quality of movement within periods of critical importance for the task. Spectral decomposition is routinely used in electrophysiology (e.g., Cohen, 2014) but its applicability to EOG signals has scarcely been explored.

This study's overarching aim was to explore the utility of the EOG when studying complex behavior such as putting a golf ball to a target. We had three objectives. First, to evaluate the validity and accuracy of EOG-based QE measurements when compared with those obtained from ET data. Second, to compare ET-based and EOG-based temporal and spectral measurements in their ability to predict behavioral indices of performance. Finally, to explore the crossmodal link between oculomotor activity and kinematics parameters of movement quality.

2 | METHOD

2.1 | Participants

Sixteen participants (14 males, 2 females) were recruited through convenience sampling among local golf clubs and within the University's student population. The participants varied in expertise, indexed by golf handicap

(a lower number corresponds with better ability) and amount of recent practice (see Figure 1). Their median age was 24.5 years (*IQR* = 6.25, *range* = [18, 36]) and their median height was 180 cm (*IQR* = 10.75, *range* = [170, 190]). All participants were right-handed putters; however, responses to the Edinburgh Handedness Inventory (Oldfield, 1971) identified 14 right-handed individuals (laterality quotient: *Mdn* = 88.85, *IQR* = 36.58, *range* = [20, 100]) and two left-handed individuals (laterality quotient: *Mdn* = −54.70, *IQR* = 25.29, *range* = [−80, −29.41]). Participants reported no neurological, cardiovascular, or ocular disorder that could not be corrected through contact lenses. Participants provided informed consent and were treated in respect of the Helsinki Declaration on Human Rights. Ethical approval was granted by the University Ethics Committee.

2.2 | Procedure and task

Participants completed a single 2-hour laboratory session and were instructed to refrain from alcohol, caffeine, and nicotine in the three hours before testing. Upon arrival in the laboratory, participants were briefed and completed the Edinburgh Handedness Inventory (Oldfield, 1971). Following the application of psychophysiological and kinematics sensors (described below), participants performed 15 practice putts and then completed the experimental task, which consisted of 60 putts separated into 4 blocks of 15 putts. The shortest interval between consecutive putts had a median of 25.77 s across participants (*IQR* = 3.09, range = [21.15, 30.90]), ensuring an adequate preparation for each putt. Prior to each block, participants completed the gaze calibration to ensure accurate ET measurements and to record the EOG data necessary for the conversion from voltages to degrees of visual angle. For more details on the calibration, see the online supplemental material.

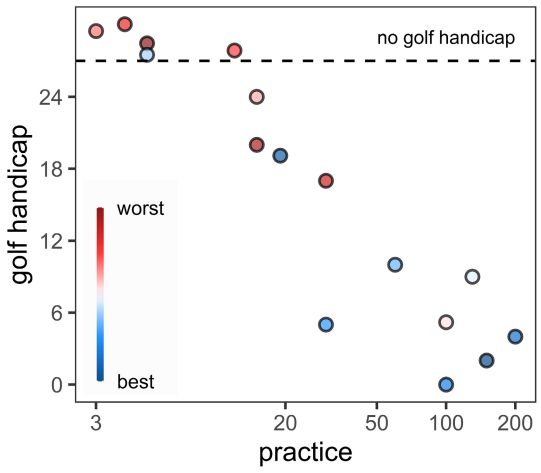

FIGURE 1 Scatterplot of self-report golf handicap (*Mdn* = 9, *IQR* = 13.55, *range* = [0, 24]) as a function of self-report recent practice, operationalized as times played golf in the past 12 months (*Mdn* = 24.5, *IQR* = 89.75, *range* = [3, 200]). Values on the *x*-axis are log spaced. Five participants reported no official golf handicap, so they were placed above a dotted line, and a small vertical jitter was added for them to avoid overlap between two observations reporting the same practice amount. Color is used to rank participants' accuracy at the putting task (median radial error) from worst (red) to best (blue).

Participants putted regular-sized golf balls (Titleist Pro V1) to a 4-m distant target on a green flat putting surface, an AstroTurf with a Stimpmeter reading of 12.46 feet corresponding to a fast speed. The target consisted of a bicolored dot sticker (1.9 cm diameter white on top of 2.5 cm diameter red). Participants were encouraged to use their own putter but were given the option to use our blade-style putter (Odyssey White Hot Pro 2.0, length = 33 inches). To promote an ecologically valid interpretation, participants were encouraged to think of the task as a lag putt with the goal to 'get the final position of the ball as close as possible to the target'. No performance feedback or comments were given during the task.

Participants were instructed to perform the following sequence of events for each putt. First, stand upright in a relaxed position and maintain their gaze on the ball positioned a step away in front of their feet. Second, upon hearing an acoustic tone (duration = 200 ms, frequency = 1200 Hz, generated by an Arduino-controlled buzzer; Arduino, Italy) step forward and adopt their usual putting stance with the putter head positioned to the right of the ball and aligned with an infrared sensor used to detect the initiation of the backswing movement. The position of this sensor was adjusted by the experimenter during the practice to fit the participant's technique. Third, rest their eyes on a red dot sticker (diameter = 0.8 cm) positioned 9 cm to the right of ball until receiving verbal confirmation to proceed by the experimenter. This fixation was used as reference to recalibrate the eye tracker online prior to each putt. Fourth, feel free to explore the visual scene and, whenever ready, execute the putt. Finally, step back to allow the experimenter to reposition the golf ball for the next putt. Figure 2 illustrates the setup as recorded by the scene camera of the mobile eye tracker.

2.3 | Data recording and processing

2.3.1 | Performance

After each putt, the experimenter took a photo using a digital camera (X20 4K 20MP Action Camera, Campark, Hong Kong) mounted on the extension pole of a tripod and positioned above the putting area (cf. Neumann & Thomas, 2008). Each photograph was processed offline in MATLAB to

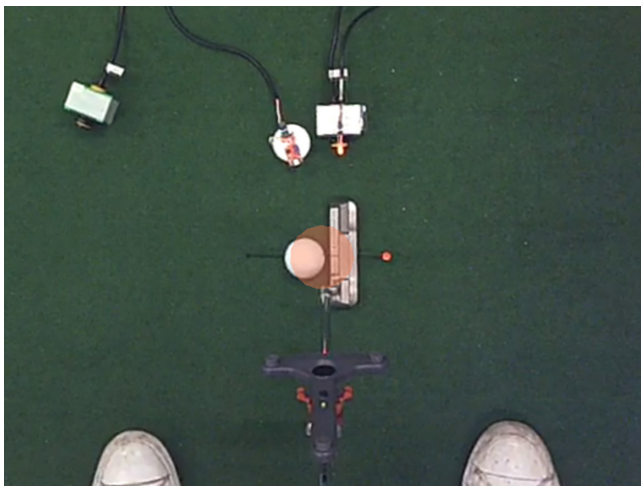

FIGURE 2 Visual scene recorded by the SMI mobile eye tracker (SensoMotoric Instruments, Germany) as the participant was about to initiate the backswing. The shaded orange circle represents the gaze location and has a diameter of 3° of visual angle. The red sticker on the right of the putter head was used as additional calibration prior to each putt. The bottom of the figure shows the SAM PuttLab sensor (Science Motion GmbH, Germany) mounted on the shaft of the putter. The top of the figure shows the three sensors used to mark the instants of movement initiation and impact on the EOG signal.

identify the positions (in cm) of the target and of the ball in a rectangular 2D space defined by the anteroposterior and mediolateral axes with origin on the target position.

Robust versions of the bivariate radial and variable errors described by Hancock et al. (1995) were obtained by replacing the mean with the median in the formulas. Namely, radial and variable errors were computed, separately for each participant, respectively as $\mathrm{Mdn}\left(\sqrt{(x_i-x_0)^2+(y_i-y_0)^2}\right)_t$ and $\sqrt{\mathrm{Mdn}\left((x_i-x_c)^2+(y_i-y_c)^2\right)_t}$ where x_i and y_i denote the coordinates of the final ball position for the trial i, x_0 and y_0 denote the coordinates of the target position, x_c and y_c denote the coordinates of the centroid computed as $\left(\mathrm{Mdn}(x_i)_t, \mathrm{Mdn}(y_i)_t\right)$, and $\mathrm{Mdn}(.)_t$ denotes the median across all trials. In lay terms, radial and variable errors can be respectively interpreted as accuracy and consistency. Upon a preliminary exploration of the dataset, we found that radial and variable errors were highly and positively correlated (Spearman's $\rho = .94$, $p < .001$), indicating that participants who were on average more accurate were also more consistent. Therefore, the analyses reported in the following sections were conducted on radial error only, after applying a log transformation to unskew the data distribution.

2.3.2 | Eye tracking

Binocular eye movements were recorded through the mobile SMI system (SensoMotoric Instruments, Germany), composed of ETG 2w wearable glasses interfaced via USB with a laptop running the iViewETG software. This system tracked head-centered binocular positional eye movements through corneal reflection (sampling rate: 60 Hz) and overlaid them on the visual-scene video (sampling rate: 30 Hz) to show the gaze location. To ensure an adequate mapping of the gaze location onto the visual scene, a three-point calibration was performed prior to each block of 15 putts (for more details, see the online supplemental material) and a one-point calibration was performed as an additional check prior to each putt to ensure a valid measurement. The ET data were used to score QE durations.

ET-based QE duration: QE^{ET}

The onset and offset of the QE period were identified through the frame-by-frame inspection of the raw camera data showing the gaze location superimposed onto the visual scene, using the BeGaze software (SensoMotoric Instruments, Germany). The gaze location was marked by a semi-transparent circular cursor with diameter of 3° of visual angle (see Figure 2). The diameter of the circular cursor was determined empirically for each block of 15 putts by matching its size with that of calibration circles with diameters depending on the participant's eye-floor distance (for details, see the online supplemental material). The scoring procedure consisted of adding in-software annotations corresponding with the timings of (a) movement initiation, (b) QE onset, and (c) QE offset. Movement initiation was defined as the earliest clear indication of the putter head moving backward. QE onset and offset were identified as the farthest time points—respectively, preceding and following movement initiation—with the cursor overlapping any part of the golf ball. These annotations were exported to MATLAB to compute the durations (in seconds) of QE_{pre} as the interval between QE onset and movement initiation, QE_{post} as the interval between movement initiation and QE offset, and QE_{total} as the interval between QE onset and offset. The total QE period was distinguished into the pre- and post-movement initiation components to take into account the different stages of movement preparation and execution. If the cursor moved away from the ball and returned to the ball within 3 frames (i.e., 99 ms) the QE period was not interrupted. Similarly, if it disappeared and reappeared on the ball, we considered it as either a blink or a temporary tracking loss, and the QE period was not interrupted. If the cursor did not overlap the ball at movement initiation or if the time on the ball lasted less than 100 ms, no QE onset and offset were identified. As noted by Chia et al. (2017), there is no consensus on scoring trials with no QE period, with some discarding them and others assigning a duration of 0 s. Due to the impact that zeros can have on measures of central tendency, we chose to discard

these trials; however, we offer a comparative analysis of the two approaches in the supplemental material. Participants did not seem to always have their gaze on the ball at movement initiation: on average, the cursor overlapped the ball for 86.07% of trials ($Mdn = 94.99\%$), however, there was a noteworthy variability across participants ($SD = 19.63\%$, $IQR = 19.17$, $range = [36.67, 100]\%$).[1]

2.3.3 | Electrooculography

The EOG was recorded via the ActiveTwo system (BioSemi, The Netherlands)[2] from active electrodes (i.e., amplified at the sensor level) applied to six mildly abraded skin sites near the eyes: two at the outer canthi, two at the bottom of the eyes, and two above them (on forehead sites Fp1 and Fp2 as described in the 10–20 system; Jasper, 1958). Signals were recorded in monopolar. Two scalp electrodes were used as common mode sense and driven right leg channels to enhance the common mode rejection ratio of the signal. The signal was amplified and digitized at 2048 Hz with 24-bit resolution. Because the acquisition system was battery operated and placed far from mains-powered devices, the signals were not contaminated by alternating-current power (50 Hz). No online filter was applied.

A bespoke sensor system (see Figure 2) controlled by a set of microcontrollers (Arduino Micro, Arduino, Italy) was used to send triggers (i.e., TTL pulses) to the signal acquisition system to allow the offline identification of (a) the initiation of the backswing movement and (b) the putter-ball impact during the downswing. The system's input came from two proximity switches (E18-D80NK) and one sound switch (KY-037). Each proximity switch sent a trigger whenever an object crossed or moved away from its infrared beam whose depth was adjusted through a built-in potentiometer. One of the proximity switches was positioned to the right of the ball in line with the putter head to detect the initiation of the backswing movement. The other proximity switch was positioned approximately 20–30 cm to the left of the ball so that, offline, we could confirm whether the ball had been struck following a potential backswing initiation trigger. The sound switch was positioned near the impact zone and sent a trigger when a sound exceeded a certain intensity threshold, which was manually adjusted through a built-in potentiometer. A bespoke MATLAB program was used to identify the relevant markers within the EOG recordings and label them accordingly. This setup proved highly sensitive as it detected 100% of backswing initiations and 99.60% of impacts. Across the whole dataset, only four impacts were undetected.

Monopolar channels were converted to bipolar to obtain one horizontal and one binocular vertical channel.[3] Following this transformation, voltage increases (or decreases) on the horizontal channel indicated eye movements to the left (or right), whereas voltage increases (or decreases) on the vertical channel indicated eye movements up (or down). The EOG time series were segmented from −15 to 10 s relative to movement initiation (0 s). Voltages were converted to degrees of visual angle by dividing the voltages by the scaling factor obtained for each block through the calibration procedure described in the online supplemental material (for a similar approach, see also Acuña et al., 2014, and Gonzalez et al., 2017). Following these preliminary processing steps, the EOG signals were processed separately to compute QE durations and time-frequency power, described in turn below.

EOG-based QE duration: QE^{EOG}

Each 25-s time series was linearly detrended (by subtracting the least-square line) to minimize any low-frequency drift, and the horizontal and vertical EOG time series were zeroed to the value at time 0 s. Then, a combined channel was computed as the vectorial sum of the horizontal and vertical channels by applying the Pythagorean theorem (cf. Henn & Cohen, 1973). Two algorithms, described below, were used independently to identify QE onsets and offsets. QE durations were computed as the intervals between QE onset and movement initiation (QE_{pre}), between movement initiation and QE offset (QE_{post}), and between QE onset and QE offset (QE_{total}).

[1]Interestingly, the participant with the lowest percentage of trials with gaze on the ball at movement initiation (36.67%) was an expert golfer, who reported a golf handicap of 2 and 150 practice sessions in the previous 12 months. In a conversation immediately after testing, this participant revealed that his strategy was to end the backswing/begin the foreswing on an imaginary spot ahead of his right foot to maintain consistency in the club path. The ET data confirmed that for most putts his gaze moved to a spot ahead of his right foot prior to or shortly after initiating the backswing. The interested reader can inspect the individual EOG waveforms in the supplemental material. Accordingly, for those trials with a fixation on the ball at movement initiation, his QE_{post} durations were short ($Mdn = 135.45$ ms). This participant was the best performer in terms of radial error (15.54 cm) and the second best for variable error (14.05 cm).

[2]Beside the EOG, electrophysiological signals also included EEG and ECG. These additional data were collected to address questions not directly related with oculomotor activity and have not yet been analyzed.

[3]The horizontal channel was computed as LO1 – LO2, where LO1 and LO2 indicate the left and right outer canthus, respectively. The binocular vertical channel was computed as $2^{-1} \sum(\text{Fp1} + \text{Fp2}) - 2^{-1} \sum(\text{IO1} + \text{IO2})$, where IO1 and IO2 indicate the sites below the left and the right eye, Fp1 and Fp2 indicate the left and right forehead sites, and $2^{-1} \sum(.)$ represents the mean of the two time series.

For each of the two algorithms described below, the scoring procedure yielded 240 different measurements for each QE measure (pre, post, total): 3 channels (horizontal, vertical, combined) × 8 filter options × 10 thresholds (depending on the algorithm; described below). Sample data (horizontal EOG in Figure 3) and MATLAB code for extracting QE durations from EOG signals are available in the repository described in the Data and Code Availability section.

Dispersion algorithm

This algorithm identified QE onsets and offsets by first smoothening each EOG time series through median filtering and then comparing the signal against a threshold. When computed over a time series, the median filter replaces each data point with the median of the data points within a surrounding time window. The length of this window—referred to as filter order—determines the smoothness of the resulting signal: a longer window corresponds to a larger filter order and yields a smoother waveform. Median filtering was preferred over high-pass filtering to avoid low-frequency distortions that would impair the ability to identify ocular fixations (cf. Acuña et al., 2014; Marmor et al., 2011; Yoshikawa et al., 2019). Median filters are particularly useful for highlighting stepwise features in the EOG waveforms because they reduce short-time bumps and valleys (ocular overshoots, blinks, and high-frequency activity) while preserving flat and vertical portions of the signal indicative of fixations and saccades respectively (Juhola, 1991). Figure 3 illustrates the effect of the median filter on the raw EOG signal. To avoid committing a priori to a certain filter order and to inform future similar research, we employed a range of filter orders including 64, 128, 192, 256, 384, 512, 640, and 768 points, which, with our sampling rate, corresponded respectively to window lengths of 31.25, 62.5, 93.75, 125, 187.5, 250, 312.5, and 375 ms. Each time series was zeroed to the value corresponding with movement initiation so that at time 0 s (i.e., movement initiation) the visual angle was 0°. Finally, the visual angle of each data point was evaluated against a dispersion threshold (e.g., 3°) starting from 0 s and sliding one point at the time backward for QE onset and forward for QE offset. When the algorithm detected a 100-ms window of data points entirely suprathreshold, the data point immediately preceding the farthest point of the window was marked as QE onset or offset, depending on the direction of the evaluation. To avoid committing a priori to a certain threshold and to offer an unbiased comparative evaluation, we used 10 thresholds linearly spaced from 1° to 10°. Each filter setting and threshold was applied on independent iterations of signal processing.

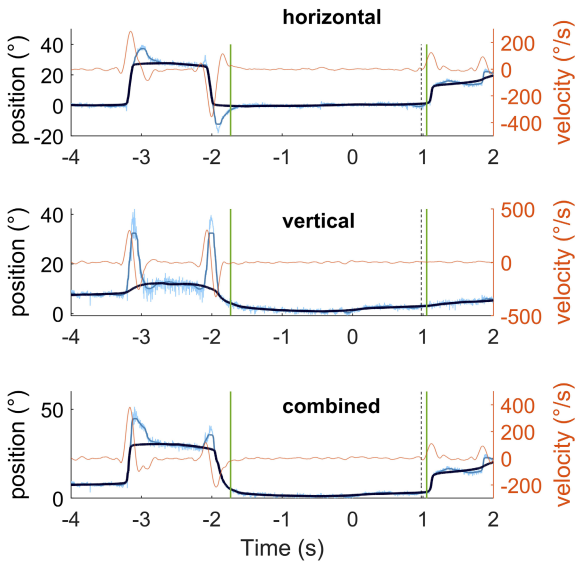

FIGURE 3 Effects of filters on saccadic EOG data. The blue lines represent the positional time series (light blue: no filter; blue: with 125-ms median filter; dark blue: with 1000-ms median filter). The red line represents the velocity time series obtained through a 5th-degree 767-point Savitzy-Golay 1st-derivative filter applied to the unfiltered signals. The two green vertical lines indicate QE^{ET} onset and offset. The black dashed vertical line indicates the putter-ball impact. Time 0 s coincided with the initiation of the backswing movement. Note that, while the time series were zeroed to the value at time 0 s, the value at time 0 s itself represented not one point but either a weighted average of the neighboring points (when using the median filter) or a derivative computed on polynomials fitted including also neighboring points (when using the Savitzky–Golay filter). The extent to which points were considered neighbors depended on the filter order.

Velocity algorithm

This algorithm identified QE onsets and offsets by first computing eye movement velocities (°/s) and then comparing the signals against a threshold. Eye movement velocities were obtained through Savitzky–Golay differential filtering, a procedure that computes the n^{th} derivative of a signal over time while minimizing differentiation noise (Savitzky & Golay, 1964). This filter performed the following steps. First, it selected a window of a certain length (i.e., filter order), fitted a polynomial of a certain degree through the least-square method, then, it computed the first derivative of that polynomial, and finally replaced the central point of the window with the central point of the differentiated polynomial. Figure 3 illustrates the effect of the differentiation filter on the raw EOG signal. To explore a wide range of settings, we used filter orders of 511, 639, 767, and 895 points (ca. 250, 312.5, 375, and 437.5 ms) for a 3rd-degree polynomial and filter orders of 767, 895, 1023, and 1151 points (ca. 375, 437.5, 500, and 562.5 ms) for a 5th-degree

polynomial. The velocity time series was zeroed to the value corresponding with movement initiation so that at time 0 s (i.e., movement initiation) the velocity was 0°/s. Finally, the velocity of each data point was compared with a threshold (described below), sliding one data point at the time backward or forward from time 0 s. When a suprathreshold velocity was detected, the time point that preceded it was marked as QE onset or offset depending on the sliding direction. We evaluated 10 velocity thresholds, logarithmically spaced between 25 and 60°/s and rounded to the nearest integer: 25, 28, 30, 33, 37, 41, 45, 49, 54, 60°/s. We probed such a wide range because of the novelty of this approach and to inform future research towards a narrower threshold range. Each filter setting and threshold was applied on independent iterations of signal processing.

Time-frequency EOG power

The time series were zeroed to the value at time 0 s, and the combined channel was computed as the vectorial sum of the horizontal and vertical channels. A Multi-Taper short-time Fast Fourier Transform (FFT) decomposition was applied to the three EOG channels (horizontal, vertical, and combined) to extract their time-frequency content (see Figure 4). In addition, the same processing—except for the preliminary voltage-to-degrees conversion described above—was performed on a bipolar mastoid channel examined for inferential control, that is to evaluate whether any potential EOG power effect was also evident from a cephalic non-EOG channel. In lay terms, this transformation enabled us to distinguish slower and faster oscillatory activity as a function of time. In more detail, 97,500-ms time windows with 87.5% overlap were identified with center points ranging from −4 to +2 s relative to movement initiation (i.e., 0 s). Each 500-ms window was mean-centered and multiplied by one of four Slepian tapers in separate iterations. No zero-padding was applied. This procedure generated, for each taper iteration, complex-valued FFT coefficients with a temporal resolution of 1/16 s (62.5 ms), a frequency resolution of 2 Hz, and a frequency precision of 10 Hz. Separately for each taper iteration, power was computed in the time-frequency plane as the elementwise product of each FFT coefficient with its complex conjugate (equivalent to amplitude squared). Power values obtained for each taper iteration were then averaged (arithmetic mean), and the values within the [−4, 2] s time interval and the [0, 100] Hz frequency interval were retained for the analyses. MATLAB coding tutorials for Multi-Taper FFT and baseline normalization are widely available in literature (e.g., Cohen, 2014; Prerau et al., 2017). Broadly speaking, there are two ways to examine power: relative to a baseline or by itself. To avoid biasing future studies towards either approach, we computed both types and evaluated them separately.

Relative power: RP^{EOG}. A baseline is often employed in physiological analyses to minimize non-task-related interindividual differences. We defined the baseline as the [−1, 0] s interval relative to the onset of the putting prompt (0 s), that is, while participants were resting their eyes on the ball prior to adopting their putting stance. The time between the putting prompt and movement initiation had a median of 13.14 s across participants ($IQR = 4.39$, $range = [9.35, 18.46]$). Power in the baseline window was computed using the same decomposition procedure described above except for using 17 overlapping windows. A decibel (dB) conversion was performed separately for each channel through the formula $10\log_{10}\frac{A_{time,freq,trial}}{B_{freq}}$, where $A_{time,freq,trial}$ identified power in each time-frequency point for each trial for the peri-movement "activity" period and B_{freq} identified "baseline" power for each frequency after that values across time and trials were median-averaged. After the dB conversion, positive (or negative) values within the activity period indicate increased (or decreased) activity in that specific time-frequency compared to the baseline period.

Absolute power: AP^{EOG}. Values that are not expressed relative to a baseline are not influenced by the signals recorded during a separate interval. However, they are potentially biased by interindividual differences. For example, small phasic changes may be masked by tonically large or small values when running between-participant statistical analyses. To address this bias, we computed the Mdn-scaled log transformation (cf. Gallicchio et al., 2016), whereby each participant's values were first scaled by the median computed across all time points and trials separately for each frequency and channel, and then submitted to a $10\log_{10}$ transformation. After this transformation, positive (or negative) values indicate greater (or less) power, scaled in a way that each participant has an equal number of positive and negative values across all frequencies and channels.

2.3.4 | Kinematics

The SAM PuttLab system (version 2010, Science Motion GmbH, Germany) was used to record the three-dimensional kinematics of the golf club following the laser-based calibration protocol described by Marquardt (2007). The acquired data were available for 95.42% of trials

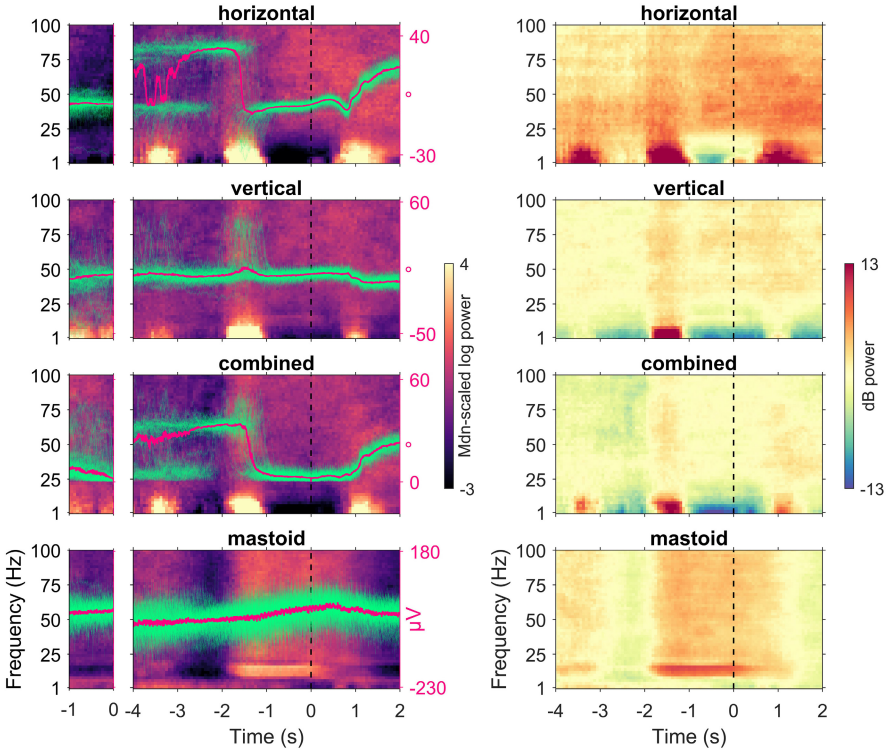

FIGURE 4 Spectrograms showing EOG time-frequency power and amplitude for each channel (horizontal, vertical, combined, and mastoid). The left panels show Mdn-scaled log power (AP^{EOG}) in the baseline and peri-movement period. Note that, for visual purposes, median-scaling in this figure was performed across all trials and time points from the *pooled* baseline and peri-movement periods. However, to keep the peri-movement period independent from the baseline in the analyses, median-scaling was performed on the peri-movement period only. The right panels show dB power (RP^{EOG}) in the peri-movement period relative to the baseline. Each time-frequency value is the median across trials. Time 0 s indicates the putting prompt for the baseline and the movement initiation for the putting period. On the left panels, the thick magenta line represents the median EOG amplitude across trials, whereas the thin green lines represent each trial's amplitude (as degrees of visual angle for the EOG channels and as μV for the mastoid channel). This figure shows one participant's data for illustrative purposes; however, all participants' data are reported in the online supplemental material.

($Mdn = 100\%$, $IQR = 1.65\%$, $range = [56.67,\ 100]\%$). The system uses small lightweight ultrasonic transmitters attached to the club shaft (see Figure 2) and sends pulses to a receiver hosting a set of microphones. The acquired positional data have a spatial resolution of ca. 0.1 mm and a temporal resolution of 210 Hz. The key events—backswing initiation, top of backswing, and end of downswing movement—were automatically identified by the SAM PuttLab software. The time series for velocity, acceleration, and jerk (i.e., the first, second, and third derivative of position with respect to time) in the mediolateral axis were submitted to a kernel-based differentiation filter using filter orders respectively of 90, 110, and 130 (Marquardt & Mai, 1994). The resulting time series (time resolution of 70 Hz) were further processed in MATLAB to compute two indices of movement quality related to smoothness: log dimensionless jerk (LDLJ) and spectral arc length (SPARC). Both indices were computed separately during the backswing phase (i.e., from movement initiation to top of backswing) and the downswing phase (i.e., from top of backswing to end of swing after the putter-ball impact).

LDLJ defines smoothness by quantifying the extent to which movement acceleration changes across time by the temporal analysis of jerk. We computed it as $-\log\left(\frac{(t_2-t_1)^3}{s^2_{peak}}\int_{t_1}^{t_2} j(t)^2 dt\right)$, where $\int_{t_1}^{t_2} j(t)^2 dt$ indicates the integral of the squared jerk time series $j(t)$ over the time points in the $[t_1, t_2]$ interval, $\frac{(t_2-t_1)^3}{s^2_{peak}}$ is a normalization factor that makes LDLJ dimensionless with s^2_{peak} being the maximum speed in the $[t_1, t_2]$ interval (Hogan & Sternad, 2009), and log(.) being a natural log transformation used to increase the index's sensitivity by reducing ceiling effects (Balasubramanian et al., 2012).

SPARC (Balasubramanian et al., 2015) quantifies the complexity of the velocity profile as the arc length of its frequency representation. We computed it as $-\int_0^{\omega_c}\sqrt{\left(\frac{1}{\omega_c}\right)^2+\left(\frac{d\hat{V}(\omega)}{d\omega}\right)^2}d\omega$, where $\hat{V}(\omega)$ is the velocity spectral amplitude $V(\omega)$ scaled by its largest value, $V(\omega)$ is the spectral amplitude of velocity over the angular frequencies ω and obtained through FFT, with the data being zero-padded to reach a length of $2^{\lceil \log_2(n)+4 \rceil}$ time points with n indicating the number of time points in the velocity

time series within the interval of interest, $\log_2(.)$ being the base-2 log transformation, the number 4 being a constant, and $\lceil . \rceil$ indicating the ceiling function, $\frac{d\widehat{V}(\omega)}{d\omega}$ is the derivative of $\widehat{V}(\omega)$ with respect to ω, ω_c is the largest angular frequency so that $\omega \leq 2\pi$ 10 Hz and $\widehat{V}(\omega) \geq 0.05$ with 10 Hz and 0.05 being constants, and $\int_0^{\omega_c} . d\omega$ is the integral over the $[0, \omega_c]$ angular frequency interval.

For both LDLJ and SPARC, we used the Riemann sum computed through the trapezoidal method to approximate the integrals. Both measures are unitless, and the negative sign at the beginning of each formula makes them directly proportional to movement smoothness (i.e., larger values correspond to greater smoothness).

2.4 | Data analysis

We adopted a stepwise approach. First, we evaluated validity and accuracy of the EOG-based QE durations by comparing them to those obtained through ET. Second, we used nested cross-validation to rank the five techniques describing oculomotor activity—eye-tracking-based QE (QE^{ET}), EOG-based QE with dispersion algorithm ($QE^{EOG\text{-}dispersion}$), EOG-based QE with velocity algorithm ($QE^{EOG\text{-}velocity}$), absolute time-frequency EOG power (AP^{EOG}), and relative time-frequency EOG power (RP^{EOG})—for their estimated ability to predict golf putting accuracy on out-of-sample data. Third, we selected the best performing technique and examined its association with performance accuracy and movement smoothness. For all analyses, trial-level data were median-averaged to yield one value per participant for each processing parameter.

Power analyses conducted through G*Power 3.1.9.4 (Faul et al., 2007) indicated that with a sample size of 16, statistical significance threshold of .05, and statistical power of .80, our two-tail analyses were sensitive to detect large effects corresponding with correlation coefficients more extreme than ±.50. The participant-level dataset used for these analyses is available in the repository described in the Data and Code Availability section.

2.4.1 | Validity and accuracy of QE^{EOG}

Concurrent validity

These analyses aimed to describe the changes in the Pearson correlation coefficient computed between QE^{EOG} and QE^{ET} as a function of phase (pre, post, total) and EOG processing options (channel, filter, threshold). We evaluated the statistical significance of each coefficient through nonparametric permutation testing corrected for multiple comparisons using the extreme-value method (for $p < .05$ with 5000 iterations; Cohen, 2017). Coefficients were computed as $\mathrm{P}\left(\mathrm{Mdn}_{t=1}^{m}\left(QE_{p,t}^{ET}\right) : \mathrm{Mdn}_{t=1}^{m}\left(QE_{p,t}^{EOG}\right)\right)_p$, where $QE_{p,t}^{ET}$ and $QE_{p,t}^{EOG}$ indicate the QE durations for participant p and trial t obtained, respectively, through ET and EOG, $\mathrm{Mdn}_{t=1}^{m}(.)$ indicates the median across all m trials, and $\mathrm{P}(x:y)_p$ indicates the Pearson correlation coefficient computed between two vector variables x and y across all participants. Correlation coefficients larger than .10, .30, and .50 were used as references to describe respectively 'small', 'medium', and 'large' effects (see reference values in Cohen, 1992). In addition, we considered coefficients greater than .80 as corresponding to 'very large' effect.

Accuracy

These analyses aimed to evaluate the accuracy of QE^{EOG} scores by computing the temporal discrepancy (ms) with QE^{ET} scores, as a function of phase (pre, post, total) and EOG processing options (channel, filter, threshold). We assessed temporal discrepancy through two scores: median absolute deviation (MdnAD) and median deviation (MdnD). The former evaluates the expected unsigned temporal discrepancy over multiple trials, whereas the latter evaluates any directional bias due to the EOG algorithm consistently under or overestimating the QE durations compared to QE^{ET}. MdnAD and MdnD scores were computed, respectively, as $Mdn_{p=1}^{n}\left(\left|Mdn_{t=1}^{m}\left(QE_{p,t}^{ET}\right) - Mdn_{t=1}^{m}\left(QE_{p,t}^{EOG}\right)\right|\right)$ and $Mdn_{p=1}^{n}\left(Mdn_{t=1}^{m}\left(QE_{p,t}^{ET}\right) - Mdn_{t=1}^{m}\left(QE_{p,t}^{EOG}\right)\right)$, where $QE_{p,t}^{ET}$ and $QE_{p,t}^{EOG}$ indicate the QE measurements for participant p and trial t obtained, respectively, through ET and EOG, $Mdn_{t=1}^{m}(.)$ indicates the median across all m trials, and $Mdn_{p=1}^{n}(.)$ indicates the median across all n participants, and $|.|$ indicates the absolute value.

2.4.2 | Technique selection

This set of analyses aimed to rank the utility of each of five oculomotor techniques when predicting performance accuracy. In line with our exploratory aims and to maximize the reliability of the results, we adopted a machine-learning approach to within-study replication, broadly referred to as cross-validation, whereby a predictive model is fit (i.e., trained) on a subset of the data and then fit (i.e., evaluated) on a subset of the data that was held out from model fitting (Muller et al., 1983). Below, we describe in turn the predictive model and cross-validation procedure.

Our predictive model was a Theil–Sen univariate linear regression (Sen, 1968; Theil, 1950). We chose it to ensure robustness of fit and linear interpretability of its main model parameter, the regression slope. In this regression model, the output variable (i.e., target) was

radial error, whereas the only input variable (i.e., feature) was in turn each of the ET or EOG measurement corresponding with the following hyperparameters[4]: 3 stages (pre, post, total) for QE^{ET}; 3 channels (horizontal, vertical, combined) × 8 median filter settings × 10 thresholds × 3 stages (pre, post, total) for both $QE^{EOG\text{-}dispersion}$ and for $QE^{EOG\text{-}velocity}$; and 4 channels (horizontal, vertical, combined, mastoid) × 437 time-frequency areas of interest (AOI) for AP^{EOG} and RP^{EOG}. Time-frequency AOIs were computed to reduce the data dimensionality and, therefore, to decrease the computational resources needed to perform the analyses. Each AOI was defined as a 0.5-s × 10-Hz rectangle, covering the (−4, 2] s time interval and the (0, 100] Hz frequency interval, with 50% overlap between contiguous AOIs. More in detail, the rectangular AOIs were built so that, for any frequency, there were 23 8-point time intervals (i.e., (−4, −3.5], (−3.75, −3.25], ..., (1.5, 2] s), and for any time, there were 19 5-point frequency intervals (i.e., (0, 10], (5, 15), ..., (90, 100] Hz). EOG power values within each AOI were median averaged to yield a single value.

We used nested cross-validation (Varma & Simon, 2006) to estimate the out-of-sample generalization error yielded by each of the five oculomotor techniques when predicting performance accuracy. This procedure uses two cross-validation loops nested within one another. In the outer loop, the full dataset was iteratively split into an outer-training set and a test set, and in the inner loop, the outer-training set was further split iteratively into an inner-training set and a validation set. The iterative splits were obtained through the leave-one-out procedure, hence yielding 15 inner-loop iterations for each of the 16 outer-loop iterations.

Within each outer loop, the predictive model was fit on the inner-training set using each hyperparameter individually as input variable. Then, separately per each of the five techniques, we identified the best hyperparameter as that yielding the lowest median squared logarithmic error across the validation sets ($MdnSLE_{validation}$), computed as $\mathrm{Mdn}_{i=1}^{n}\left(\left[\log\left(y_i^{obs}+1\right)-\log\left(y_i^{pre}+1\right)\right]^2\right)_{inner}$, where y_i^{obs} indicates the radial error observed in the validation set i, y_i^{pre} indicates the radial error in the validation set i as predicted through the inner-training set, and $\mathrm{Mdn}_{i=1}^{n}(.)_{inner}$ indicates the median across the n (15) inner loops. In lay terms, the purpose of inner cross-validation was to identify the best validated hyperparameters (a procedure also known as hyperparameter tuning) per each oculomotor technique and feed them to the outer cross-validation loop which performed the so-called *technique selection*, described next.

For each outer loop, the predictive model was fit on the outer-training set using each technique's best hyperparameter as predictor and then evaluated on the test set by computing the logarithmic errors. This procedure yielded a pool of 16 test-set logarithmic errors (one per each outer loop) which were used to estimate (a) the generalization error of each technique as the median squared logarithmic error across the test sets ($MdnSLE_{test}$) and (b) its stability through the interquartile range of the logarithmic errors ($IQRLE_{test}$). $MdnSLE_{test}$ and $IQRLE_{test}$ were computed, respectively, as $\mathrm{Mdn}_{j=1}^{m}\left(\left[\log\left(y_j^{obs}+1\right)-\log\left(y_j^{pre}+1\right)\right]^2\right)_{out}$ and $\mathrm{IQR}_{j=1}^{m}\left(\log\left(y_j^{obs}+1\right)-\log\left(y_j^{pre}+1\right)\right)_{out}$, where y_j^{obs} indicates the radial error observed in the test set j, y_j^{pre} indicates the radial error in the test set j as predicted from the outer-training set, and $\mathrm{Mdn}_{j=1}^{m}(.)_{out}$ and $\mathrm{IQR}_{j=1}^{m}(.)_{out}$ indicate, respectively, the median and the interquartile range across the 16 m outer loops.

The three next and final analytic steps focused on the best-performing oculomotor technique selected through nested cross-validation. Because the details of these analyses are technique-dependent, it is worth anticipating that the technique yielding the smallest generalization error and greater stability was relative time-frequency EOG power (RP^{EOG}).

2.4.3 | Hyperparameter tuning

This set of analyses was subsequent to the technique selection described above and aimed to identify the RP^{EOG} best-validated hyperparameters for predicting subsequent radial error (performance accuracy). To do so, we used the same Theil–Sen regression model described for nested cross-validation. However, differently from the hyperparameter tuning occurring for each outer loop of nested cross-validation, this procedure used all available data for a single cross-validation. The full dataset was iteratively split into training and validation sets through the leave-one-out procedure, which yielded 16 loops. The validation error was computed for each hyperparameter (4 channels × 437 AOIs) as the median squared logarithmic error ($MdnSLE_{validation}$): $\mathrm{Mdn}_{k=1}^{v}\left(\left[\log\left(y_k^{obs}+1\right)-\log y_k^{pre}+1\right]^2\right)$, where y_k^{obs} indicates the radial error observed in the validation set k, y_k^{pre} indicates the radial error in the validation set l as predicted through the training set, and $\mathrm{Mdn}_{k=1}^{v}(.)$ indicates the median across the v (16)

[4]The term hyperparameter identifies each data processing choice in an algorithm (e.g., channel, filter order). The prefix "hyper" is used in the machine-learning literature to distinguish this type of processing parameter from the parameters estimated by a predictive model (e.g., intercept and slope for a regression model).

cross-validation loops. The hyperparameters with relatively small $MdnSLE_{validation}$ (i.e., below the 5th percentile) were highlighted and the associated full-dataset Theil–Sen regression slopes were reported.

2.4.4 | Oculomotor antecedents of performance

This set of analyses aimed to examine the changes in correlation coefficients describing the linear (Pearson's r) and monotonic (Spearman's ρ) association between radial error (performance accuracy) and RP^{EOG}, as a function of each 97 time × 51 frequency points in each of the four channels (horizontal, vertical, combined, mastoid). Cluster-based multiple-comparison correction was applied separately for each of the two sets of correlations (5000 iterations; Cohen, 2017). To describe the effect size of significant results, coefficients within significant clusters were Fisher z transformed,[5] and the back-transformed median and *IQR* were reported. The results of the same analyses conducted on the other four oculomotor techniques are reported in the online supplemental material.

2.4.5 | Kinematic concomitants of oculomotor activity

This set of analyses aimed to explore the functional meaning of oculomotor activity by describing the changes in the linear (Pearson's r) and monotonic (Spearman's ρ) association between the two indices of movement smoothness in the two movement stages (i.e., backswing LDLJ, downswing LDLJ, backswing SPARC, downswing SPARC) and RP^{EOG} as a function of each 97 time × 51 frequency points in each of the four channels (horizontal, vertical, combined, mastoid). To identify significant time-frequency clusters, we used nonparametric permutation testing (i.e., participants' values were permuted) corrected for multiple comparisons across the four smoothness indices and the four channels using the cluster-size method ($p < .05$, 5000 iterations; Cohen, 2017). To describe the effect size of significant results, coefficients within significant clusters were Fisher z transformed, and the back-transformed median and IQR were reported to describe the effect size. The results of the same analyses conducted on the other four oculomotor techniques are reported in the online supplemental material.

[5]The Fisher z transformation (i.e., inverse hyperbolic tangent) is applied to correlation coefficients prior to further computations (Corey et al., 1998).

3 | RESULTS

3.1 | QE^{EOG} validity and accuracy

The first set of analyses aimed to test the concurrent validity between QE^{ET} and QE^{EOG} through Pearson correlations. The second set of analyses aimed to evaluate the accuracy of QE^{EOG} measurements against QE^{ET} measurements through two discrepancy scores. Across a wide range of processing options, correlation coefficients were 'very large' (Figure 5a,b) and discrepancy scores were small (Figure 5c–f). To provide reference values for optimal parameters, the dispersion algorithm applied to 256-point median filtered horizontal channel with a 3° threshold yielded correlation coefficients of .97, .74, and .99, MdnAD scores were 161, 184, and 88 ms, and MdnD scores were 70, 32, and 16 ms for QE_{pre}, QE_{post}, and QE_{total}, respectively. The velocity algorithm applied to the horizontal channel filtered through a 5th-degree 767-point Savitzky–Golay differentiation filter and using a 33°/s threshold yielded correlation coefficients of .99, .90, and .99, MdnAD scores of 127, 65, and 133 ms, and MdnD scores of 127, −47, and 99 ms for QE_{pre}, QE_{post}, and QE_{total}, respectively. Additional analyses were run to evaluate the intra-individual concurrent validation and its consistency across participants on the same combination of processing parameters described above. We computed the Pearson's correlation coefficient across trials separately for each participant, then submitted each pool of coefficients to a one-sample Wilcoxon exact signed-rank test against the zero-median null hypothesis and reported the back-transformed median of the Fisher z transformed correlation coefficients. When using the dispersion algorithm and the parameters combination of horizontal channel, 256-point median filter, and a 3° threshold, the coefficients were consistently greater than 0 (QE_{pre}: $Mdn = .70$, $W = 134$, $p < .001$; QE_{post}: $Mdn = .24$, $W = 124$, $p = .002$; QE_{total}: $Mdn = .46$, $W = 136$, $p < .001$). Similarly, when using the velocity algorithm and the parameters combination of horizontal channel, 5th-degree 767-point Savitzky–Golay differentiation filter, and a 33°/s threshold, the coefficients were consistently greater than 0 (QE_{pre}: $Mdn = .73$, $W = 130$, $p < .001$; QE_{post}: $Mdn = .39$, $W = 136$, $p < .001$; QE_{total}: $Mdn = .64$, $W = 132$, $p < .001$). The QE durations obtained from EOG with these settings are reported alongside those obtained from ET in Table 1.

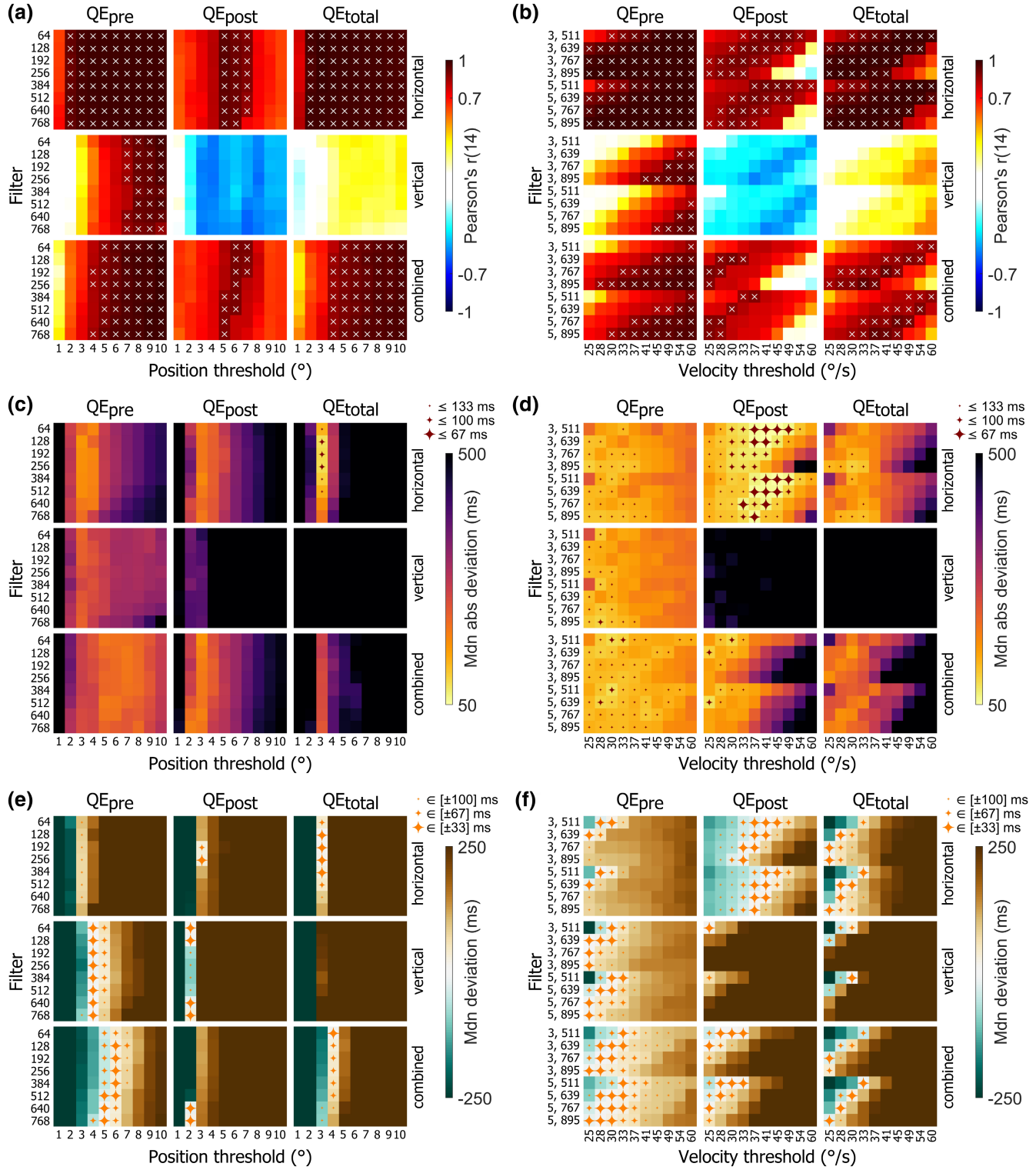

FIGURE 5 Validity and accuracy of QE^{EOG} measurements. QE^{EOG}:QE^{ET} Pearson's correlation coefficients are reported in panels a and b. The × symbol indicates a statistically significant result ($p < .05$) obtained through permutation testing with the extreme-value correction for multiple comparison (5000 iterations; Cohen, 2017) conducted separately for each of the two algorithms. Median absolute deviation scores (MdnAD) are reported in panels c and d. Dark-red stars highlight scores lower or equal than 133 ms (small star), 100 ms (medium star), and 67 ms (large star). Median deviation scores (MdnD) are reported in panels e and f. Orange stars highlight scores between −100 and 100 ms (small star), between −67 and 67 ms (medium star), and between −33 and 33 ms (large star). The left panels (a, c, and e) pertain to $QE^{EOG\text{-}dispersion}$ and the right panels (b, d, and f) to $QE^{EOG\text{-}velocity}$. Within each panel, scores are reported as a function of channel (horizontal, vertical, combined), filter parameters (filter order for the dispersion algorithm and polynomial degree and filter order for the velocity algorithm), threshold (10 values depending on the algorithm), and QE phase (pre, post, total).

3.2 | Technique selection

Nested cross-validation indicated that the RP^{EOG} technique yielded the smallest generalization error (92.77% smaller than QE^{ET}) followed by AP^{EOG} (67.76% smaller than QE^{ET}), $QE^{EOG\text{-}velocity}$ (19.86% smaller than QE^{ET}), QE^{ET}, and finally $QE^{EOG\text{-}dispersion}$ with the largest generalization error (80.84% larger than QE^{ET}). RP^{EOG} also had the greatest stability, followed by AP^{EOG}, $QE^{EOG\text{-}velocity}$, QE^{ET}, and finally $QE^{EOG\text{-}dispersion}$ (Table 2).

A descriptive analysis of the nested cross-validation hyperparameter tuning of the RP^{EOG} technique (Table 3) indicated that the horizontal channel, a time window between 0.75 and 1.25 s, and a frequency window between 75 and 100 Hz were consistently selected for the outer loops. The outer-training sets' slopes associated with these rounds of hyperparameter tuning were negative, suggesting an association between lower radial error (i.e., better putting accuracy) and larger relative EOG power within the channel-time-frequency space mentioned above.

3.3 | Hyperparameter tuning

The follow-up cross-validated Theil–Sen regression analysis for tuning the RP^{EOG} hyperparameters (Figure 6a) indicated that the lowest predictive errors were mostly localized to the horizontal channel, in the AOI overlapping movement execution (within the [−0.75, 1.25] s interval), and with a frequency content larger than 40 Hz. For these hyperparameters, the full-dataset Theil–Sen regression slope was consistently negative ($Mdn = -0.08$,

TABLE 1 Median (and IQR) of the QE durations (ms) across participants, after that trial-level values were median-averaged within each participant.

Technique	Pre	Post	Total
QE^{ET}	1569.75 (731.26)	905.73 (409.75)	2582.25 (1532.15)
$QE^{EOG\text{-}dispersion}$	1562.38 (942.57)	1073.49 (618.84)	2523.32 (1336.61)
$QE^{EOG\text{-}velocity}$	1552.98 (746.64)	884.28 (488.59)	2417.48 (1521.00)

Note: Each cell represents $\mathrm{Mdn}^{n}_{p=1}\left(\mathrm{Mdn}^{m}_{t=1}\left(QE^{q}_{p,t,s}\right)\right)$, where $QE^{q}_{p,t,s}$ indicates the QE duration for participant p in trial t obtained through the technique q (ET, EOG-dispersion, EOG-velocity) for the movement stage s (pre, post, total), and $\mathrm{Mdn}^{n}_{p=1}(.)$ and $\mathrm{Mdn}^{m}_{t=1}(.)$ indicate, respectively, the median across all n participants and the median across all m trials. $QE^{EOG\text{-}dispersion}$ values are reported for a 3° threshold with a 256-point median filter applied to the horizontal channel. $QE^{EOG\text{-}velocity}$ values are reported for a 33°/s threshold with a 5th-degree 767-point Savitzky–Golay differentiation filter applied to the horizontal channel.

TABLE 2 Nested cross-validation results for the five oculomotor techniques (QE^{ET}, $QE^{EOG\text{-}velocity}$, $QE^{EOG\text{-}dispersion}$, AP^{EOG}, RP^{EOG}).

Technique	$MdnSLE_{test}\cdot10^{4}$	$IQRLE_{test}\cdot10^{2}$	$RangeLE_{test}\cdot10^{2}$
$QE^{EOG\text{-}dispersion}$	124.02	22.74	[−15.36, 27.88]
QE^{ET}	68.58	16.37	[−14.88, 20.22]
$QE^{EOG\text{-}velocity}$	54.96	15.36	[−32.66, 26.27]
AP^{EOG}	22.11	11.61	[−8.48, 25.57]
RP^{EOG}	4.96	5.49	[−7.68, 31.16]

Note: $MdnSLE_{test}$ is reported as a measure of generalization error. $IQRLE_{test}$ is reported as inverse index of technique stability. $RangeLE_{test}$ indicates the range of the test-set logarithmic errors. Rows are sorted for descending $MdnSLE_{test}$. Values are reported as multiplied by 10^4 and 10^2.

TABLE 3 Hyperparameter tuning for the RP^{EOG} technique for the 16 outer cross-validation loops.

Hyperparameter combination	Count	Mdn slope	IQR slope
Channel: horizontal; AOI: (0.75, 1.25] s × (80, 90] Hz	7	−0.11	0.0018
Channel: horizontal; AOI: (0.75, 1.25] s × (90, 100] Hz	4	−0.13	0.0081
Channel: horizontal; AOI: (0.75, 1.25] s × (75, 85) Hz	4	−0.11	0.0033
Channel: horizontal; AOI: (0.75, 1.25] s × (85, 95) Hz	1	−0.12	–

Note: The dash replaces a non-meaningful IQR (computed across one value). Each row identifies a combination of hyperparameters (channel and AOI). Each combination is reported along with how many times it was selected across the outer loops (i.e., count out of 16), the median of the slopes obtained from the outer-training sets (i.e., Mdn slope), and their interquartile range (i.e., IQR slope).

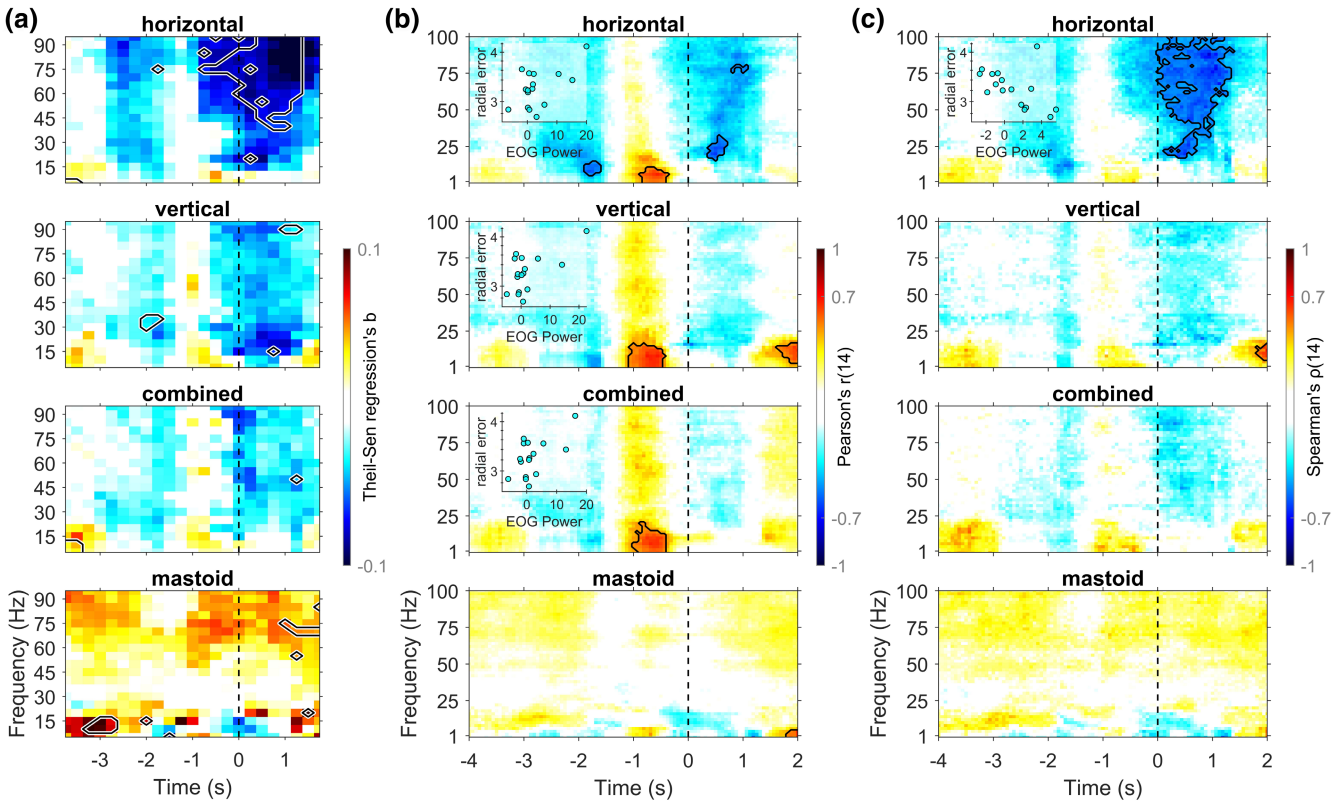

FIGURE 6 Panel a: Full-dataset Theil–Sen regression slopes describing RP^{EOG} hyperparameters (4 channels × 437 AOIs) as predictors of performance accuracy (i.e., radial error). The contour lines highlight the slopes corresponding with cross-validation errors below the 5th percentile. Panels b and c: Pearson and Spearman correlation coefficients between performance accuracy (i.e., radial error) and RP^{EOG} as a function of 4 channels, 97 times, and 51 frequencies. The contour lines highlight statistically significant results ($p < .05$) obtained through cluster-corrected permutation testing. For each time-frequency plot, the largest cluster, among those larger than 20 time-frequency points, was further explored visually through scatterplots.

IQR = 0.02, *range* = [−0.12, −0.06]), indicating that participants showing larger increases within this AOI putted more accurately.

3.4 | Oculomotor antecedents of performance accuracy

Pearson's correlation tests revealed a positive linear association between RP^{EOG} and radial error, indicating that worse performance accuracy was preceded by larger RP^{EOG} within a time-frequency cluster identified by the final second before movement initiation (namely, within the intervals [−0.875, −0.375] s, [−1.0625, −0.4375] s, and [−1, −0.4375] s respectively for the horizontal, vertical, and combined channels) and by a frequency content lower than ca. 20 Hz (namely, within the intervals [0, 10] Hz, [0, 16] Hz, and [0, 20] Hz respectively for the horizontal, vertical, and combined channels) across all three EOG channels but not in the control mastoid channel. For the cluster in the horizontal channel, Pearson's *r*s had a median of .55 (*IQR* = 0.08, *range* = [.50, .67]) corresponding to a large effect. For the remaining clusters, the reader is directed to Figure 6b.

Spearman's correlation tests highlighted a negative monotonic association between RP^{EOG} and radial error, indicating that better performance was preceded by larger RP^{EOG} within a time-frequency cluster overlapping movement execution ([0, 1.375] s) and with frequency content larger than 18 Hz limited to the horizontal channel only. Within this cluster, Spearman's ρs had a median of −.54 (*IQR* = 0.08, *range* = [−.77, −.50]) corresponding to a large effect (Figure 6c).

3.5 | Kinematic concomitants of oculomotor activity

Pearson's and Spearman's correlation tests revealed a positive linear and monotonic association between greater RP^{EOG} and movement smoothness immediately before and during movement execution (Figure 7). For both LDLJ and SPARC, this association was more evident during the downswing and for the EOG channels.

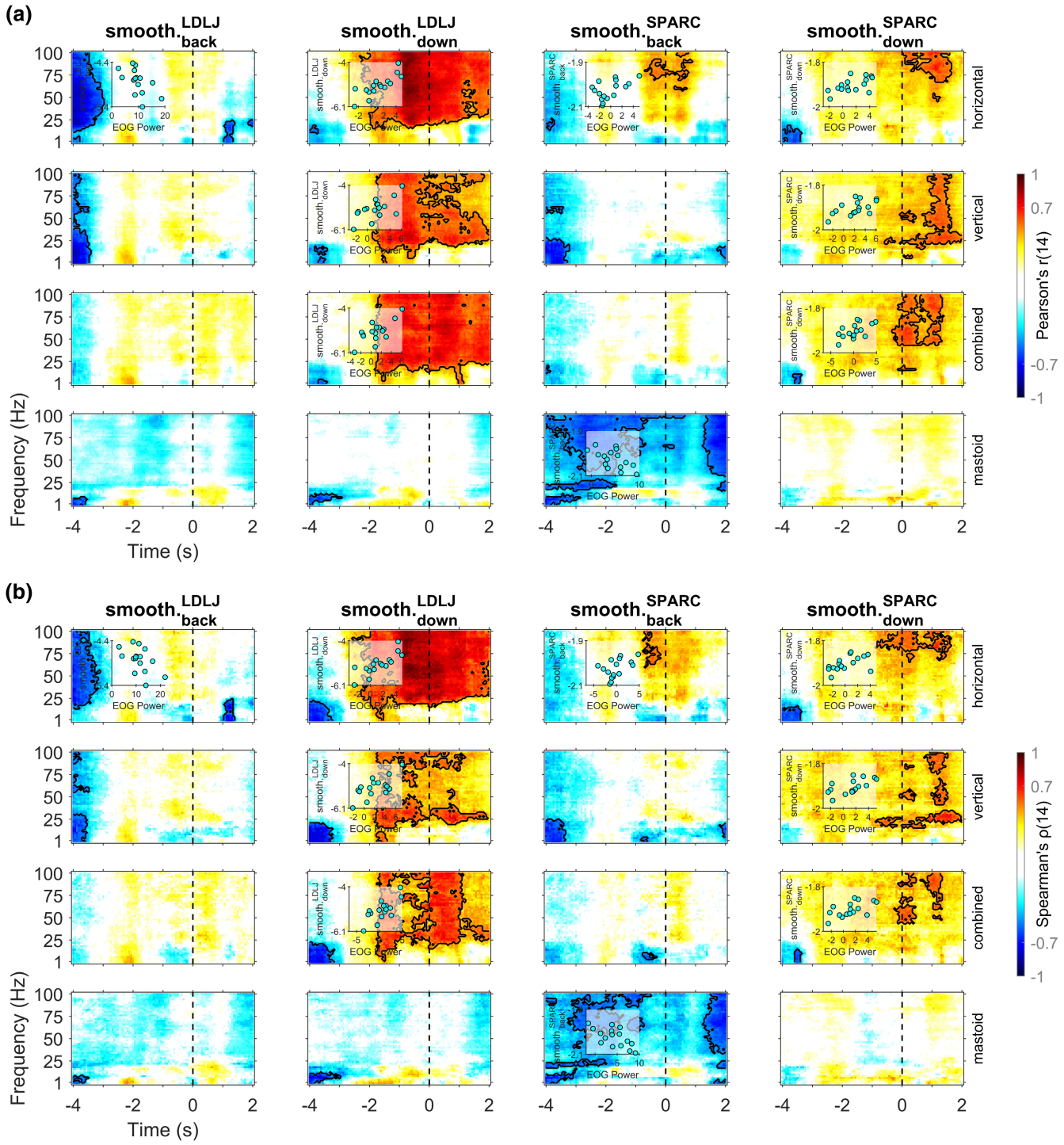

FIGURE 7 Correlation coefficients (panel a: Pearson's *r*; panel b: Spearman's ρ) between smoothness metrics (LDLJ for backswing, LDLJ for downswing, SPARC for backswing, SPARC for downswing) and RP^{EOG} as a function of 4 channels, 97 times, and 51 frequencies. The contour lines highlight statistically significant results ($p < .05$) obtained through cluster-corrected permutation testing. For each time-frequency plot, the largest cluster (among those larger than 50 time-frequency points and ending after −1 s) was further explored visually through scatterplots.

For downswing LDLJ and the horizontal channel, the largest Pearson's *r* cluster was identified in the [−2.3125, 2] s time interval and [16, 100] Hz frequency interval with a median of .68 (*IQR* = 0.31, *range* = [.50, .92]). The largest Spearman's ρ cluster was identified in the [−2.5625, 2] s time interval and [0, 100] Hz frequency interval with a median of .72 (*IQR* = 0.34, *range* = [.50, .94]). For downswing SPARC and the horizontal channel, the largest Pearson's *r* cluster was identified in the [0.0625, 1.625] s time interval and [68, 100] Hz frequency interval with a median of .55

(*IQR* = 0.09, *range* = [.50, .71]). The largest Spearman's ρ cluster was identified in the [−0.875, 1.625] s time interval and [68, 100] Hz frequency interval with a median of .54 (*IQR* = 0.06, *range* = [.50, .69]). These coefficients corresponded to a large to very large effect. For all clusters, the reader is directed to Figure 7.

4 | DISCUSSION

This study had three goals. First, to evaluate validity and accuracy of EOG-based QE measurements. Second, to compare the utility of various temporal and spectral indices of oculomotor activity when predicting behavioral performance. Third, to explore the cross-modal link between eye movements and movement quality. We found that the EOG can provide valid and accurate QE measurements for certain processing parameters and algorithms at both group and individual level. However, EOG-based metrics accounting for both temporal and spectral content proved more useful when predicting movement outcome and moment quality.

4.1 | QE^{EOG} validity and accuracy

The EOG signals showed a relative flatness around movement initiation, which corresponded with the QE period (see Figure 3 for a sample trial and the supplemental material for all trials). This correspondence provides face validity to EOG-based QE measurements. To formally identify and quantify such flatness, we tested two algorithms, each with various processing parameters. Then, we compared these measurements with those obtained through ET to evaluate their validity and accuracy. These analyses, reported in Figure 5, revealed the following general trends.

First, the most valid and accurate measurements emerged when using the horizontal channel alone. The lesser utility of the vertical channel, alone or combined with the horizontal channel, may be attributed to biological factors, as for example the eyelid movements confound the vertical channel more than the horizontal (Yee et al., 1985), or to task-specific factors, as all key visual elements in golf putting (i.e., the ball and the target) lie along the mediolateral axis and therefore require horizontal saccades.

Second, when applied to the horizontal channel, the dispersion algorithm performed best with a threshold of 3°. Because ET-based QE durations were also scored using a 3° threshold, this finding provides a quality check validating the specificity of the dispersion algorithm. In combination with the 3° threshold, a median filter order corresponding to a window between 62.5 and 125 ms provided more valid and accurate measurements. Although larger filter orders could simplify the waveforms more by enhancing the stepwise morphology of the EOG waveforms, it appeared that some short-time features (e.g., ocular overshoots) may be useful when detecting the onset of a fixation following a large saccade and therefore should not be filtered out (we point the reader to the difference between the 125-ms and the 1000-ms median filters in Figure 3 around the QE^{ET} onset).

Third, overall, the velocity threshold algorithm yielded more valid and accurate measurements than the dispersion algorithm. The superiority of the velocity algorithm was especially evident for QE_{post}. In comparison with QE_{post}, the relative worse performance of the velocity algorithm for QE_{pre} may be due to the greater tendency for ocular overshoots and blinks to occur immediately before the QE onset than immediately after the QE offset. Any of the differential filter parameters explored in this study appeared to be valid when applied in combination with a certain velocity threshold. While the optimal velocity threshold in saccade detection algorithms is greatly influenced by data types and signal processing choices, it is noteworthy that commonly used algorithms for the automatic saccade detection from ET data also apply velocity thresholds similar to the ones that resulted optimal in our study (e.g., 30°/s; Nyström & Holmqvist, 2010).

4.2 | Temporal and spectral features

A data-driven selection conducted among several temporal and spectral oculomotor features grouped under five broad categories, QE^{ET}, $QE^{EOG\text{-}dispersion}$, $QE^{EOG\text{-}velocity}$, AP^{EOG}, and RP^{EOG}, indicated that the latter (i.e., RP^{EOG}) was the technique yielding the lowest and most stable estimation of the out-of-sample generalization error when predicting subsequent performance accuracy (Table 2). The hyperparameter tuning embedded within the nested cross-validation (Table 3), as well as that conducted in the follow-up non-nested cross-validation (Figure 6a), indicated that the RP^{EOG} parameters most reliably predicting performance accuracy were those defining high-frequency activity recorded from the horizontal channel at a time overlapping movement execution. The evaluation of the regression slopes indicated that individuals with greater activity within this time-frequency AOI performed more accurately. These findings partly converged with those resulting from the linear and monotonic correlation analyses (Figure 6b,c). Overall, these analyses revealed two main clusters.

The first cluster consisted of lower-frequency activity occurring immediately before movement initiation and emerged mainly through linear correlations. The inspection of Pearson's *r* coefficients indicated that greater activity in this cluster had a large detrimental

effect on subsequent performance accuracy. This cluster occurred before movement initiation and appeared, although only as a trend, in the control mastoid channel, suggesting that this effect may not entirely be attributed to eye movements. Accordingly, this effect may indicate a suboptimal motor preparation consisting of not only lower-frequency oculomotor activity (e.g., saccades) but also head adjustments, reflecting a postural re-orienting that happened too close in time to the initiation of the movement. This finding is aligned with previous research indicating that psychomotor quietening before movement initiation is a feature of expertise in target sports (e.g., Cooke et al., 2014).

The second cluster consisted of higher-frequency activity occurring during movement execution and emerged through cross-validated regressions as well as monotonic correlations. The inspection of regression slopes and Spearman's ρ coefficients indicated that greater activity within this cluster had a large beneficial effect for subsequent performance accuracy. Its occurrence during the QE period (Table 1), its specificity to the horizontal EOG channel while being absent in the control mastoid channel, and its high-frequency content suggest that this cluster may reflect fixational activity. Indeed, oculomotor activity within a frequency range overlapping this cluster has been previously observed when individuals fixate a visual target (30–150 Hz; Young & Sheena, 1975) and can be linked to fixational movements such as ocular tremor that prevent sensorial decay at the retinal level (Krauzlis et al., 2017). The golf putting swing is inevitably accompanied by head and trunk rotation and, therefore, this high-frequency cluster may reflect compensatory vestibulo-ocular adjustments promoting the persistence of a clear retinal image. It is noteworthy that the high-frequency cluster correlated positively with various indices of movement smoothness (Figure 7), potentially suggesting that a stable visual percept, supported by high-frequency fixational eye movements, facilitates a less jerky, less controlled, and smoother movement execution.

Interestingly, the duration of the QE was not associated with individual performance (see supplemental material). While this is in contrast with most of the QE literature to date, there is increasing experimental evidence that the location and duration of the gaze are not as important for performance as initially thought (Harris et al., 2021; Lee, 2015).

4.3 | Limitations and future directions

Despite probing a vast parameter space, a residual QE^{ET}-QE^{EOG} temporal discrepancy remained. This residual discrepancy may be intrinsic to a series of factors including, for example, data type (video frames for ET, voltages for the EOG), scoring procedure (subjective evaluation for ET, computational analysis for the EOG), gaze location data (visually evaluated for ET, assumed from instructions for EOG), or differences in temporal resolution (greater for the EOG). Future research interested in decreasing this residual discrepancy could attempt to address the major shortcoming of the EOG—its poor spatial resolution—by supplementing it with a head kinematics sensor.

Because the EOG signals depend on the electrode placement and the corneoretinal potential, which in turn is influenced by individual and ambient characteristics, we recommend researchers to perform a frequent gaze calibration whenever using the EOG to examine oculomotor activity, as done in this study. In absence of individual calibration data, researchers should avoid adopting a one-size-fits-all approach and evaluate how the results vary as a function of the processing parameters (cf. Gallicchio et al., 2018; Gallicchio & Ring, 2020).

The demonstrated strong associations of time-frequency EOG activity with behavioral performance and movement quality describe novel phenomena. Because the current EOG technology is not mature enough to accurately detect small eye movements and no other studies, that we are aware of, have examined the contribution of different ocular movement types to movement outcome or quality, our interpretations of low-frequency activity as gross orienting (e.g., saccades) and of high-frequency activity as fixational movements await confirmatory experimental testing. Despite the precise biological source of the high-frequency activity remains undefined, it remains that this activity regards an ocular, or at least periocular, phenomenon occurring at a time of relative ocular quiescence.

4.4 | Conclusion

Compared to the frame-by-frame visual inspection of ET data, EOG signal processing reduces the opportunity for human error, it increases the objectivity of data analysis, it obviates the need of multiple human scorers, it increases the reproducibility of the results, and it allows to tune processing parameters through repeated data scoring. While temporal features of the EOG signals can be used to approximate measurements obtained through ET under controlled experimental settings, using both temporal and spectral EOG features provides unique information that increases the utility of the EOG signals when predicting behavioral performance.

AUTHOR CONTRIBUTIONS

Germano Gallicchio: Conceptualization; data curation; formal analysis; investigation; methodology; project

administration; resources; software; validation; visualization; writing – original draft; writing – review and editing. **Donghyun Ryu:** Conceptualization; investigation; methodology; writing – review and editing. **Mudit Krishnani:** Data curation; investigation; writing – review and editing. **Guy L. Tasker:** Data curation; investigation; writing – review and editing. **Alessandra Pecunioso:** Data curation; investigation; writing – review and editing. **Robin C. Jackson:** Conceptualization; resources; supervision; writing – review and editing.

DATA AVAILABILITY STATEMENT

Sample data and MATLAB code accompanying this study are openly available on Zenodo (Gallicchio, 2023). The data that support the findings of this study are openly available on Figshare (Gallicchio et al., 2023).

ORCID

Germano Gallicchio https://orcid.org/0000-0002-2180-8848
Donghyun Ryu https://orcid.org/0000-0001-8054-4929
Mudit Krishnani https://orcid.org/0009-0003-5055-9662
Guy L. Tasker https://orcid.org/0009-0004-7019-7993
Alessandra Pecunioso https://orcid.org/0000-0001-8087-5234
Robin C. Jackson https://orcid.org/0000-0001-7983-3870

REFERENCES

Acuña, O. V., Aqueveque, P., & Pino, E. J. (2014). Eye-tracking capabilities of low-cost EOG system. In *36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (Vol. *2014*, pp. 610–613). IEEE. https://doi.org/10.1109/EMBC.2014.6943665

Balasubramanian, S., Melendez-Calderon, A., & Burdet, E. (2012). A robust and sensitive metric for quantifying movement smoothness. *IEEE Transactions on Biomedical Engineering*, *59*(8), 2126–2136. https://doi.org/10.1109/TBME.2011.2179545

Balasubramanian, S., Melendez-Calderon, A., Roby-Brami, A., & Burdet, E. (2015). On the analysis of movement smoothness. *Journal of Neuroengineering and Rehabilitation*, *12*(1), 1–11. https://doi.org/10.1186/10.1186/s12984-015-0090-9

Chia, S. J., Chow, J. Y., Kawabata, M., Dicks, M., & Lee, M. (2017). An exploratory analysis of variations in quiet eye duration within and between levels of expertise. *International Journal of Sport and Exercise Psychology*, *15*(3), 221–235. https://doi.org/10.1080/1612197X.2015.1114503

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Cohen, M. X. (2014). *Analyzing neural time series data: Theory and practice*. MIT Press. https://doi.org/10.7551/mitpress/9609.001.0001

Cohen, M. X. (2017). *MATLAB for brain and cognitive scientists*. MIT Press.

Cooke, A., Kavussanu, M., Gallicchio, G., Willoughby, A., McIntyre, D., & Ring, C. (2014). Preparation for action: Psychophysiological activity preceding a motor skill as a function of expertise, performance outcome, and psychological pressure. *Psychophysiology*, *51*(4), 374–384. https://doi.org/10.1111/psyp.12182

Corey, D. M., Dunlap, W. P., & Burke, M. J. (1998). Averaging correlations: Expected values and bias in combined Pearson rs and Fisher's z transformations. *The Journal of General Psychology*, *125*(3), 245–261. https://doi.org/10.1080/00221309809595548

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Gallicchio, G. (2023). GermanoGallicchio/Quiet-Eye-EOG: v1.0 (v1.0). *Zenodo*. https://doi.org/10.5281/zenodo.8411093

Gallicchio, G., Cooke, A., & Ring, C. (2018). Assessing ocular activity during performance of motor skills using electrooculography. *Psychophysiology*, *55*(7), 1–12. https://doi.org/10.1111/psyp.13070

Gallicchio, G., Finkenzeller, T., Sattlecker, G., Lindinger, S., & Hoedlmoser, K. (2016). Shooting under cardiovascular load: Electroencephalographic activity in preparation for biathlon shooting. *International Journal of Psychophysiology*, *109*, 92–99. https://doi.org/10.1016/j.ijpsycho.2016.09.004

Gallicchio, G., & Ring, C. (2020). The quiet eye effect: A test of the visual and postural-kinematic hypotheses. *Sport, Exercise & Performance Psychology*, *9*(1), 143–159. https://doi.org/10.1037/spy0000162

Gallicchio, G., Ryu, D., Krishnani, M., Tasker, G., Pecunioso, A., & Jackson, R. (2023). Dataset (Version 1). *Figshare*. https://doi.org/10.6084/m9.figshare.24250753.v1

Gonzalez, C. C., Causer, J., Grey, M. J., Humphreys, G. W., Miall, R. C., & Williams, A. M. (2017). Exploring the quiet eye in archery using field-and laboratory-based tasks. *Experimental Brain Research*, *235*(9), 2843–2855. https://doi.org/10.1007/s00221-017-4988-2

Hancock, G. R., Butler, M. S., & Fischman, M. G. (1995). On the problem of two-dimensional error scores: Measures and analyses of accuracy, bias, and consistency. *Journal of Motor Behavior*, *27*(3), 241–250. https://doi.org/10.1080/00222895.1995.9941714

Harris, D. J., Wilson, M. R., & Vine, S. J. (2021). A critical analysis of the functional parameters of the quiet eye using immersive virtual reality. *Journal of Experimental Psychology: Human Perception and Performance*, *47*(2), 308–321. https://doi.org/10.1037/xhp0000800

Henn, V., & Cohen, B. (1973). Quantitative analysis of activity in eye muscle motoneurons during saccadic eye movements and positions of fixation. *Journal of Neurophysiology*, *36*(1), 115–126. https://doi.org/10.1152/jn.1973.36.1.115

Hogan, N., & Sternad, D. (2009). Sensitivity of smoothness measures to movement duration, amplitude, and arrests. *Journal of Motor Behavior*, *41*(6), 529–534. https://doi.org/10.3200/35-09-004-RC

Jasper, H. H. (1958). The ten-twenty electrode system of the international federation. *Electroencephalography and Clinical Neurophysiology*, *17*, 37–46.

Juhola, M. (1991). Median filtering is appropriate to signals of saccadic eye movements. *Computers in Biology and Medicine*, *21*(1–2), 43–49. https://doi.org/10.1016/0010-4825(91)90034-7

Krauzlis, R. J., Goffart, L., & Hafed, Z. M. (2017). Neuronal control of fixation and fixational eye movements. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1718), 20160205. https://doi.org/10.1098/rstb.2016.0205

Lebeau, J. C., Liu, S., Saenz-Moncaleano, C., Sanduvete-Chaves, S., Chacon-Moscoso, S., Becker, B. J., & Tenenbaum, G. (2016). Quiet eye and performance in sport: A meta-analysis. *Journal of Sport and Exercise Psychology*, *38*(5), 441–457. https://doi.org/10.1123/jsep.2015-0123

Lee, D. H. (2015). *The role of the quiet eye in golf putting* (Doctoral dissertation, University of Exeter). Open Research Exeter. http://hdl.handle.net/10871/17707.

Mann, D. T., Coombes, S. A., Mousseau, M. B., & Janelle, C. M. (2011). Quiet eye and the Bereitschaftspotential: Visuomotor mechanisms of expert motor performance. *Cognitive Processing*, *12*(3), 223–234. https://doi.org/10.1007/s10339-011-0398-8

Marmor, M. F., Brigell, M. G., McCulloch, D. L., Westall, C. A., Bach, M., & International Society for Clinical Electrophysiology of Vision. (2011). ISCEV standard for clinical electrooculography (2010 update). *Documenta Ophthalmologica*, *122*(1), 1–7. https://doi.org/10.1007/s10633-011-9259-0

Marquardt, C. (2007). The sam puttlab: Concept and PGA tour data. *International Journal of Sports Science & Coaching*, *2*, 101–120. https://doi.org/10.1260/174795407789705479

Marquardt, C., & Mai, N. (1994). A computational procedure for movement analysis in handwriting. *Journal of Neuroscience Methods*, *52*(1), 39–45. https://doi.org/10.1016/0165-0270(94)90053-1

Muller, K. E., Otto, D. A., & Benignus, V. A. (1983). Design and analysis issues and strategies in psychophysiological research. *Psychophysiology*, *20*(2), 212–218. https://doi.org/10.1111/j.1469-8986.1983.tb03290.x

Neumann, D. L., & Thomas, P. R. (2008). A camera-based scoring system for evaluating performance accuracy during a golf putting task. *Behavior Research Methods*, *40*(3), 892–897. https://doi.org/10.3758/BRM.40.3.892

Nyström, M., & Holmqvist, K. (2010). An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. *Behavior Research Methods*, *42*(1), 188–204. https://doi.org/10.3758/BRM.42.1.188

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Prerau, M. J., Brown, R. E., Bianchi, M. T., Ellenbogen, J. M., & Purdon, P. L. (2017). Sleep neurophysiological dynamics through the lens of multitaper spectral analysis. *Physiology*, *32*(1), 60–92. https://doi.org/10.1152/physiol.00062.2015

Savitzky, A., & Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, *36*, 1627–1639. https://doi.org/10.1021/ac60214a047

Sen, P. K. (1968). Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association*, *63*(324), 1379–1389. https://doi.org/10.1080/01621459.1968.10480934

Shackel, B. (1960). Pilot study in electro-oculography. *The British Journal of Ophthalmology*, *44*(2), 89–113. https://doi.org/10.1136/bjo.44.2.89

Theil, H. (1950). A rank-invariant method of linear and polynomial regression analysis. *Indagationes Mathematicae*, *12*(85), 173.

Varma, S., & Simon, R. (2006). Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics*, *7*(1), 1–8. https://doi.org/10.1186/1471-2105-7-91

Vickers, J. N. (1992). Gaze control in putting. *Perception*, *21*(1), 117–132. https://doi.org/10.1068/p210117

Vickers, J. N. (1996). Visual control when aiming at a far target. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(2), 342–354. https://doi.org/10.1037/0096-1523.22.2.342

Vickers, J. N. (2007). *Perception, cognition, and decision training: The quiet eye in action*. Human Kinetics.

Xu, S., Sun, G., & Wilson, M. R. (2021). Neurophysiological evidence of how quiet eye supports motor performance. *Cognitive Processing*, *22*(4), 641–648. https://doi.org/10.1007/s10339-021-01036-3

Yee, R. D., Schiller, V. L., Lim, V., Baloh, F. G., Baloh, R. W., & Honrubia, V. (1985). Velocities of vertical saccades with different eye movement recording methods. *Investigative Ophthalmology & Visual Science*, *26*(7), 938–944.

Yoshikawa, N., Hirao, T., & Masaki, H. (2019). Quiet eye duration evaluated by electrooculogram during archery performance. *Psychophysiology*, *56*(S1), S43. https://doi.org/10.1111/psyp.13502

Young, L. R., & Sheena, D. (1975). Survey of eye movement recording methods. *Behavior Research Methods & Instrumentation*, *7*(5), 397–429. https://doi.org/10.3758/BF03201553

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

FIGURE S1. The calibration sheet was a white A4 paper sheet with four markers (black circles, diameter = 8 mm) positioned at the edge of the four sides. The distance between the left and right markers was 25 cm and the distance between the top and bottom markers was 17 cm. At the center of the sheet there was a black cross and three concentric circles with diameters measuring 1°, 3°, and 5° of visual angle when seen from a certain height. For each of the three angles, the linear lengths of the diameters (i.e., in cm) depended on the eye-floor distance.

FIGURE S2. EOG gaze calibration processing steps demonstrated for 1-s median-filtered horizontal (top) and vertical (bottom) EOG signals. For each channel, the light blue and dark gray waveforms show the EOG data respectively before and after linear detrending. The red portions of the detrended EOG identify the central 20% of the data between consecutive acoustic tones. The absolute differences between the medians of each central section are summarized by the boxplot on the right. The median

of these absolute differences was used as the voltage associated with a saccade between two fixation markers.

FIGURE S3. Scaling factor as a function of Median filter order and Channel. Error bars indicate the SE of the mean.

FIGURE S4. Scatterplots showing performance accuracy (log-transformed radial error, in cm) as a function of QE pre, post, and total durations (in ms). Each marker represents the median across trials for each participant. QE durations are scored using both the 'discard' approach, whereby trials with a QE_{total} shorter than 100 ms or with no fixation on the ball are discarded from analysis (blue circles), and the 'zero' approach, whereby those trials are given a duration of zero (orange xs). It is noteworthy that, for all participants except two, adopting either approach made a negligible difference; however, this is due to the robustness of median averaging values across trials: averaging across trials through the arithmetic mean would have produced larger differences between the 'discard' and 'zero' approaches. The blue-dotted and orange-solid lines indicate the least-square best fit respectively for the 'discard' and 'zero' approaches.

FIGURE S5. Pearson's (panels a and b) and Spearman's (panels c and d) correlation tests between performance accuracy (radial error) and QE pre, post, and total durations as a function of channel, filter settings, and thresholds. The results obtained with dispersion algorithm are reported in panels a and c, whereas those obtained with the velocity algorithm are in panels b and d. No significant result ($p < .05$) emerged through permutation testing with extreme-value multiple-comparison correction (5000 iterations; Cohen, 2017) conducted separately for each panel.

FIGURE S6. Correlation coefficients between putting performance (radial error) and RPEOG as a function of 4 channel, XX times, and YY frequencies are reported in panels b (Pearson's r) and c (Spearman's ρ). The contour lines highlight statistically significant results ($p < .05$) obtained through cluster-based multiple-comparison-corrected permutation testing. For each time-frequency plot, the largest cluster, among those larger than 20 time-frequency points, was further explored through scatterplots.

FIGURE S7. Pearson's (panels a–d and i–l) and Spearman's (panels e–h and m–p) correlation tests between smoothness (LDLJ during backswing, LDLJ during downswing, SPARC during backswing, SPARC during downswing) and QE durations (QE_{pre}, QE_{post}, and QE_{total}) obtained from EOG data scored through the dispersion algorithm (panels a–h) and the velocity algorithm (panels i–p), as a function of channel, filter settings, and thresholds. No significant result ($p < .05$) emerged through permutation testing with extreme-value multiple-comparison correction (5000 iterations; Cohen, 2017) conducted separately for each panel.

FIGURE S8. Correlation coefficients (panel a: Pearson's r; panel b: Spearman's ρ) between smoothness metrics (LDLJ for backswing, LDLJ for downswing, SPARC for backswing, SPARC for downswing) and AP^{EOG} as a function of 4 channels, 97 times, and 51 frequencies. The contour lines highlight statistically significant results ($p < .05$) obtained through cluster-corrected permutation testing. For each time-frequency plot, the largest cluster (among those larger than 50 time-frequency points and ending after −1 s) was further explored visually through scatterplots.

TABLE S1. Pearson's r and Spearman's ρ describing the linear and monotonic correlations between smoothness (LDLJ during backswing, LDLJ during downswing, SPARC during backswing, SPARC during downswing) and QE durations (QE_{pre}, QE_{post}, and QE_{total}) obtained from ET data, separately for the discard and zero method of scoring trials with no QE duration. Note: uncorrected p values.

How to cite this article: Gallicchio, G., Ryu, D., Krishnani, M., Tasker, G. L., Pecunioso, A., & Jackson, R. C. (2024). Temporal and spectral electrooculographic features in a discrete precision task. *Psychophysiology*, *61*, e14461. https://doi.org/10.1111/psyp.14461