IAES International Journal of Robotics and Automation (IJRA)
Vol. 15, No. 4, December 2026, pp. 874∼882
ISSN: 2722-2586, DOI: 10.11591/ijra.v15i4.pp874-882

Cue-locked gaze and time-limited decoding in motor imagery control of assistive robots

Fouad Fakhet, Rachida Djeradi
Speech Communication and Signal Processing Laboratory (LCPTS), Faculty of Electronics and Computer Science, Université des Sciences et de la Technologie Houari Boumediène (USTHB), Algiers, Algeria

Article Info

Article history:

Received Aug 21, 2026
Revised Oct 10, 2026
Accepted Nov 10, 2026

Keywords:

Assistive robotics
Brain-computer interface
Eye movements
Motor imagery
Shared control
Shortcut learning

ABSTRACT

Motor-imagery decoders for assistive robots are trained and validated while a directional cue is on screen; a robot in use shows no cue. We ask what such decoders learn and what survives cue removal, using two public datasets. Under a lateral cue visible throughout imagery, mean horizontal gaze position alone classifies left from right across 30 participants at an area under the curve of 0.87, against 0.68 for spatially filtered brain signals; the shift begins 413 ms after the cue. Where a central arrow is removed mid-imagery, the eye signal falls to chance. The brain signal does not follow the cue, but its cross-subject component is time-limited: a calibration-free decoder falls from 0.79 to 0.59, declining before the cue disappears, while a per-user calibrated decoder is unchanged. A pre-registered closed-loop simulation prices both: calibrated task success transfers to cue-free operation, calibration-free success falls from 70% to 45%, and a gaze-driven controller reaches the wrong target in 42% of cue-free episodes. Three mitigations are tested in the same loop: gaze as a zero-calibration channel when a target is visible, a confidence gate that cuts wrong-target commands from 10% to 4%, and a per-user reliability check.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Fouad Fakhet
Speech Communication and Signal Processing Laboratory (LCPTS), Faculty of Electronics and Computer Science, Université des Sciences et de la Technologie Houari Boumediène (USTHB)
BP 32 El Alia, Bab Ezzouar, Algiers 16111, Algeria
Email: ffakhet@usthb.dz

1. INTRODUCTION

Service robots for people with severe motor impairment, such as powered wheelchairs, reaching arms and exoskeletons, need a command channel that does not depend on limb movement. Motor imagery (MI), the mental rehearsal of a movement decoded from the scalp electroencephalogram (EEG), is one of the most widely used [1]: the user imagines moving the left or right hand, a classifier maps the EEG pattern to a command, and the robot acts [2]–[5].

MI decoders are built under a laboratory protocol that deployment does not reproduce. In every public dataset the trial opens with a directional visual cue that tells the participant which movement to imagine; accuracy is measured while that cue is, or has just been, on screen. A robot in use shows no cue. Whatever part of the learned signal belongs to the cue rather than to the imagery is absent at deployment, and the robot has no way to know. Two vectors carry that risk. The eyes: a lateral cue elicits an orienting saccade, after which eye position predicts the class label [6]. And time: the neural response to an instruction is strongest just after it arrives.

Brain-actuated wheelchairs [3], [4], [7] and robotic arms [5], [8] are all validated under cued protocols; Aljalal *et al.* [1] identify reliability outside the laboratory as the open problem. Gaze is a mature input for assistive arms [9], and hybrid interfaces of EEG and the electro-oculogram (EOG) restore daily-living function [10]. In recent shared-control designs gaze selects while MI confirms [11], [12], and passive eye movements are proposed as a state monitor [13]. All of these assume the two channels are independent. Calibration-free MI decoding is pursued with alignment methods [14], transfer and adaptive learning [15], [16], and reinforcement-signal-driven adaptation from error-related potentials [17]. None of these methods addresses what the training cue put into the signal. The decoded signal itself is event-related desynchronisation of sensorimotor rhythms, strongest shortly after onset [18], extracted with common spatial patterns [19]. In machine learning, a model that exploits a training-set correlate absent at deployment has learned a shortcut [20].

Whether cue-evoked eye movements *substitute* for the MI signal was asked for the P300 speller [21], which led to gaze-independent speller designs, and once for MI, where Meng *et al.* [22] found within-subject decoding unaffected by fixation. What a decoder given both channels actually relies on, what either signal does when the cue is withdrawn, and what the difference costs a robot in closed loop, have not been tested. Nor has the cue question been separated from the temporal decay that non-stationarity work documents [23], [18].

Using two public datasets whose cue designs differ in exactly the relevant way, we i) measure how much of cross-subject classification a single gaze feature carries and ii) time the gaze response; iii) test which signals survive cue removal for calibrated and calibration-free decoders; iv) replay real decoder outputs through a pre-registered closed-loop task; and v) test three mitigations in the same loop, reporting the one that failed. The contribution is a set of reproducible design rules that prevent gaze-dependent shortcuts and time-limited miscalibration, with the complete code and pre-registered specifications released.

Both failure modes are silent: the robot reaches the *wrong* target rather than stopping. For assistive use—feeding, reaching, steering a wheelchair—a wrong-target action is not a nuisance but a safety event, and cued offline validation cannot expose it because the cue is present in every fold.

2. METHOD

2.1. Datasets

Eye-BCI [24] (CC0, Synapse syn64005218): 31 participants, 63 sessions, 2520 left/right MI trials, 62-channel EEG at 1000 Hz and Tobii TX300 eye tracking at 300 Hz. A fixation cross is followed by a red rectangle *left or right* of it for 4 s, during which the participant imagines grasping with that hand. The cue is lateral and visible throughout; there is no cue-free window. Because the embedded eye-tracker timestamps are mis-paired in some sessions, EEG and gaze are aligned from the eye tracker's event markers (60 of 63 sessions within 12 ms; participant S06 unalignable and excluded); the procedure and per-session diagnostics are released with the code.

BCI Competition IV 2a [25]: 9 participants, 22 EEG and 3 EOG channels at 250 Hz. A *central arrow* appears at 2.0 s and is removed at 3.25 s; imagery continues to 6.0 s. This gives 2.75 s of imagery with no cue. Training sessions only, left/right trials, the dataset's artefact-flagged trials excluded. EOG1−EOG3 (right minus left canthus) is the horizontal derivation.

2.2. Decoders and evaluation

Eye features: eleven per trial over 0.5–4.0 s (gaze position mean and s.d. in x and y; fixation, saccade and blink counts and durations; pupil mean and s.d.). EEG: common spatial patterns (CSP, four filter pairs, 8–30 Hz, log-variance) [19] with linear discriminant analysis (LDA). Eye-BCI eye models use a random forest (200 trees); every 2a decoder and every simulation decoder uses LDA, so decoders differ only in their inputs. Imputation, scaling and CSP are fitted inside the training fold. *Calibration-free* means leave-one-subject-out (LOSO); *calibrated* means within-subject five-fold. Performance is the area under the ROC curve (AUC) or accuracy as stated. Participants whose mean horizontal gaze differs by less than 35 px between right- and left-cue trials are termed fixating.

2.3. Gaze latency and cue windows

Gaze latency: cue-locked horizontal gaze, baseline-corrected, averaged by class and differenced; latency is the first time the difference reaches 50% of its plateau and stays there for 100 ms, and participants with a pre-cue difference above 15% of plateau are excluded. 2a EOG: sign-preserving AUC of the window-mean horizontal EOG after removing the slope fitted to the pre-cue segment (method C). No correction (A) and

whole-epoch detrending (B) are also reported, because the result depends on the choice. 2a EEG: CSP+LDA in equal 1.25 s windows, 2.00–3.25 s (cue present) and 4.75–6.00 s (cue absent), calibrated and calibration-free. Calibration-free decoding is repeated on sub-windows, channel subsets, the EOG channels alone, EEG with the EOG regressed out, and sliding 0.5 s windows. Eye-BCI EEG is compared between 0.5–1.5 s and 3.0–4.0 s. The evaluation is a within-participant pretest–posttest design: the cue-present window is the pretest, the equal-length cue-free window the posttest, and every contrast is paired at the participant level (replicate-level validation). A randomized design is not available, because cue timing is fixed by each dataset; the one-group paired comparison, each participant their own control, is the strongest design the data admit, and its limitation—condition and elapsed time cannot be fully separated—is treated in the limitations. Tests are paired two-sided Wilcoxon across participants; effect sizes are reported as matched-pairs rank-biserial correlation r_{rb} and Cohen's d_z, with bootstrap 95% confidence intervals on the paired difference (10 000 resamples); every test run is reported, uncorrected. Non-parametric tests are used because nine participants give no basis for distributional assumptions. All alignment offsets, thresholds and analysis steps are released as numbered scripts with per-session diagnostics, so the synchronisation and gaze-reliability procedures can be replicated on new recordings.

2.4. Closed-loop simulation (pre-registered)

Metrics, windows and decision rules were committed before the run. Out-of-fold posteriors $p(\text{right})$ from three decoders—EEG, EOG, EEG+EOG—are replayed through a one-dimensional reach. The effector starts at 0 and the target is at ± 3. Each command draws a trial of the intended class and moves one step in the decoded direction; an episode ends at either end or after nine commands. Draws are independent with replacement; the task measures transfer of decoder behaviour between cued and cue-free operation, not absolute performance. 2000 episodes per cell, fixed seed. Metrics: task success (correct target), wrong-target rate (the robot completes an unintended action), commands to success, and overstatement (offline cued accuracy minus cue-free success). Declared variants: confidence gate $\tau = 0.2$ (no command when $|p - 0.5| < \tau$) and target distance 5. Added afterwards and labelled so: a 0.5 s cue-onset run, and a calibration-free arm trained on the other eight participants.

2.5. Mitigation study (pre-registered)

A second specification, committed before its run, evaluates five controllers in the same task, all with calibration-free decoders. They are gaze only (G), EEG only (E) and naive fusion of the two feature sets (F). A1 is an arbitration that routes to gaze when the robot has displayed a target and to EEG otherwise; A2 adds a per-user gaze-reliability gate. The gate uses ten labelled gaze trials, no EEG, and passes if the left–right separation has the pooled sign and exceeds 50 px (Eye-BCI) or 0.5 pooled s.d. (2a); gate failure routes every command to EEG. Scenarios: target visible throughout (both datasets), target absent (2a), and a 50/50 mixture (2a). Five predictions were registered and each outcome is reported.

3. RESULTS AND DISCUSSION

3.1. Under a lateral cue, one gaze feature carries the cross-subject signal

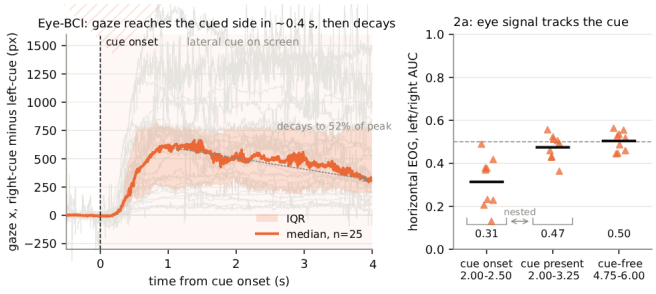
On Eye-BCI, eye features reach calibration-free AUC 0.898 against 0.678 for CSP-EEG. Mean horizontal gaze position alone reaches 0.866; the other ten eye features together reach 0.511, and pupil and fixation features are at chance. Twenty-seven of 30 participants shift gaze toward the cued side (median separation 481 px) and average 0.940; the three who hold fixation average 0.519 (Mann–Whitney $p = 0.005$). Gaze separation predicts eye-only accuracy ($r = 0.53$, $p = 0.003$) but not CSP-EEG accuracy ($r = 0.07$, 95% CI $[-0.30, 0.42]$), which argues against an ocular contribution to the EEG result without excluding it.

3.2. The gaze response is a fast orienting movement that decays

In 25 of 27 participants gaze reaches half of its final displacement 413 ms after the cue (IQR 383–443 ms; bootstrap 95% CI of the median 393–443 ms) and 90% at 680 ms. In 21 it is complete within one second of a four-second imagery period in Figure 1(a). The deflection then declines—smaller late than early in 20 of 25 participants, $p = 0.0004$—but never returns to baseline. A response that saturates within half a second, with a 60 ms interquartile range across people, is an orienting response to a stimulus, not a correlate of sustained imagery.

3.3. Removing the cue removes the eye signal; calibrated EEG is unchanged

In 2a the horizontal EOG separates left from right in the 0.5 s after cue onset, with the same sign in all nine participants (AUC 0.31, $p = 0.004$). Once the cue is gone it is at chance (0.50, $p = 0.82$; Figure 1(b), Table 1). Table 1 is the sensitivity analysis for EOG drift correction: the conclusion depends on the treatment, so all three treatments are reported. Whole-epoch detrending restores a cue-free effect, but the whole-epoch slope is itself class-predictive (AUC 0.62, $p = 0.008$), so it removes part of a sustained deflection. The pre-cue slope, where no class signal exists, is the valid estimate. Calibrated EEG accuracy is 0.72 with the cue and 0.68 without ($p = 0.43$).



(a) (b)

Figure 1. The eye signal follows the cue (a) Eye-BCI, lateral cue on throughout: horizontal gaze on right-minus left-cue trials, median and interquartile range over the 25 participants with a timeable deflection (grey, individual traces) and (b) 2a, central arrow removed at 3.25 s: horizontal-EOG AUC per participant (dots) and mean (bars); the onset window is nested in the cue-present window

Table 1. 2a horizontal-EOG AUC under three drift treatments (mean of 9; p, Wilcoxon against 0.5; participants below 0.5). Method C is used throughout; B is shown because it would reverse the conclusion

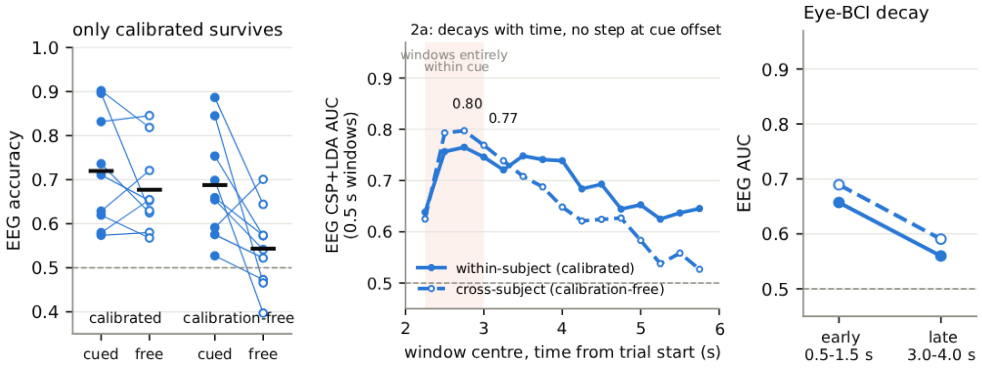
Drift treatment	onset 2.00–2.50 s	cue present 2.00–3.25 s	cue free 4.75–6.00 s
A none	0.223 ($p = 0.004$; 9/9)	0.427 ($p = 0.13$; 7/9)	0.454 ($p = 0.13$; 6/9)
B whole-epoch detrend	0.198 ($p = 0.004$; 9/9)	0.388 ($p = 0.020$; 8/9)	0.344 ($p = 0.004$; 9/9)
C pre-cue slope only	0.314 ($p = 0.004$; 9/9)	0.474 ($p = 0.36$; 5/9)	0.504 ($p = 0.82$; 4/9)

3.4. Calibration-free EEG is time-limited, not cue-locked

The calibration-free decoder falls from AUC 0.794 with the cue to 0.587 without, worse in 9 of 9 participants ($p = 0.004$, $r_{rb} = 1.00$, $d_z = 1.34$; Table 2; Figure 2(a)). The signal it loses is not ocular. It is weakest at cue onset, where the saccade occurs (0.63), and strongest afterwards (0.81). Central channels carry it better than frontal ones (0.76 vs 0.68), the EOG channels alone decode only at onset, and regressing the EOG out of the EEG leaves it intact (0.79 cued). Nor is it cue-locked. In sliding windows it peaks at 0.80 around 2.5–3.0 s and is already declining before the arrow is removed, with no step at offset (Figure 2(b)). The calibrated decoder holds near 0.74 until 4.25 s. Eye-BCI, where the cue never leaves the screen, decays identically between early and late windows, calibrated ($0.657 \rightarrow 0.560$) and calibration-free ($0.689 \rightarrow 0.591$), with no difference in rate ($p = 0.93$; Figure 2(c)). What generalises across people is the early response; it wanes regardless of the cue.

Table 2. Paired cued-versus-cue-free contrasts ($n = 9$, 2a): difference Δ, bootstrap 95% CI, Wilcoxon p, rank-biserial r_{rb} and Cohen's d_z. Positive Δ means worse without the cue. Values are generated from the result tables

Contrast	Δ	95% CI	p	r_{rb}	d_z
Calibrated EEG accuracy	+0.043	[−0.020, +0.113]	0.43	0.33	0.39
Calibration-free EEG AUC	+0.207	[+0.122, +0.313]	0.004	1.00	1.34
Calibration-free EEG accuracy	+0.144	[+0.055, +0.233]	0.027	0.82	0.97
Sim. success, calibrated EEG	+0.058	[−0.045, +0.166]	0.43	0.33	0.34
Sim. success, calibration-free EEG	+0.242	[+0.069, +0.398]	0.039	0.78	0.90
Sim. success, gaze controller (onset)	+0.211	[+0.075, +0.347]	0.020	0.87	0.96



(a) (b) (c)

Figure 2. The EEG signal is time-limited, not cue-locked (a) 2a, equal cue-present and cue-free windows, paired per participant, (b) 2a, sliding 0.5 s windows at their centre; the band marks windows lying entirely within the cue, and (c) Eye-BCI, early versus late windows, cue on throughout. Solid and filled, calibrated; dashed and open, calibration-free

3.5. What the difference costs in closed loop

Table 3 reports the simulation. The calibrated EEG decoder's success transfers to cue-free operation ($p = 0.43$), and its offline cued accuracy overstates cue-free success by 0.017. The calibration-free decoder's does not: success falls from 69.5% to 45.3% ($p = 0.039$), the wrong target is reached in 35.7% of cue-free episodes, and offline accuracy overstates deployed success by 0.235. A gaze-driven controller on the cue-onset window reaches the correct target in 62.1% of cued episodes—two participants above 90%—and 41.0% of cue-free ones ($p = 0.020$, 8 of 9 worse), with the wrong-target rate doubling to 0.42. In the pre-registered 1.25 s windows the EOG decoder was at chance with and without the cue, so its registered collapse was not testable; the sensitivity variants leave every comparison unchanged ($p \geq 0.36$).

Table 3. Closed-loop simulation, mean over 9 participants. Pre-registered and calibration-free runs use 1.25 s windows; the exploratory run uses the 0.5 s cue-onset window and an equal cue-free window. Values are generated from the result tables

Run	Decoder	Window	Offline acc.	Task success	Wrong target
pre-registered	EEG	cued	0.719	0.760	0.088
	EEG	cue-free	0.677	0.702	0.111
	EOG	cued	0.500	0.402	0.428
	EOG	cue-free	0.501	0.418	0.416
	EEG+EOG	cued	0.720	0.763	0.088
	EEG+EOG	cue-free	0.671	0.691	0.119
calibration-free	EEG	cued	0.688	0.695	0.220
	EEG	cue-free	0.543	0.453	0.357
	EEG+EOG	cued	0.690	0.690	0.224
	EEG+EOG	cue-free	0.538	0.444	0.371
exploratory	EOG	cued onset	0.629	0.621	0.203
	EOG	cue-free	0.504	0.410	0.416
	EEG	cued onset	0.607	0.579	0.174
	EEG	cue-free	0.615	0.582	0.180

3.6. Mitigations tested in closed loop

With a target visible, a gaze decoder that needs no per-user calibration reaches 0.82 task success against 0.49 for calibration-free EEG (27 of 30 participants, $p < 10^{-5}$; Table 4). Its wrong-target rate is a third of the EEG decoder's (0.10 vs 0.36). Naive fusion reaches the same success: an LDA on EEG and gaze features becomes a gaze decoder. The registered prediction that explicit arbitration would beat fusion or a single decoder *failed*: A1 equals fusion with a target visible ($p = 0.86$) and beats nothing in the mixed scenario (all $p = 0.65$, $n = 9$). Two mitigations do hold. A confidence gate of $\tau = 0.2$ cuts wrong-target commands from 0.10 to 0.04 for 0.04 of success on Eye-BCI; on 2a, where posteriors are weaker, it halves success. The

gaze-reliability gate correctly refuses the three fixating participants, but ten trials is too few: across random draws, fixators pass 15% of the time; forty trials gives 5% against 99% for the rest.

Table 4. Mitigation study: task success / wrong-target rate, calibration-free decoders, mean over participants (Eye-BCI $n = 30$, 2a $n = 9$). G gaze, E EEG, F fusion, A1 arbitration by target visibility, A2 with the gaze-reliability gate.

Controller	Eye-BCI, target visible	2a, target visible	2a, target absent	2a, mixed
G	0.816 / 0.103	0.646 / 0.192	0.495 / 0.499	0.573 / 0.343
E	0.489 / 0.361	0.697 / 0.217	0.440 / 0.378	0.570 / 0.293
F	0.808 / 0.110	0.703 / 0.219	0.432 / 0.386	0.573 / 0.298
A1	0.816 / 0.103	0.646 / 0.192	0.440 / 0.378	0.548 / 0.276
A2	0.823 / 0.096	0.734 / 0.139	0.440 / 0.378	0.585 / 0.258
A2, $\tau = 0.2$	0.784 / 0.038	0.350 / 0.055	0.205 / 0.095	0.285 / 0.074

3.7. Design rules for MI-controlled service robots

3.7.1. Present cues centrally and hold fixation

The gaze shortcut tracks cue geometry: dominant under a lateral rectangle shown for 4 s, a half-second transient under a central arrow with a held fixation cross. The confound is made by the protocol and is preventable at design time.

3.7.2. Run a gaze-ablation test before deployment

Record eye position or EOG during calibration and report accuracy with and without gaze features, and with EEG alone in a cue-free window. If performance depends on gaze, the system is a gaze interface and should be engineered as one. The test costs an EOG pair.

3.7.3. Use gaze as a declared, zero-calibration channel when a target is visible—and verify it per user

With a target on screen, gaze outperforms calibration-free MI by a wide margin and needs no training session. It is useless without a target, and for fixating users a pooled gaze decoder can steer the robot the wrong way. A forty-trial per-user gaze check provides that verification; a ten-trial one does not.

3.7.4. Evaluate on the window the robot will use; calibrate per user or gate on confidence

The cross-subject EEG signal is strongest one to two seconds after cue onset and wanes thereafter, cue or no cue. A calibration-free decoder validated early overstated deployed success by 0.24 here. Report time-resolved accuracy, evaluate late, and prefer per-user calibration.

3.8. Why the failure is silent

Both failing controllers—gaze-latched and calibration-free EEG—raised their wrong-target rate while their commands-to-success barely changed: the robot did not hesitate, it went the wrong way. For users with advanced amyotrophic lateral sclerosis or locked-in syndrome, voluntary eye movement may itself be absent, so a gaze-dependent decoder fails exactly where it is most needed [21].

3.9. Practical implications for deployment

For clinical and daily-life use, rules 2–4 applied in order form the pre-handover procedure; one step is not covered by them: choose the confidence-gate threshold from replayed posteriors against a wrong-target ceiling, since the gate that makes a gaze channel safer silences a low-spread EEG decoder. In daily-living terms a wrong-target rate of 0.36 means one cue-free command in three ends in an unintended action; and the rules carry over to wheelchair heading, lateral arm reach and exoskeleton stepping alike. For BCI engineers the summary is one sentence: report class-wise accuracy, abstention rate and a cue-free evaluation, because a single cued accuracy figure predicts neither safety nor throughput. Tables 3 and 4 give the comparative view across all cued and cue-free conditions.

3.10. From simulation to hardware

The evaluation is simulation-only by design: it replays real human decoder outputs, so it measures the transfer of decoder behaviour, not robot dynamics. The roadmap is: i) Hardware-in-the-loop replay, feeding the same recorded posteriors to a physical wheelchair or tabletop-arm controller so that dynamics, latency and stopping behaviour enter; ii) Online cue-free sessions with able-bodied users under the design rules, gated by the gaze-ablation test; and iii) Supervised trials with motor-impaired users, where the wrong-target rate, not

accuracy, is the primary endpoint. The released code is structured so that step (i) replaces only the simulator's step function. Adaptive transfer learning and reinforcement-signal-driven decoders [16, 17] are the natural candidates for closing the calibration-free gap within this roadmap, provided they are evaluated cue-free under the same rules.

3.11. Limitations

The large gaze shortcut was measured in a dataset with no cue-free window. Its deployed failure is therefore inferred from 2a, not replayed from the same participants. 2a has nine participants and one session each; its nulls have limited power (detectable $d_z \approx 1.0$ at 80%). In 2a the cue-free window is also the late window, so cue removal and elapsed time are separated only by the pre-offset decline and the Eye-BCI comparison. The calibration-free simulation arm was added after the first specification; the simulation is a kinematic abstraction with no dynamics, shared autonomy or physical robot. Participant S06 could not be aligned and is excluded; S20 is excluded from latency statistics for a pre-cue gaze difference.

4. CONCLUSION

Under a lateral cue, a motor-imagery decoder can reach high cross-subject accuracy from eye position alone—a signal that appears within half a second of the cue and disappears with it—while the cross-subject EEG component wanes within seconds of onset, so a calibration-free decoder validated early overstates what it will deliver. Central cues, a gaze-ablation test, gaze declared and verified as its own channel, and evaluation on the deployment window—with per-user calibration or a confidence gate—address them. Beyond deployment, the study is a compact teaching case for robotics and human–machine interaction courses: a protocol confound, a pre-registered test, a failed prediction reported honestly, and design rules derived from closed-loop cost. The released scripts map one-to-one onto the tables and figures and carry both pre-registered specifications with their version history, so the analyses can be re-run or extended; hardware-in-the-loop replay and trials with motor-impaired users are next.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Fouad Fakhet	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓		✓	
Rachida Djeradi	✓	✓		✓						✓		✓		

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal Analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject Administration
Fu : **Fu**nding Acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

ETHICAL APPROVAL

This study is a secondary analysis of two public, fully de-identified datasets whose original collection was approved by the ethics committees named in the source publications; no new data were collected and no additional approval was required.

DATA AVAILABILITY

Both datasets are public: Synapse accession syn64005218 (CC0) and `https://www.bbci.de/competition/iv/`. The analysis code, synchronisation and gaze-ablation procedures, both pre-registered

specifications with their version history, and every result table are supplied as supplementary material and will be deposited publicly upon acceptance.

REFERENCES

[1] M. Aljalal, S. Ibrahim, R. Djemal, and W. Ko, "Comprehensive review on brain-controlled mobile robots and robotic arms based on electroencephalography signals," *Intelligent Service Robotics*, vol. 13, pp. 539–563, 2020, doi: 10.1007/s11370-020-00328-5.

[2] J. D. R. Millán, "Combining brain-computer interfaces and assistive technologies: state-of-the-art and challenges," *Frontiers in Neuroscience*, vol. 1, 2010, doi: 10.3389/fnins.2010.00161.

[3] F. Galán, M. Nuttin, E. Lew, P. Ferrez, G. Vanacker, J. Philips, and J. d. R. Millán, "A brain-actuated wheelchair: Asynchronous and non-invasive Brain–computer interfaces for continuous control of robots," *Clinical Neurophysiology*, vol. 119, no. 9, pp. 2159–2169, 2008, doi: 10.1016/j.clinph.2008.06.001.

[4] T. Carlson and J. del R. Millan, "Brain-Controlled Wheelchairs: A Robotic Architecture," *IEEE Robotics & Automation Magazine*, vol. 20, no. 1, pp. 65–73, 2013, doi: 10.1109/mra.2012.2229936.

[5] J. Meng, S. Zhang, A. Bekyo, J. Olsoe, B. Baxter, and B. He, "Noninvasive electroencephalogram based control of a robotic arm for reach and grasp tasks," *Scientific Reports*, vol. 6, no. 1, 2016, doi: 10.1038/srep38565.

[6] M. Plöchl, J. P. Ossandón, and P. König, "Combining EEG and eye tracking: identification, characterization, and correction of eye movement artifacts in electroencephalographic data," *Frontiers in Human Neuroscience*, vol. 6, 2012, doi: 10.3389/fnhum.2012.00278.

[7] L. Tonin, S. Perdikis, T. D. Kuzu, J. Pardo, B. Orset, K. Lee, M. Aach, T. A. Schildhauer, R. Martínez-Olivera, and J. d. R. Millán, "Learning to control a BMI-driven wheelchair for people with severe tetraplegia," *iScience*, vol. 25, no. 12, p. 105418, 2022, doi: 10.1016/j.isci.2022.105418.

[8] Q. Ai, M. Zhao, K. Chen, X. Zhao, L. Ma, and Q. Liu, "Flexible coding scheme for robotic arm control driven by motor imagery decoding," *Journal of Neural Engineering*, vol. 19, no. 5, p. 056008, 2022, doi: 10.1088/1741-2552/ac84a9.

[9] A. Fischer-Janzen, T. M. Wendt, and K. Van Laerhoven, "A scoping review of gaze and eye tracking-based control methods for assistive robotic arms," *Frontiers in Robotics and AI*, vol. 11, 2024, doi: 10.3389/frobt.2024.1326670.

[10] S. R. Soekadar, M. Witkowski, C. Gómez, E. Opisso, J. Medina, M. Cortese, M. Cempini, M. C. Carrozza, L. G. Cohen, N. Birbaumer, and N. Vitiello, "Hybrid EEG/EOG-based brain/neural hand exoskeleton restores fully independent daily living activities after quadriplegia," *Science Robotics*, vol. 1, no. 1, 2016, doi: 10.1126/scirobotics.aag3296.

[11] A. Dillen, M. Omidi, F. Ghaffari, B. Vanderborght, B. Roelands, O. Romain, A. Nowé, and K. De Pauw, "A shared robot control system combining augmented reality and motor imagery brain–computer interfaces with eye tracking," *Journal of Neural Engineering*, vol. 21, no. 5, p. 056028, 2024, doi: 10.1088/1741-2552/ad7f8d.

[12] C. Palanichamy, S. P. Thirumoorthi, K. Lakshminarayanan, D. Madathil, and M. H. Rahman, "Multimodal brain-computer interface for robotic control: integration of real-time gaze tracking and EEG-based motor imagery," *Medical & Biological Engineering & Computing*, vol. 64, pp. 781–792, 2025, doi: 10.1007/s11517-025-03489-6.

[13] A. D'Aquino and T. Schack, "Bridging cognition and control through passive eye movement integration in motor imagery brain-computer interfaces," *Frontiers in Human Neuroscience*, vol. 20, 2026, doi: 10.3389/fnhum.2026.1849674.

[14] H. He and D. Wu, "Transfer learning for brain–computer interfaces: a Euclidean space data alignment approach," *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 2, pp. 399–410, 2020, doi: 10.1109/TBME.2019.2913914.

[15] D. Wu, Y. Xu, and B.-L. Lu, "Transfer learning for EEG-based brain–computer interfaces: a review of progress made since 2016," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 14, no. 1, pp. 4–19, 2022, doi: 10.1109/tcds.2020.3007453.

[16] K. Zhang, N. Robinson, S.-W. Lee, and C. Guan, "Adaptive transfer learning for EEG motor imagery classification with deep Convolutional Neural Network," *Neural Networks*, vol. 136, pp. 1–10, 2021, doi: 10.1016/j.neunet.2020.12.013.

[17] A. Xavier Fidêncio, F. Grün, C. Klaes, and I. Iossifidis, "Hybrid brain-computer interface using error-related potential and reinforcement learning," *Frontiers in Human Neuroscience*, vol. 19, 2025, doi: 10.3389/fnhum.2025.1569411.

[18] G. Pfurtscheller and F. Lopes da Silva, "Event-related EEG/MEG synchronization and desynchronization: basic principles," *Clinical Neurophysiology*, vol. 110, no. 11, pp. 1842–1857, 1999, doi: 10.1016/S1388-2457(99)00141-8.

[19] H. Ramoser, J. Muller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 4, pp. 441–446, 2000, doi: 10.1109/86.895946.

[20] R. Geirhos, J.-H. Jacobsen, C. Michaelis, R. Zemel, W. Brendel, M. Bethge, and F. A. Wichmann, "Shortcut learning in deep neural networks," *Nature Machine Intelligence*, vol. 2, no. 11, pp. 665–673, 2020, doi: 10.1038/s42256-020-00257-z.

[21] P. Brunner, S. Joshi, S. Briskin, J. R. Wolpaw, H. Bischof, and G. Schalk, "Does the "P300" speller depend on eye gaze?" *Journal of Neural Engineering*, vol. 7, no. 5, p. 056013, 2010, doi: 10.1088/1741-2560/7/5/056013.

[22] J. Meng, Z. Wu, S. Li, and X. Zhu, "Effects of gaze fixation on the performance of a motor imagery-based brain-computer interface," *Frontiers in Human Neuroscience*, vol. 15, 2022, doi: 10.3389/fnhum.2021.773603.

[23] P. Shenoy, M. Krauledat, B. Blankertz, R. P. N. Rao, and K.-R. Müller, "Towards adaptive classification for BCI," *Journal of Neural Engineering*, vol. 3, no. 1, pp. R13–R23, 2006, doi: 10.1088/1741-2560/3/1/r02.

[24] E. Guttmann-Flury, X. Sheng, and X. Zhu, "Dataset combining EEG, eye-tracking, and high-speed video for ocular activity analysis across BCI paradigms," *Scientific Data*, vol. 12, no. 1, p. 587, 2025, doi: 10.1038/s41597-025-04861-9.

[25] M. Tangermann, K.-R. Müller, A. Aertsen, N. Birbaumer, C. Braun, C. Brunner, R. Leeb, C. Mehring, K. J. Miller, G. R. Müller-Putz, G. Nolte, G. Pfurtscheller, H. Preissl, G. Schalk, A. Schlögl, C. Vidaurre, S. Waldert, and B. Blankertz, "Review of the BCI Competition IV," *Frontiers in Neuroscience*, vol. 6, p. 55, 2012, doi: 10.3389/fnins.2012.00055.

BIOGRAPHIES OF AUTHORS



Fouad Fakhet is a Ph.D. student in telecommunications at USTHB, Algiers, Algeria, where he also works as an engineer and conducts research within the LCPTS laboratory. His interests include brain–computer interfaces and multimodal signal processing. He designed and ran the analyses, wrote the software, and drafted the manuscript. He can be contacted at email: ffakhet@usthb.dz.



Rachida Djeradi is a full professor with the Department of Telecommunications, Faculty of Electronics and Computer Science, USTHB, Algiers, Algeria, where she heads the PCPS team of the LCPTS laboratory. Her research spans speech signal processing, acoustic phonetics, and natural language understanding. She contributed to the conceptualisation and methodology, validated the results, and supervised the work. She can be contacted at email: rdjeradi@usthb.dz.