

ErgoAssist: Cognition-Aware Posture Feedback in Wearable Ergonomic Systems

SARMISTHA SARNA GOMASTA, University of Massachusetts Amherst, USA

BHAWANA CHHAGLANI, University of Massachusetts Amherst, USA

VP NGUYEN, University of Massachusetts Amherst, USA

PRASHANT SHENOY, University of Massachusetts Amherst, USA

Prolonged digital device use has made poor posture and musculoskeletal discomfort pervasive among knowledge workers. Existing ergonomic wearables rely solely on posture thresholds, frequently interrupting users during high-focus moments and leading to alert fatigue and abandonment. Yet posture and cognitive load are closely coupled, and most systems remain cognitively unaware. We present *ErgoAssist*, a head-worn ergonomic assistant that detects poor posture using IMU-based head tracking and estimates task-induced cognitive load using a consumer-grade EEG headband for continuous everyday use. In a controlled lab study, *ErgoAssist* achieves 81% posture classification and 90.2% task induced cognitive load estimation accuracy under leave-one-subject-out evaluation. In a preliminary real-time deployment, cognition-aware alerting reduces alert frequency by 81%, improves perceived usability by 43%, task performance by 25%, and improves posture correction rate by 38%, delivering fewer but better-timed interventions rather than merely suppressing alerts.

CCS Concepts: • **Human-centered computing** → **Ubiquitous and mobile computing**; **Human computer interaction (HCI)**.

Additional Key Words and Phrases: Multi-modal Sensing, Cognitive Modeling, Cognition-aware Ergonomic Feedback

ACM Reference Format:

Sarmistha Sarna Gomasta, Bhawana Chhaglan, Vp Nguyen, and Prashant Shenoy. 2026. ErgoAssist: Cognition-Aware Posture Feedback in Wearable Ergonomic Systems. *Proc. ACM Hum.-Comput. Interact.* 10, 5, Article MHCI8780 (August 2026), 40 pages. <https://doi.org/10.1145/3821691>

1 Introduction

Prolonged use of digital devices has made discomfort, musculoskeletal strain, eye strain, and fatigue pervasive among office workers and students [12, 27]. In particular, forward head posture (FHP) substantially increases load on the cervical spine, with even modest neck flexion leading to fatigue and long-term strain [7, 83]. In response, there has been a rich literature on sensing user posture using different modalities like wireless [74], acoustics [11, 95], and vision [17]. A growing body of commercial and research systems leverage inertial measurement units (IMUs) to monitor posture and provide corrective feedback through wearables and workstation aids [33, 77]. Despite their promise, posture-aware ergonomic systems often fail to achieve sustained adoption, deployment, and social acceptance. The major challenges faced by these posture assistants are *alert*

Authors' Contact Information: Sarmistha Sarna Gomasta, Manning College of Information and Computer Sciences, University of Massachusetts Amherst, Amherst, Massachusetts, USA, sgomasta@umass.edu; Bhawana Chhaglan, Manning College of Information and Computer Sciences, University of Massachusetts Amherst, Amherst, Massachusetts, USA, bchhaglan@umass.edu; Vp Nguyen, Manning College Information and Computer Sciences, University of Massachusetts Amherst, Amherst, Massachusetts, USA, phuc@umass.edu; Prashant Shenoy, Manning College of Information and Computer Sciences, University of Massachusetts Amherst, Amherst, Massachusetts, USA, shenoy@cs.umass.edu.



This work is licensed under a Creative Commons Attribution 4.0 International License.

© 2026 Copyright held by the owner/author(s).

ACM 2573-0142/2026/8-ARTMHCI8780

<https://doi.org/10.1145/3821691>

fatigue, increased anxiety or over-correction, and inaccurate feedback [15, 112]. When interventions occur too frequently, users become desensitised, ignore alerts, or actively disable them, reducing overall effectiveness and increasing annoyance. Studies of alerting systems in healthcare and human-computer interaction have documented how excessive or poorly timed alerts significantly degrade user experience and compliance [2, 51]. Existing systems typically issue alerts whenever a posture threshold is violated, regardless of the user's mental state or task context. As a result, users are frequently interrupted during moments of high concentration, leading to annoyance and eventual abandonment of the system [9]. These observations highlight a fundamental limitation of posture-only ergonomic assistants: they treat feedback as a purely physical correction problem, rather than an interruptibility-sensitive interaction [2, 56]. Prior work suggests that posture and cognition are tightly coupled. Poor posture can negatively impact cognitive performance [58], while high cognitive load often exacerbates postural degradation and muscular strain [71]. Importantly, moments of high cognitive demand are also when users are least receptive to interruptions [87]. Yet, *most existing ergonomic systems remain cognitively unaware*, issuing alerts when users are deeply focused, leading to annoyance. This motivates a key question: *How can ergonomic feedback be made more effective by being cognition-aware?* Rather than alerting users whenever posture deteriorates, an intelligent ergonomic assistant should reason about *when* to intervene delivering fewer, better-timed corrections during moments of low cognitive focus or natural breaks in attention. *We explore how we can sense holistic state of user (including both their posture and cognitive context) to generate human-centered adaptive interventions.*

Recent advances in wearable sensing make this possible. Consumer-grade head-worn EEG devices and ear-integrated sensors enable continuous, non-invasive estimation of task-induced cognitive load in everyday settings [26, 39, 65, 66]. When combined with posture sensing, these modalities offer complementary signals: IMUs capture physical strain, while EEG provides insight into task-induced cognitive load and receptivity to feedback. We study how these physical and physiological signals vary over time and correlate with each other when a user is performing different tasks, and leverage these insights to design context-sensitive interventions. In this work, we present *ErgoAssist*, a wearable-based cognition-aware ergonomic assistant that integrates posture sensing with cognitive-state-aware feedback scheduling. Our system continuously monitors user's neck posture using an IMUs and estimates task-induced cognitive load using wearable EEG. Instead of issuing posture alerts immediately, the system adapts feedback timing based on the user's cognitive context, delaying or suppressing alerts during periods of high focus and prioritizing interventions during low-load moments or natural breakpoints [87].

We evaluate our approach through a controlled user study (N=24) in which participants perform three tasks (Rest, Stroop, Numerical Calculation) under multiple postural conditions while wearing our head-mounted IMU and EEG band. We compare two feedback strategies: (1) posture-only alerts triggered by neck angle thresholds, and (2) cognition-aware alerts that fuse posture and task-induced cognitive load to schedule feedback. Our results provide empirical evidence that posture-only alerts are often poorly timed, frequently interrupting users during periods of high cognitive focus. We show that cognition-aware alerts are significantly less frequent (81% less), less disruptive, and more effective, leading to improved perceived usability (43% more), higher posture correction rate (61.9% to 100%), and improved task performance (25% more) compared to posture-only feedback. Overall, this paper makes the following contributions:

- (1) We design and implement a wearable ergonomic assistant that integrates IMU-based posture sensing with EEG-based task-induced cognitive load estimation to continuously track the holistic state of the user. The system achieves **90.2%** accuracy for task-induced cognitive load classification and **81.4%** posture classification under LOSO evaluation.

- (2) We analyze the interaction of physical (postural) and physiological (cognitive) signals during cognitively demanding tasks, and examine how cognitive state influences posture behavior.
- (3) We predict perceived discomfort levels reported by users using multimodal sensor fusion and lightweight machine learning models, achieving 75% accuracy with a latency of **12.98 ms**.
- (4) Through a preliminary real-time within-subject study (N=13), we show that cognition-aware interventions reduce alert frequency by **81%**, higher posture correction rates by **38%**, and improve perceived usability by **43%** compared to posture-only feedback.

2 Background and Motivation

In this section, we motivate the need for a cognition-aware ergonomic assistants in everyday work environments to promote good postural habits without affecting performance.

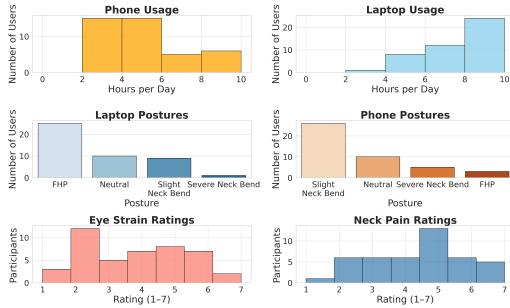
2.1 Need for ergonomic assistant: A User Survey

Prolonged screen-based work is increasingly associated with musculoskeletal discomfort, eye strain, and fatigue [12, 27]. FHP is particularly concerning as even modest neck flexion increases mechanical load on the cervical spine and accumulates into pain over time [7]. To understand real-world challenges, we conduct an IRB-approved pre-survey with 50 participants aged 18–45. Participants report an average of 7.36 hours/day on laptops and 4.24 hours/day on phones, with 77.3% and 75.0% respectively using ergonomically poor postures. FHP is the most common laptop posture, contributing to neck pain and cervical spondylosis (tech neck [8, 19]). Overall, 93.2% of users report eye strain and 95.5% report neck or upper back pain, yet none use any posture assistant (Figure 1a). Thus, there is a need for an ergonomic assistant that continuously monitors our state and provide timely alerts.

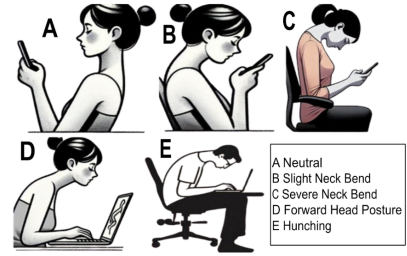
2.2 Towards Cognition-Aware Ergonomic Assistance

A growing body of commercial and research systems use IMUs to detect slouching or excessive flexion and trigger corrective feedback [33, 77]. However, these systems often struggle with sustained adoption in real-world settings. A key reason is that alerts are frequently delivered at the wrong time. Posture thresholds are evaluated independent of the user's task-induced cognitive load, task demands, or receptivity to interruption, leading to repeated notifications during high-focus moments that feel disruptive and contribute to alert fatigue [9]. Interruption science further shows that timing matters, and alerts delivered during high-focus moments are more disruptive and less likely to be heeded than those aligned with natural breakpoints [2, 24, 56, 87]. Recent advances in consumer-grade wearable EEG devices now enable practical non-invasive monitoring of task-induced cognitive load in everyday contexts [26, 39], enabling real-time integration of cognitive state into systems that decide *when* to deliver feedback. This is further supported by the strong bidirectional relationship between posture and cognition. Cognitive load can degrade posture and increase muscle strain [55], and poor posture can impair cognitive performance and alter neural patterns [58, 71]. Yet most existing posture feedback tools such as Upright [33] and SuperBreak [103] rely solely on posture thresholds independent of cognitive context, risking over-intervention during precisely those moments when users are least receptive [9, 78].

Taken together, these observations motivate the need for cognition-aware ergonomic assistants that reason about a holistic user state encompassing both physical strain (posture) and psychological state (task-induced cognitive load and receptivity). By fusing IMU-based posture sensing with EEG-based task-induced cognitive load estimation, a wearable assistant can deliver better-timed interventions, suppressing or deferring alerts during high-focus periods and prioritizing feedback



(a) User Pre-survey Results.



(b) Posture Categories [19].

Fig. 1. User Pre- Survey context. (a) Results from the motivation survey and (b) Demo posture categories included in survey.

during natural breakpoints. Our proposed system aims to sense holistic user state and focuses on delivering *human-centered adaptive interventions*.

3 Related Work

Wearable sensing has been widely adopted for monitoring either physical ergonomics or cognitive states. However, few systems have explored their integration to understand the interplay between mental workload and physical posture.

3.1 Posture Sensing for Ergonomic Risk

Accurate sensing of head and neck posture has been widely explored to detect ergonomically harmful configurations such as forward head flexion and sustained non-neutral alignment. Vision-based systems [17, 32, 73, 125] leverage pose estimation and skeletal tracking to classify a wide range of postures with high spatial fidelity. However, they are computationally intensive, raise privacy concerns, and degrade under occlusion or low-light conditions. Wireless sensing approaches [35, 74, 95, 108] infer posture through RF signal perturbations without requiring body-worn devices, but are highly sensitive to environmental dynamics such as furniture layout and user position, reducing robustness across settings. Pressure-based systems [82, 113] instrument seating interfaces to estimate body posture via contact distribution, yet remain constrained to fixed locations and cannot directly capture head or neck kinematics. Similarly, motion capture systems provide high-precision joint tracking and serve as a gold standard, but require specialized infrastructure and are impractical outside controlled laboratory environments [77, 117, 120]. Wearable sensing provides a more practical alternative for continuous monitoring. IMU-based systems estimate orientation, angular velocity, and derived neck angles to classify posture in real time [33, 60, 72, 93]. While accurate, these approaches often require rigid placement on the cervical spine or upper back, impacting comfort and long-term adherence. An increasingly practical direction is to leverage opportunistic sensing from existing head-worn devices, such as headphones or smart wearables, to estimate head kinematics [19, 37, 46, 61]. These systems utilize embedded inertial sensors to continuously capture orientation and motion dynamics with minimal user burden. Such approaches enable real-time, fine-grained tracking of head posture in everyday settings without requiring dedicated instrumentation, making them well-suited for scalable ergonomic monitoring.

3.2 Wearable EEG for Cognitive Load Monitoring

Wearable brain sensing now supports inference of cognitive states such as mental workload, fatigue, and attentional engagement, with research moving from bulky EEG caps to lightweight headbands, earbuds, and EEG-equipped headphones [34, 39, 48, 63, 67]. Knierim et al. showed that headphone-based EEG can classify cognitive load in lab and field settings [63], and Kwon et al. demonstrated continuous workload monitoring using wearable EEG bands [67]. Most approaches represent EEG activity using band-limited spectral features, with frontal theta, alpha, and beta power systematically modulated by cognitive workload [14, 20, 94, 96]. Frontal theta typically increases with task difficulty and resource demands, whereas alpha decreases under higher workload and is sensitive to fatigue and attentional load [54, 94, 106]. Beta activity has been associated with visual attention and short-term memory processes in working-memory and vigilance tasks [76, 92, 121]. Ratios such as theta-beta and alpha-beta, alongside spectral and sample entropy features, are widely used to discriminate cognitive load levels [20, 75, 96]. Recent work has also explored machine learning approaches on wearable EEG, ranging from classical models to deep learning architectures, enabling robust classification under real-world noise and motion artifacts [4, 52, 124]. Despite this, real-world robustness remains limited. Dry electrodes, in-motion conditions, and uncontrolled environments are the three primary sources of signal degradation in wearable EEG [1, 80], yet few studies address their combined effect in mobile settings. Real-world benchmarking shows that consumer-grade devices exhibit larger drops in spectral stability and weaker neurometric discrimination during multitasking and driving, with motion sensitivity and sparse electrode coverage as the main factors [114]. Some existing systems also estimate cognitive state based on user behavior or physical context [28, 42], limiting their applicability for adaptive interactive systems. This creates opportunities to integrate cognitive state awareness into interactive systems that adapt in real time to both user state and context.

3.3 Multimodal Sensor Fusion for User State Modeling

Prior studies have explored physical and physiological data streams and their combination for fatigue and ergonomics [16, 59, 101, 116]. For example, Graña et al. demonstrated that combining an IMU's motion features with EEG features yielded more accurate recognition of natural daily activities than IMU alone [91]. Nica et al. [84] demonstrated that fusing EEG, ECG, and EMG data enables improved prediction of mental and physical fatigue, especially under multitasking conditions. Their work underscores the importance of multimodal sensing in understanding how cognitive and physical demands co-occur. Demirel et al. [26] address cybersickness in VR environments using head and physiological sensors. The field of neuroergonomics also promotes this integration, emphasizing how mental workload and ergonomic performance are tightly coupled, particularly in safety-critical and prolonged-use scenarios [25, 79]. This motivates the need for an assistant that continuously monitors both how you sit/stand and how you think/feel, enabling context-sensitive support. While prior work has explored combining physiological and physical signals for user state recognition, these approaches primarily focus on improving classification accuracy for tasks such as activity recognition or fatigue detection. They do not address how multimodal sensing can inform decision-making for real-time interventions. In particular, existing systems do not leverage multimodal signals to determine when feedback should be delivered in interactive settings, which is critical for minimizing disruption and improving user compliance.

Overall, prior work has independently explored posture sensing, cognitive load estimation, multimodal activity sensing and interruptibility-aware feedback. However, these domains remain largely disconnected. Posture-aware systems lack cognitive awareness and cognitive sensing systems do not inform intervention timing. To this end, *ErgoAssist* integrates IMU-based posture sensing and

EEG-based cognitive load estimation to drive adaptive feedback scheduling. Rather than focusing solely on state recognition, our system addresses the key question of when to intervene, enabling fewer but better-timed ergonomic corrections.

4 ErgoAssist Architecture

We design *ErgoAssist*, a head-worn wearable system that provides cognition-aware posture assistance by sensing posture via a head tracker and cognitive state via an EEG headband. The system continuously tracks the holistic user state during everyday computer use and explores the relationship between physical and physiological signals during cognitively demanding tasks. The system pipeline is shown in Figure 2. The core modules of *ErgoAssist* are: (1) a head-worn multimodal sensing device, (2) a signal processing and machine learning pipeline for posture, discomfort, and task-induced cognitive load inference, and (3) an interruptibility-aware feedback module to balance ergonomic benefit and productivity.

4.1 Data Acquisition

The headworn prototype consists of two components: a Superware Head Tracker [98] and a Frenz Brainband [39]. The head tracker includes a 6 axis IMU (3-axis accelerometer, gyroscope, and compass) providing pitch, roll, and yaw at 50 Hz with angular accuracy $< 1^\circ$. We collect EEG using the Frenz Brainband, a 4-channel wearable device (LF-FpZ, OTE_L-FpZ, RF-FpZ, OTE_R-FpZ) sampled at 125 Hz. We use the *frenzt toolkit* (v0.3.2) to stream timestamped EEG signals and spectral band powers over Bluetooth. From these, we extract δ (1–4 Hz), θ (4–7 Hz), α (8–13 Hz), β (13–30 Hz), and γ (30–45 Hz) bands. IMU data are collected asynchronously via a USB serial interface. EEG and IMU streams are aligned using system timestamps while retaining their native sampling rates (EEG: 125 Hz, IMU: 50 Hz).

4.2 Signal Processing

ErgoAssist performs continuous, real-time signal processing on multimodal sensor streams to extract posture and cognition-related features. We perform calibration and noise removal to ensure robustness to inter-participant variability and everyday motion artifacts. IMU and EEG signals are processed independently and later fused at the model level.

4.2.1 IMU Data Pre-processing. We record IMU data at 50 Hz while participants hold three head postures: D1 (neutral), D2 (mild forward flexion), and D3 (deep forward flexion). Among yaw, roll, and pitch, we use *pitch* as our primary signal because it best captures sagittal head–neck flexion, consistent with biomechanical literature on FHP [19, 36]. A key challenge in wearable-based posture monitoring is inter-participant variability in neutral head orientation. Raw pitch values during the same labeled posture can differ significantly across users due to anatomical differences and sensor placement variations [64]. We mitigate this by normalizing pitch relative to each user’s own neutral posture rather than absolute sensor coordinates. To remove startup and ending artifacts, we discard the first and last 10 seconds of each IMU recording. This alignment matches the boundary trimming applied to EEG, ensuring that subsequent multimodal analysis is performed on temporally corresponding segments. Then, for each user, we compute a neutral baseline P_{base} as the median pitch across all their D1 segments:

$$P_{\text{base}} = \text{median}(\text{pitch}_{\text{raw}}^{\text{D1}}).$$

We define the baseline-centered flexion signal as

$$\text{flex}(t) = \text{pitch}_{\text{raw}}(t) - P_{\text{base}}.$$

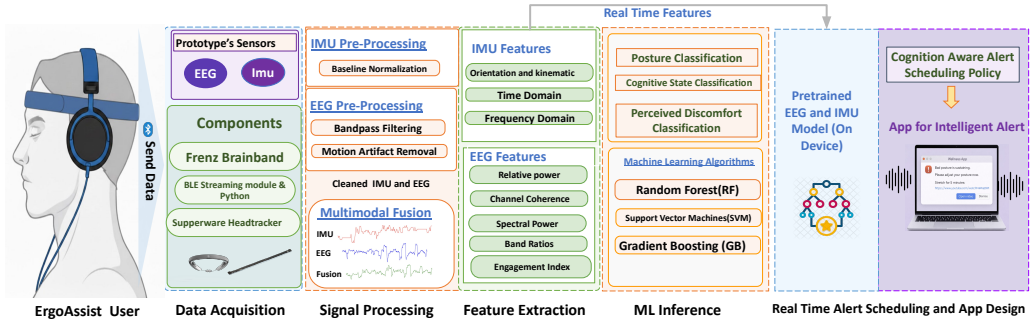


Fig. 2. *ErgoAssist* system pipeline. The prototype captures EEG and IMU signals and runs pretrained sensor-based machine learning models in real time to infer cognition-aware alerts.

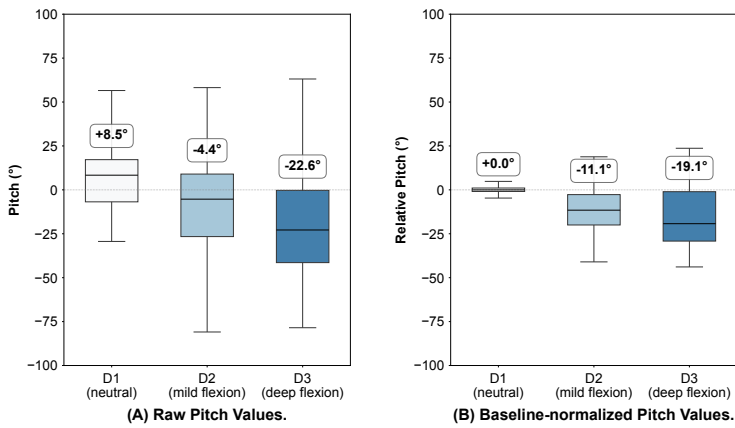


Fig. 3. Distribution of raw and baseline-normalized pitch across posture conditions. Personalized baseline normalization reduces participant-specific offset and makes the progression from D1 to D3 more interpretable.

This shifts each user's neutral posture to approximately 0° , so D2 and D3 are expressed as deviations from their personal neutral position. As shown in Figure 3, baseline normalization improves the interpretability of posture-dependent pitch patterns by reducing participant-specific offset in the raw signal. In the unnormalized representation, the pitch distributions exhibit substantial inter-subject variability. After personalized baseline calibration, the median relative pitch decreases monotonically from D1 to D3, showing clearer flexion ordering across posture conditions. Although the normalized distributions still partially overlap, they exhibit clearer ordering across D1, D2, and D3, with separation that is sufficient for binary and three-class posture classification.

4.2.2 EEG Data Pre processing. We collect raw EEG data at a sampling rate of 125 Hz using the FRENZ Brainband. We discard the first and last 10 seconds of each EEG recording to remove boundary transients caused by device settling, participant adjustment, and task onset and offset. This step is not due to synchronization error. EEG and IMU streams are aligned using their timestamps, and the same boundary trimming is applied to IMU recordings so that both modalities remain temporally aligned for downstream multimodal analysis. We validate physiological sensitivity using a standard eyes-open versus eyes-closed benchmark, where eyes-closed recordings show increased alpha-band

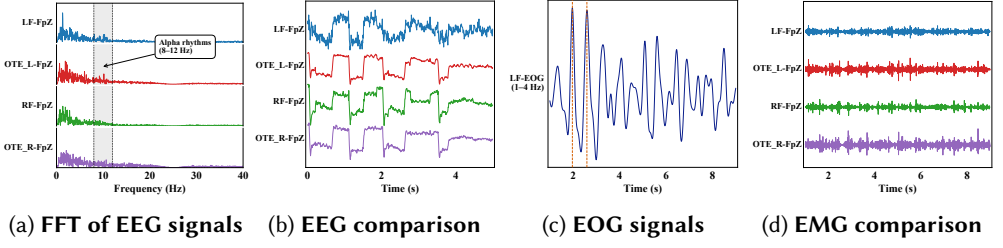


Fig. 4. EEG, EOG, and EMG signals acquired with the FRENZ Brainband. The figure illustrates frequency characteristics and time-domain of four EEG channels, eye blinks captured as EOG activity, and posture related muscle activity captured as EMG.

power (8–12 Hz) consistent with canonical neurophysiological activity (Figure 4a). Per-channel Z-normalization is applied before visualization (Figure 4b), and the device additionally captures blink-related EOG (Figure 4c) and posture-related EMG activity (Figure 4d). EEG recordings often exhibit a low signal-to-noise ratio due to contamination from ocular activity, muscle contractions, and environmental interference [81]. These artifacts overlap with typical EEG frequency bands and introduce bias in downstream features [5]. **Band-pass Filtering** We first band-pass filter (0.5–40 Hz) EEG signal using a zero-phase fourth-order Butterworth filter [122]. The high-pass cutoff reduces slow drifts caused by sweat, electrode polarization, and skin–electrode interface instabilities and the low-pass cutoff attenuates high-frequency muscle activity and other non-neural noise outside typical EEG rhythms. Zero-phase filtering (forward and reverse application) preserves the temporal alignment of neural events.

4.2.3 IMU-based Motion Artifact Removal. For multimodal data processing, we suppress motion-induced artifacts by leveraging synchronized IMU data to identify contaminated EEG segments. We compute angular velocity from IMU orientation signals and estimate RMS velocity using a 1 s sliding window. For each recording, we define an adaptive threshold based on the RMS velocity distribution, capturing subject- and trial-specific movement variability. Segments exceeding this threshold are labeled as motion-contaminated [89]. We then convert these labels into a binary IMU-derived motion mask at the native 50 Hz IMU rate and map this mask onto the 125 Hz EEG sampling grid using nearest-neighbor assignment, so that each EEG sample inherits the motion label of the temporally closest IMU sample. Because this step transfers binary artifact labels rather than interpolating continuous signal values, it does not introduce synthetic IMU measurements. We then band-pass filter the EEG (0.5 to 40 Hz), segment it into 5 s windows with 50% overlap, and discard any window in which more than 40% of samples are flagged as motion-contaminated. As shown in Figure 5, periods of increased head motion captured by the IMU align with large-amplitude, broadband distortions in the raw EEG signal. After IMU-guided artifact rejection, these motion-related components are substantially attenuated in both the time-domain waveform and the corresponding spectrogram.

4.3 Feature Extraction

IMU Feature Extraction. After IMU data processing (Section 4.2.1), we segment the pitch signal into 5 s windows with 50% overlap. From each window, we extract time-domain, derivative-based, and frequency-domain features to capture both posture magnitude and posture dynamics. We focus on extracting all features from pitch as pitch better reflects FHP than yaw and roll [19]. Time domain features such as Mean, median, and root mean square summarize the sustained level of neck

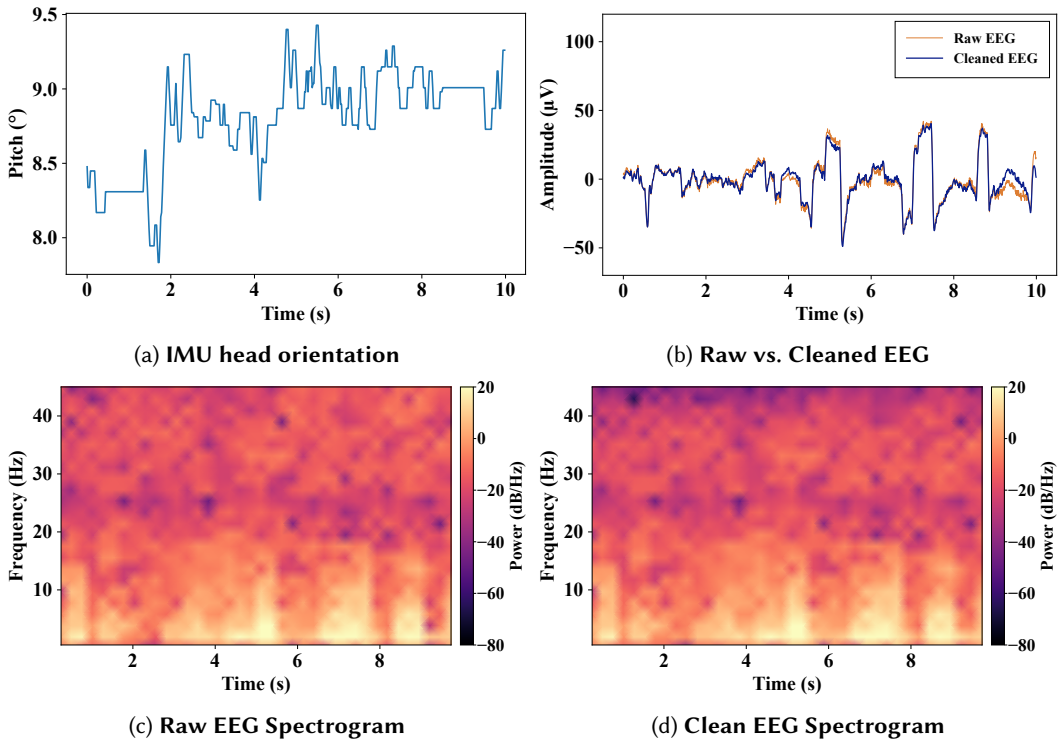


Fig. 5. EEG preprocessing pipeline for motion artifact removal. The figure demonstrates the multimodal preprocessing approach: (a) IMU data (Pitch) capturing head movements, (b) temporal comparison of raw EEG and cleaned EEG signal, (c) spectrogram of raw EEG, and (d) spectrogram of cleaned EEG.

flexion [57]. Standard deviation, range, interquartile range, skewness, and kurtosis characterize within-window variability and waveform shape [70, 126]. Zero-crossing rate and line length capture short-term fluctuations and movement irregularity [90]. We also compute velocity and acceleration based features from the first and second temporal derivatives of pitch to describe postural adjustment dynamics [57]. Derivative signals of this kind, analogous to jerk in accelerometer-based HAR, have been shown to improve discrimination between posture classes [126]. In addition, we extract Hjorth activity, mobility, and complexity as compact descriptors of signal power and spectral shape [88]. In the frequency domain, dominant frequency, spectral entropy, and low- and high-frequency power ratios characterize the rate and regularity of postural change [70, 126]. IMU-based kinematic features of this kind have been validated for posture recognition across multiple body placements and sensor configurations [110, 127]. To improve interpretability, we report model-specific feature importance scores with a linear-kernel SVM weight analysis for the same IMU feature set in Section 6.1. We standardize the resulting feature vectors and use them to train machine learning models for both multi-class and binary posture classification.

EEG Feature Extraction. We focus on EEG band-power ratios that serve as established cognitive workload (CWL) biomarkers, including θ/β , θ/α , and α/β to capture changes in mental effort and attentional engagement [96]. Accordingly, we focus on the θ , α , and β frequency bands. We do not extract features from the δ band or higher γ frequencies, as these ranges are more susceptible to motion, muscle, and environmental noise in wearable EEG recordings. We extract absolute band power because it captures the strength of oscillatory activity within each band [38, 99]. We also

Table 1. Hyperparameter search spaces. The same spaces apply to IMU posture, task-induced cognitive load, and unimodal discomfort classifiers.

Model	Search Space
RF	$n_{\text{est}} \in \{100, 200, 300, 500\}$, $\text{depth} \in \{5, 8, 10, \text{None}\}$, $\text{min_leaf} \in \{1, 3, 5, 8\}$
GB	$n_{\text{est}} \in \{100, 200, 300\}$, $\text{lr} \in \{0.01, 0.05, 0.1, 0.2\}$, $\text{depth} \in \{2, 3, 5\}$
SVM	$C \in \{0.1, 1, 10\}$, $\gamma \in \{\text{scale}, \text{auto}, 0.01\}$
Linear SVM	$C \in \{0.1, 1, 10\}$

compute relative band power to normalize each band by total spectral power and reduce sensitivity to subject-specific amplitude differences [118]. Inter-band power ratios and the engagement index are included because they more directly reflect workload-related shifts between slower and faster rhythms than band powers alone [86, 119]. We further compute magnitude-squared coherence for all channel pairs averaged within each band to quantify cross-channel functional coupling [68, 97]. Finally, we estimate periodic alpha power by fitting a spectral parameterization model over 1 to 30 Hz and integrating the periodic component after removing the aperiodic background. This isolates oscillatory alpha activity from broadband spectral slope effects and improves physiological specificity [115]. To improve feature interpretability, Section 6.2 reports model-specific feature rankings together with a linear-kernel SVM baseline on the same EEG feature set.

Multimodal Feature Extraction. We temporally align EEG and IMU features using matched 5 s windows with 50% overlap and retain only windows that pass IMU-guided artifact rejection (described in section 4.2.3). For each trial, we concatenate the modality-specific feature vectors into a single multimodal representation. We then aggregate window-level features to train machine learning models for our downstream classification tasks.

4.4 ML Inference

We select classifier settings to balance predictive performance and computational efficiency for lightweight wearable inference. To validate these choices, we perform a post hoc nested LOSO grid search over the spaces in Table 1, with best configurations reported in Table 2. For SVM, we search $C \in \{0.1, 1, 10\}$ and $\gamma \in \{\text{scale}, \text{auto}, .01\}$ to cover standard regularization and kernel ranges [53]. Results show that selected configurations remain competitive for binary IMU posture and task-induced cognitive load tasks, with small F1 differences. Larger gaps appear in multiclass settings, indicating higher sensitivity to hyperparameters. For example, in IMU binary classification, grid-selected SVM ($C=10$, $\gamma=.01$) achieves $F1=.82$, compared to $F1=.78$ for our selected setting ($C=1.0$, $\gamma=\text{scale}$). We apply the same validation to unimodal discomfort classification. For multimodal discomfort classification, hyperparameters are selected via nested cross-validation with inner grid search due to the higher-dimensional feature space and limited data (Table 3). For class imbalance, we apply SMOTE ($k=5$) only on training folds, with class weighting used when the minority class has fewer than two samples. Test partitions are never oversampled to prevent leakage.

4.4.1 IMU-Based Posture Classification. To support on-device inference, we use four traditional machine learning models for IMU-based posture classification: Radial basis function Support Vector Machine (SVM), Linear SVM, Random Forest (RF), and Gradient Boosting (GB). These models provide a good tradeoff between accuracy, interpretability, and computational efficiency for real-time use. For the main reported experiments, model parameters were chosen to balance predictive performance and computational efficiency for on-device deployment. Linear SVM uses a linear

Table 2. Best hyperparameter configurations and LOSO performance identified by nested grid search. Search spaces are given in Table 1. All metrics are macro-averaged.

Task	Sub-task	Model	Best configuration	Acc	F1
IMU posture classification	Binary	RF	$n_{\text{est}}=200$, depth=5, min_leaf=1	.796	.786
		GB	$n_{\text{est}}=200$, lr=0.2, depth=3	.783	.762
		SVM	$C=10$, $\gamma=0.01$	.822	.818
		Linear SVM	$C=0.1$	.796	.793
	Multiclass	RF	$n_{\text{est}}=200$, depth=8, min_leaf=5	.658	.643
		GB	$n_{\text{est}}=300$, lr=0.01, depth=5	.658	.649
		SVM	$C=10$, $\gamma=\text{scale}$	.612	.598
		Linear SVM	$C=1$	.658	.641
EEG-based task-induced cognitive load classification	Binary	RF	$n_{\text{est}}=200$, depth=5, min_leaf=3	.902	.902
		GB	$n_{\text{est}}=300$, lr=0.2, depth=3	.916	.916
		SVM	$C=1$, $\gamma=0.01$	.832	.832
		Linear SVM	$C=10$	.867	.867
	Multiclass	RF	$n_{\text{est}}=100$, depth=8, min_leaf=8	.685	.615
		GB	$n_{\text{est}}=100$, lr=0.05, depth=5	.643	.562
		SVM	$C=10$, $\gamma=\text{scale}$	.636	.548
		Linear SVM	$C=0.1$	.643	.603
Unimodal discomfort (IMU only)	Binary	RF	$n_{\text{est}}=100$, depth=8, min_leaf=5	.580	.577
		GB	$n_{\text{est}}=100$, lr=0.1, depth=2	.565	.554
		SVM	$C=0.1$, $\gamma=\text{scale}$	.623	.622
		Linear SVM	$C=0.1$	.551	.545
Unimodal discomfort (EEG only)	Binary	RF	$n_{\text{est}}=100$, depth=8, min_leaf=5	.522	.511
		GB	$n_{\text{est}}=200$, lr=0.1, depth=3	.580	.579
		SVM	$C=0.1$, $\gamma=0.01$	.565	.549
		Linear SVM	$C=0.1$	.522	.515

[†] Multimodal discomfort uses prospective nested cross-validation; search space in Table 3.

kernel with $C=1.0$ and balanced class weighting; $C=1.0$ provides moderate regularization that is well-suited to normalised inertial features without requiring heavy tuning. SVM uses $C=1.0$ and $\gamma=\text{scale}$, where scale automatically adapts the kernel width to the feature variance, avoiding manual bandwidth selection. RF uses $n_{\text{est}}=300$, max_depth=10, and min_leaf=5 as the moderate depth and leaf size prevent overfitting on the limited per-subject windows while retaining sufficient model capacity. GB uses $n_{\text{est}}=200$, lr=0.1, and max_depth=3; shallow trees with a moderate learning rate are standard for structured sensor overfitting issue [40]. We evaluate these classifiers on a multiclass task with labels $D1$, $D2$, and $D3$, and on a binary task separating neutral posture ($D1$) from combined slouched postures ($D2+D3$). We use two validation strategies: leave-one-subject-out cross-validation and a within-subject 80/20 split. In each fold, StandardScaler is fit only on the training partition to prevent data leakage.

4.4.2 EEG-Based Task-induced Cognitive Load Classification. We use the same four machine learning models for EEG-based activity recognition, applied to the EEG features described in Section 4.3. For the main reported EEG experiments, RF uses $n_{\text{est}}=300$ with balanced class weighting in the binary setting: 300 trees provide stable importance estimates for the high-dimensional EEG feature set. GB uses 300 boosting iterations with lr=0.1 and max_depth=3 because EEG spectral features are noisy and high-dimensional as deeper ensembles risk overfitting across subjects [40]. SVM uses $C=1.0$ and $\gamma=\text{scale}$ with balanced class weighting to capture non-linear decision boundaries. Linear SVM uses a linear kernel with $C=1.0$ and balanced class weighting as

it serves as an interpretable baseline whose feature importance can be examined through held-out permutation importance. We evaluate on both binary (Rest vs. Active) and multiclass (Rest, Stroop, Numerical) tasks using an 80/20 stratified split and LOSO cross-validation. For the binary task, we apply SMOTE with $k=5$ nearest neighbours only to the training partition when the minority class has at least two samples. Class weighting is used instead when the minority class is too small to support synthetic generation, and test partitions are never oversampled. For the multiclass task, we do not apply SMOTE or class weighting. Normalization follows the same training-only procedure as IMU classification.

4.4.3 Perceived Discomfort Classification. We perform unimodal and multimodal discomfort classification on the labelled dataset (Section 5) using 80/20 stratified split and LOSO evaluation. We select the top $K=10$ features per training partition using mutual information, reducing K when features or trials are limited. Features are standardized with StandardScaler fit on training data only to prevent leakage. The RF classifier uses 100 decision trees with square-root feature sampling at each split, a minimum leaf size of 8, and balanced class weighting; the larger leaf size regularizes the model given the limited labelled trials (57). The GB classifier uses 100 boosting iterations with $lr=0.1$ and $max_depth=3$ as shallow trees with a moderate learning rate consistently prevent overfitting [40]. The SVM uses $C=1.0$, $\gamma=scale$, and balanced class weighting. The Linear SVM uses a linear kernel with $C=1.0$ and balanced class weighting; $C=1.0$ provides standard regularization appropriate for the small normalised feature set. To address class imbalance, we apply random undersampling on the training partition to balance class counts; test partitions are never resampled. For multimodal discomfort classification, we fuse EEG and IMU features (Section 4.3) into a single trial-level representation. Before training, we apply median imputation to fill missing values, remove near-zero-variance features (threshold=.01), and drop highly correlated features ($r>.9$). We then select the top $K=10$ features and standardise using training data only. RF, SVM, and GB hyperparameters are selected via nested cross-validation using a 3-fold stratified inner loop and Linear SVM uses fixed $C=1.0$ as the regularization strength is stable across the small labelled set. The search space for RF, SVM, and GB is reported in Table 3.

Overall for all evaluations, we use standard implementations from scikit-learn and imbalanced-learn. Unless explicitly stated otherwise, all unspecified hyperparameters follow library defaults.

Table 3. Hyperparameter search space for multimodal discomfort classification.

Model	Hyperparameter search space
RF	$n_{estimators} \in \{100, 500\}$; $max_depth \in \{8, 12, None\}$; $min_samples_leaf \in \{2, 3\}$; $min_samples_split \in \{4, 6\}$; $max_features = sqrt$; $class_weight = balanced$
SVM	Kernel = RBF; $C \in \{.1, 1, 10\}$; $\gamma \in \{scale, auto, .001, .01\}$; $class_weight = balanced$
GB	$n_{estimators} \in \{100, 300\}$; learning rate $\in \{.01, .1, .2\}$; $max_depth \in \{3, 5, 7\}$; $min_samples_leaf \in \{2, 3\}$; subsample $\in \{.8, 1.0\}$; balanced sample weights

4.5 Real-Time Alert Scheduling Algorithm

Model Selection. We deploy Random Forest as the single classifier in our alert policies for both posture and task-induced cognitive load inference, based on its stable performance in both 80–20 and LOSO evaluations reported in Section 6. We use SVM, GB, and Linear SVM only as comparative baselines to contextualize RF performance. We do not use ensemble combinations at inference time, as combining multiple models introduces computational overhead incompatible with the latency and memory constraints of continuous on-device wearable deployment [3, 23]. We implement

two real-time alert intervention policies. The posture-only policy (P-only) triggers alerts solely based on sustained non-neutral head posture. The cognition-aware policy (P+C) applies the same posture correction logic but modulates alert timing using an estimated task-induced cognitive load, delaying or suppressing alerts during periods of high mental engagement and releasing them during low-focus breakpoints.

Posture-Only Alert (P-only): In this alert system, an alert is generated when the user's head posture deviates by more than 10° from their personalized neutral posture (described in section 4.2.1), and remains in same posture for at least 3 seconds [24]. The posture-only alert is summarized in Algorithm 1.

Cognition-Aware Alert Policy (P+C): We trigger alerts based on the user's real-time task-induced cognitive load rather than posture alone. At each inference step, posture is classified using $f_{IMU}(X_{imu})$, where $isBad$ denotes the current prediction. $badCount$ tracks consecutive bad-posture ticks to determine $sustainedBad$, and $badDur$ tracks continuous duration for the 30-second override. Both reset when posture returns to neutral. If bad posture sustains beyond 30 seconds, the system overrides the cognitive load estimate and triggers an alert regardless of the interruptibility state. The cognition-aware alert policy is summarized in Algorithm 2.

5 Experimental Setup and Data Collection

In this experiment, we conduct a user study to determine the system design and evaluate the feasibility of low cost wearable sensors for intelligent intervention and discomfort analysis. This study was approved by our Institutional Review Board. The details of the main experiment are summarized below.

Participants. We recruit 24 participants aged 18 years or above for this IRB-approved study (mean age = 29.46 years, SD = 4.91). All participants report daily use of electronic devices and complete the study after providing informed consent. Each session lasts approximately 20 to 25 minutes, and participants receive a \$10 Amazon gift card as compensation. Detailed participant demographics are provided in Appendix A.

Data Collection Procedure. The user study was conducted in a controlled laboratory setting. After providing written consent, participants completed a brief questionnaire regarding their typical electronic device usage, posture habits, and experiences with muscle or eye fatigue during prolonged device use. The physical setup included three posture conditions defined by screen distance: D1 (70 cm, neutral), D2 (50 cm, mild forward flexion), and D3 (30 cm, deep forward flexion), as shown in Figure 6. We employed a mixed design in which each participant experienced all three posture conditions. The order of these conditions was randomized to reduce order effects and bias. At each distance, every participant perform:

- (1) **Rest Activity:** Participants sit quietly with eyes closed for approximately 1 minute, establishing a baseline for neurophysiological measurements.
- (2) **Active Activity:** Participants either complete a Stroop task (with both congruent and incongruent stimuli) or perform numerical calculation (addition, subtraction, multiplication, division) in a dynamic setting. We select the Stroop task and numerical calculation as established paradigms to induce task-induced cognitive load [85, 123]. The Stroop task captures selective attention under interference [31, 62], and numerical calculation engages sustained mental effort [22, 104]. Together, they provide complementary cognitive load patterns. Participants are randomly assigned to one task to control for task-specific bias.

We use a three-minute active task window to elicit measurable task-induced cognitive load within a controlled protocol. Prior work shows that EEG-based task-induced cognitive load estimation is

Algorithm 1: Posture-only Alert Scheduling (P-only)

Input: Sampling rate $f_s = 50$ Hz; baseline duration T_b ; alert threshold θ ; persistence window T_p ; alert interval T_i

Output: Set of alert events $\mathcal{E}$

Stage 1: Personalized Posture Calibration

Collect pitch samples for T_b seconds while the user maintains neutral posture;

$P_{\text{base}} \leftarrow \text{median}(\text{baseline pitch samples});$

$N_p \leftarrow \lceil T_p \cdot f_s \rceil;$

$\mathcal{E} \leftarrow \emptyset, \text{lastAlert} \leftarrow -\infty, \text{badCount} \leftarrow 0;$

Stage 2: Real-time Monitoring

while new pitch sample (t, pitch) is available **do**

$d \leftarrow \text{pitch} - P_{\text{base}};$

$a \leftarrow |d|;$

if $a > \theta$ **then**

$\text{badCount} \leftarrow \text{badCount} + 1;$

else

$\text{badCount} \leftarrow 0;$

$\text{sustainedBad} \leftarrow (\text{badCount} \geq N_p);$

$\text{inInterval} \leftarrow (t - \text{lastAlert} < T_i);$

if $\text{sustainedBad} \wedge \neg \text{inInterval}$ **then**

 trigger alert with deviation d ;

$\mathcal{E} \leftarrow \mathcal{E} \cup \{t\};$

$\text{lastAlert} \leftarrow t;$

$\text{badCount} \leftarrow 0;$

return $\mathcal{E}$

Algorithm 2: Cognition-aware Alert Scheduling with IMU + EEG (P+C)

Input: IMU stream sampling rate $f_s = 50$ Hz; baseline duration T_{base} ; IMU window length T_{imu} ; posture persistence window T_p ; continuous bad-posture window $T_{\text{cont}} = 30$ s; alert interval T_i ; EEG window length T_e ; pretrained IMU posture classifier $f_{\text{IMU}}(\cdot)$; EEG classifier $f_{\text{EEG}}(\cdot)$

Output: Set of alert events $\mathcal{E}$

Initialization: $\Delta t \leftarrow 1/f_s, N_p \leftarrow \lceil T_p \cdot f_s \rceil, \text{badCount} \leftarrow 0, \text{badDur} \leftarrow 0, \text{lastAlert} \leftarrow -\infty, \text{eegState} \leftarrow \text{Unknown}, \text{lastEEG} \leftarrow -\infty, \mathcal{E} \leftarrow \emptyset, \text{baselineReady} \leftarrow \text{false}, t_0 \leftarrow \text{session start time};$

Real-time Monitoring: **while** session is running **do**

 read next IMU sample at time t ;

if $\neg \text{baselineReady}$ **then**

 accumulate IMU samples for baseline;

if $t - t_0 \geq T_{\text{base}}$ **then**

 compute baseline normalization

 parameters;

$\text{baselineReady} \leftarrow \text{true};$

continue;

if not enough IMU samples for a T_{imu} window **then**

continue;

$X_{\text{imu}} \leftarrow \text{most recent IMU window of length } T_{\text{imu}};$

$X_{\text{imu}} \leftarrow \text{baseline-normalize}(X_{\text{imu}});$

$\hat{y}_{\text{imu}} \leftarrow f_{\text{IMU}}(X_{\text{imu}});$

$\text{isBad} \leftarrow (\hat{y}_{\text{imu}} = \text{Bad});$

if isBad **then**

$\text{badCount} \leftarrow \text{badCount} + 1;$

$\text{badDur} \leftarrow \text{badDur} + \Delta t;$

else

$\text{badCount} \leftarrow 0;$

$\text{badDur} \leftarrow 0;$

$\text{sustainedBad} \leftarrow (\text{badCount} \geq N_p);$

$\text{continuousBad} \leftarrow (\text{badDur} \geq T_{\text{cont}});$

if $t - \text{lastEEG} \geq T_e$ **then**

$X_e \leftarrow \text{most recent EEG window of length } T_e;$

$\text{eegState} \leftarrow f_{\text{EEG}}(X_e);$

$\text{lastEEG} \leftarrow t;$

$\text{interruptible} \leftarrow (\text{eegState} = \text{Rest});$

$\text{inInterval} \leftarrow (t - \text{lastAlert} < T_i);$

$\text{trigger} \leftarrow \text{false};$

if $\neg \text{inInterval}$ **then**

if continuousBad **then**

$\text{trigger} \leftarrow \text{true};$

else if $\text{sustainedBad} \wedge \text{interruptible}$ **then**

$\text{trigger} \leftarrow \text{true};$

if trigger **then**

 trigger alert at time t ;

$\mathcal{E} \leftarrow \mathcal{E} \cup \{t\};$

$\text{lastAlert} \leftarrow t;$

$\text{badCount} \leftarrow 0;$

$\text{badDur} \leftarrow 0;$

return $\mathcal{E}$

feasible from short windows ranging from 10–20 seconds to one minute [85, 123], and that sustained task engagement induces measurable changes in seated posture and neck muscle co-activation within comparable timeframes [55]. Both tasks (Stroop and numerical calculation) are established

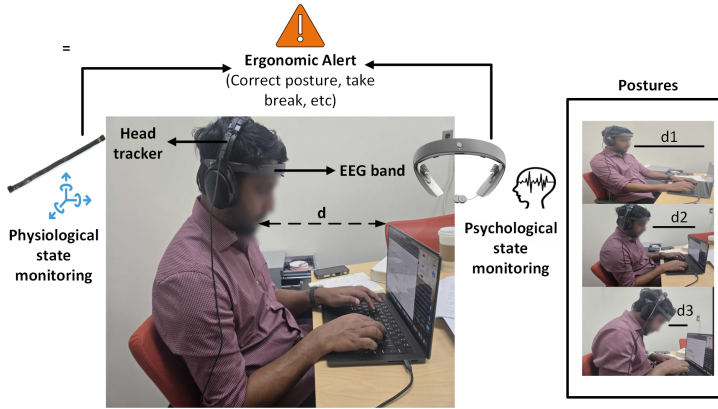


Fig. 6. Data collection setup. Participant performing activities at three screen distance conditions during *ErgoAssist* data collection.

proxies for mental effort at comparable difficulty levels [43, 85], and task order was counterbalanced across participants to control for ordering effects. Throughout the experiment, each participant wore both the Supperware head tracker [98] and the FRENZ Brainband [39] for continuous data collection.

Instructions: We do not enforce fixed angular targets for forward flexion due to inter-participant variability in cervical anatomy and natural head posture [21, 105]. Instead, we define posture conditions using controlled viewing distances D_1 , D_2 , and D_3 . We define neutral posture as a natural upright sitting position at the farthest viewing distance, $D_1 = 70$ cm. Viewing distance is measured using a measuring tape from the laptop screen to the participant's neck position. Participants are instructed to sit upright with relaxed shoulders and look at the screen comfortably, without intentional forward leaning. We do not impose a fixed angular definition of neutral posture, as natural head alignment varies across individuals [21, 105]. Instead, we illustrate representative posture angles in Figure 1b to contextualize the degree of neck flexion. At reduced viewing distances D_2 and D_3 , participants naturally adopt increasing forward flexion driven by proximity to the screen, without explicit instructions to match a specific angle. This design maintains ecological validity while ensuring consistent distance across conditions. We validate posture differences using IMU-derived pitch measurements with a clear separation across conditions (Section 4.2.1, Figure 3).

Perceived Discomfort Level Questionnaire. To assess subjective task-induced cognitive load and musculoskeletal strain, we administer the NASA Task Load Index (NASA-TLX) [47] and the Cornell Musculoskeletal Discomfort Questionnaire (CMDQ) [49] after each posture condition. Both instruments are rated on a 7-point Likert scale and provide complementary insights into perceived mental demand and physical discomfort. The full questionnaire items are provided in Appendix B.

6 ErgoAssist Performance Evaluation

In this section, we present the effectiveness of the *ErgoAssist* for posture, task-induced cognitive load, and discomfort prediction.

6.1 IMU-based Posture Classification

For binary posture classification, we group mild and severe neck flexion into a single slouched class ($D_2 + D_3$) against neutral posture (D_1), achieving consistent performance across all models.

Table 4. IMU-based posture classification performance for binary ($D1$ vs. $D2 + D3$) and multiclass ($D1/D2/D3$) tasks. Metrics include accuracy (Acc), macro-averaged precision, recall, and F1 under 80/20 and leave-one-subject-out (LOSO) evaluation.

Task	Model	80/20				LOSO validation			
		Acc	Prec.	Rec.	F1	Acc	Prec.	Rec.	F1
Binary	RF	.88	.88	.87	.87	.81	.80	.83	.80
	SVM	.86	.85	.84	.85	.79	.79	.82	.78
	GB	.84	.83	.83	.83	.80	.78	.80	.78
	LinearSVM	.89	.88	.88	.88	.78	.78	.81	.77
Multiclass	RF	.74	.73	.73	.73	.61	.60	.61	.60
	SVM	.72	.72	.71	.71	.61	.60	.61	.59
	GB	.72	.71	.71	.70	.57	.55	.56	.55
	LinearSVM	.77	.75	.75	.75	.63	.63	.63	.62

Under 80/20 split, accuracy ranges from .84 to .88, and under LOSO evaluation from .79 to .81, with RF performing best in both settings. A linear kernel SVM evaluated on the same standardized IMU feature set achieves .89 under 80/20 split and .78 under LOSO evaluation, confirming that binary posture separation is largely captured by a linear decision boundary.

For multiclass classification ($D1$ vs. $D2$ vs. $D3$), models reach .72 to .74 accuracy under 80/20 split, with $D1$ the most reliably identified class and $D2$ consistently showing the lowest scores. Under LOSO evaluation, accuracy drops to .57 to .61, indicating stronger subject-specific variability in neck flexion dynamics. The linear kernel SVM achieves .77 under 80/20 split and .63 under LOSO evaluation. Across all models, the dominant features are consistently pitch RMS, mean, and median, together with derivative-based descriptors such as acceleration and velocity variability, indicating that IMU posture classification is driven primarily by sustained flexion magnitude rather than model-specific artifacts. Overall performance is summarized in Table 4, and feature rankings are shown in Table 5.

Key Takeaway: Head-mounted IMU sensing enables reliable detection of FHP with high accuracy after personalized baseline normalization. Binary posture classification (neutral vs. slouched) is robust across users, achieving up to 88% accuracy (RF) and maintaining strong LOSO performance.

6.2 EEG-based Task-induced Cognitive Load Classification

For the binary task, RF and GB achieve the highest LOSO accuracies of .902 and .895 respectively, indicating that Rest versus Active discrimination is robust across subjects and that a substantial part of this separation is captured by a linear decision boundary. The multiclass task is more challenging as the model must separate two active conditions (*Stroop* and *Numerical*) in addition to *Rest*. Under 80/20 split, RF achieves the highest accuracy of .828, and under LOSO evaluation performance drops for all models with RF reaching .664. This generalization gap indicates that models reliably separate Rest from Active but distinguishing the two active tasks is more sensitive to subject-specific variation. Overall results are summarized in Table 6.

The feature importance analysis is summarized in Table 7. Across models, the most informative EEG features are consistently θ/β , TBR, relative theta, and relative alpha power. These patterns remain stable under LOSO evaluation, and for the multiclass task feature importance becomes more distributed but the same ratios remain dominant. The linear SVM additionally highlights periodic

Table 5. Top five IMU features for posture classification. Values denote model-specific feature importance scores. For RF, GB, and XGB, these scores are derived from tree-based importance measures. For LinearSVM, they are derived from normalized absolute coefficient magnitudes.

Task	Eval	Model	Top features (importance)
Binary	80/20	RF	rms (.19), median (.18), mean (.18), acc_std (.07), vel_max (.05)
		GB	rms (.39), median (.29), mean (.13), acc_std (.09), acc_max (.04)
		XGB	mean (.20), median (.17), vel_max (.12), acc_std (.11), rms (.09)
		LinearSVM	rms (.89), mean (.80), median (.76), acc_std (.46), vel_max (.29)
Binary	LOSO	RF	median (.28), rms (.26), mean (.26), vel_max (.02), acc_std (.02)
		GB	rms (.59), median (.13), mean (.09), vel_max (.02), acc_std (.02)
		XGB	rms (.17), median (.13), vel_mean (.10), mean (.09), ll (.04)
		LinearSVM	rms (1.00), vel_std (.27), mean (.13), acc_std (.13), range (.11)
Multiclass	80/20	RF	median (.20), mean (.19), rms (.19), vel_max (.06), acc_std (.06)
		GB	rms (.33), median (.27), mean (.14), acc_std (.09), vel_max (.08)
		XGB	mean (.22), median (.16), vel_max (.11), rms (.10), acc_std (.08)
		LinearSVM	rms (.92), mean (.89), median (.87), acc_std (.44), vel_max (.32)
Multiclass	LOSO	RF	rms (.22), median (.21), mean (.21), vel_max (.05), acc_std (.03)
		GB	rms (.40), mean (.16), median (.16), vel_max (.06), acc_std (.04)
		XGB	rms (.12), median (.11), mean (.10), vel_mean (.07), vel_max (.05)
		LinearSVM	rms (1.00), vel_std (.29), mean (.21), acc_std (.20), std (.13)

Table 6. EEG-based task-induced cognitive load classification performance for binary and multiclass tasks. Metrics include accuracy (Acc), macro-averaged precision, recall, and F1 under 80/20 and leave-one-subject-out (LOSO) evaluation.

Task	Model	80/20				LOSO validation			
		Acc	Prec.	Rec.	F1	Acc	Prec.	Rec.	F1
Binary	RF	.759	.760	.757	.757	.902	.902	.902	.902
	SVM	.793	.799	.790	.791	.818	.818	.818	.818
	GB	.828	.829	.829	.828	.895	.895	.895	.895
	Linear SVM	.793	.799	.790	.791	.804	.804	.804	.804
Multiclass	RF	.828	.800	.787	.779	.664	.564	.569	.564
	SVM	.690	.612	.622	.613	.587	.471	.477	.460
	GB	.793	.750	.765	.754	.636	.542	.546	.541
	Linear SVM	.724	.711	.746	.714	.601	.571	.573	.567

alpha features, suggesting that both spectral balance and oscillatory alpha structure contribute to task-induced cognitive load discrimination.

Window overlap sensitivity. We conduct a window-overlap ablation by varying feature-window overlap from 0% to 50% while keeping the Random Forest pipeline fixed. Table 8 reports macro F1 under both stratified 80/20 and LOSO evaluation. Under LOSO, binary macro F1 ranges from .88 to .91 and multiclass macro F1 ranges from .56 to .61, indicating broadly stable performance

Table 7. Top five EEG features for task-induced cognitive load classification under 80/20 and LOSO evaluation. For RF, GB, and XGB, values denote model-based feature importance scores. For Linear SVM, values denote normalized absolute coefficient-based importance scores.

Task	Eval	Model	Top features (importance)
Binary	80/20	RF	θ/β _Ch4 (.09), TBR_Ch4 (.08), TBR_mean (.07), θ_{rel} _Ch4 (.06), β_{rel} _Ch4 (.05)
		GB	θ_{rel} _Ch4 (.47), BandAlpha_Ch1 (.07), TBR_Ch4 (.06), θ/β _Ch4 (.05), AlphaPower_Ch1 (.05)
		XGB	BandAlpha_Ch3 (.09), α_{rel} _Ch2 (.07), TBR_mean (.07), θ/β _Ch4 (.07), θ_{rel} _Ch2 (.06)
		Linear SVM	EngagementIndex_std (1.00), α_{rel} _Ch1 (.99), θ/α _Ch3 (.69), α_{rel} _Ch3 (.61), θ_{rel} _Ch4 (.60)
Binary	LOSO	RF	θ/β _Ch4 (.07), TBR_Ch4 (.06), TBR_mean (.06), α_{rel} _Ch3 (.05), θ_{rel} _Ch4 (.05)
		GB	θ_{rel} _Ch4 (.21), α_{rel} _Ch3 (.21), TBR_Ch4 (.15), θ/β _Ch4 (.09), TBR_mean (.04)
		XGB	TBR_mean (.10), α_{rel} _Ch3 (.06), θ/β _Ch4 (.06), TBR_Ch4 (.05), θ_{rel} _Ch4 (.04)
		Linear SVM	α_{rel} _Ch1 (.95), θ/α _Ch3 (.85), α_{rel} _Ch3 (.84), EngagementIndex_std (.82), α_{rel} _Ch4 (.71)
Multiclass	80/20	RF	θ/β _Ch4 (.05), TBR_Ch4 (.04), TBR_mean (.03), α_{rel} _Ch3 (.03), θ_{rel} _Ch4 (.03)
		GB	θ/β _Ch4 (.18), α_{rel} _Ch3 (.10), θ_{rel} _Ch4 (.06), TBR_Ch4 (.06), α/β _Ch3 (.05)
		XGB	BandBeta_Ch2 (.11), θ/β _Ch4 (.05), TBR_Ch4 (.05), TBR_Ch3 (.04), θ_{rel} _Ch4 (.04)
		Linear SVM	α_{rel} _Ch1 (1.00), α_{rel} _Ch3 (.93), Coh_C1C2_Alpha (.88), α/β _Ch3 (.74), α_{rel} _Ch4 (.71)
Multiclass	LOSO	RF	θ/β _Ch4 (.04), TBR_mean (.04), TBR_Ch4 (.04), α_{rel} _Ch3 (.03), θ_{rel} _Ch4 (.03)
		GB	α_{rel} _Ch3 (.11), θ_{rel} _Ch4 (.10), θ/β _Ch4 (.09), θ/α _Ch4 (.05), TBR_Ch4 (.05)
		XGB	TBR_mean (.05), θ_{rel} _Ch4 (.05), TBR_Ch3 (.04), α_{rel} _Ch3 (.03), TBR_Ch4 (.03)
		Linear SVM	α_{rel} _Ch1 (1.00), α_{rel} _Ch3 (.78), α/β _Ch3 (.70), θ_{rel} _Ch3 (.65), α_{rel} _Ch4 (.65)

Table 8. Effect of feature-window overlap on EEG classification (macro F1, Random Forest). We report both stratified 80/20 and LOSO evaluation. LOSO is the primary cross-subject estimate.

Window Overlap	Binary 80/20	Binary LOSO	Multiclass 80/20	Multiclass LOSO
0%	.828	.881	.834	.561
10%	.793	.888	.694	.561
20%	.828	.909	.782	.585
30%	.757	.902	.726	.607
40%	.757	.909	.694	.600
50% [†]	.757	.902	.779	.564

[†] Setting used in our main experiments.

across the full overlap range. We retain 50% overlap as the default because it produces more feature windows per recording, improving feature stability by reducing sensitivity to any single window. Features from all valid windows are averaged into a single fixed-length representation per recording before any train-test split, so overlapping windows are never treated as separate training examples and the overlap does not introduce train-test information leakage [13].

Key Takeaway: Wearable EEG can reliably distinguish task-induced cognitive load such as Rest vs. Active, achieving 90% LOSO accuracy, even with lightweight models and consumer-grade hardware.

6.3 Feature Group Ablation Analysis

We conduct a group-wise feature ablation study to assess the contribution of each IMU and EEG feature group to posture and cognitive load classification respectively.

Table 9. IMU Feature Group Ablation Study using Random Forest. All metrics are macro-averaged. Top features are ranked by RF importance score (in parentheses)

Task	Feature Group	80/20		LOSO		Top-5 Features (80/20)	Top-5 Features (LOSO)
		Acc	F1	Acc	F1		
Binary	Time-domain	.87	.87	.81	.80	rms (.25), median (.25), mean (.22), ll (.07), mav (.07)	median (.30), mean (.27), rms (.27), std (.03), ll (.03)
	Derivative	.67	.64	.52	.50	acc_std (.25), vel_max (.20), vel_std (.16), acc_mean (.15), acc_max (.12)	acc_std (.22), vel_mean (.21), vel_std (.18), acc_mean (.17), vel_max (.13)
	Hjorth	.60	.57	.56	.51	hj_activity (.51), hj_mobility (.24), hj_complexity (.24)	hj_complexity (.35), hj_activity (.34), hj_mobility (.31)
	Frequency-domain	.58	.55	.55	.51	high_freq_power (.39), low_freq_power (.25), spec_entropy (.11), high_freq_ratio (.10), low_freq_ratio (.10)	high_freq_power (.21), low_freq_power (.19), spec_entropy (.19), high_freq_ratio (.17), low_freq_ratio (.17)
	All Features	.88	.87	.81	.80	rms (.19), median (.18), mean (.18), acc_std (.07), vel_max (.05)	median (.28), rms (.26), mean (.26), vel_max (.02), acc_std (.02)
Multiclass	Time-domain	.74	.73	.61	.59	median (.27), rms (.26), mean (.23), ll (.06), mav (.06)	rms (.25), median (.25), mean (.23), std (.05), ll (.04)
	Derivative	.46	.46	.35	.35	acc_std (.24), vel_max (.21), vel_std (.16), acc_mean (.15), acc_max (.13)	acc_std (.21), vel_mean (.18), vel_std (.17), vel_max (.16), acc_mean (.16)
	Hjorth	.37	.37	.36	.36	hj_activity (.50), hj_mobility (.25), hj_complexity (.25)	hj_activity (.39), hj_complexity (.32), hj_mobility (.30)
	Frequency-domain	.37	.37	.35	.35	high_freq_power (.36), low_freq_power (.26), spec_entropy (.12), low_freq_ratio (.11), high_freq_ratio (.11)	high_freq_power (.23), low_freq_power (.20), spec_entropy (.18), high_freq_ratio (.16), low_freq_ratio (.16)
	All Features	.74	.73	.61	.60	median (.20), mean (.19), rms (.19), vel_max (.06), acc_std (.06)	rms (.22), median (.21), mean (.21), vel_max (.05), acc_std (.03)

IMU Feature Group Ablation. We conduct a group-wise ablation study using Random Forest to evaluate each feature group for posture classification. Table 9 shows that time-domain features are the dominant features. Under 80/20 evaluation, the accuracy is .87 for binary and .74 for multiclass classification. The full feature set generalizes best under LOSO with .81 binary accuracy and .61 multiclass accuracy. Other feature groups perform poorly in isolation with LOSO binary accuracies between .52 and .56. This indicates they capture complementary dynamics rather than primary discriminative signals. For multiclass LOSO and 80/20, the results show the same trend. This suggests that derivative, Hjorth, and frequency-domain features improve within-subject discrimination but not cross-subject generalization. We conclude that time-domain features are the primary signal source but combining all feature groups remains necessary for robust binary cross-subject performance.

EEG Feature Group Ablation. We perform a feature-level ablation study with a Random Forest classifier to assess each EEG feature group for task-induced cognitive load classification (Table 10). Coherence features reach perfect performance under the 80/20 split (Acc: 1.00) but drop to Acc: .93 under LOSO. This suggests that connectivity patterns are partly subject-specific and can inflate performance when subjects overlap between train and test. Spectral ratio features show the most stable performance across settings (80/20: Acc: .72; LOSO: Acc: .92), indicating strong

Table 10. EEG Feature Group Ablation Study using Random Forest. All metrics are macro-averaged. Top features are ranked by RF importance score (in parentheses).

Task	Feature Group	80/20		LOSO		Top-5 Features (80/20)	Top-5 Features (LOSO)
		Acc	F1	Acc	F1		
Binary	Spectral Ratios	.72	.72	.92	.92	Ch4 θ_{rel} (.17), Ch4 θ/β (.12), Ch4 β_{rel} (.10), Ch3 α_{rel} (.07), Ch2 θ_{rel} (.06)	Ch4 θ_{rel} (.12), Ch3 α_{rel} (.10), Ch4 θ/β (.10), Ch4 β_{rel} (.07), Ch1 α_{rel} (.06)
	Absolute Power	.69	.69	.76	.76	Ch3 θ (.15), Ch1 α (.15), Ch3 α (.12), Ch3 β (.12), Ch1 θ (.10)	Ch1 α (.15), Ch1 θ (.14), Ch3 θ (.13), Ch3 β (.11), Ch3 α (.10)
	TBR+EI	.72	.72	.84	.84	TBR_ch4 (.21), TBR_mean (.18), EI_ch4 (.15), TBR_std (.12), EI_mean (.06)	TBR_ch4 (.18), TBR_mean (.17), TBR_std (.11), EI_ch4 (.10), TBR_ch2 (.08)
	Coherence	1.00	1.00	.93	.93	C1C3 α (.12), C2C4 θ (.10), C3C4 θ (.10), C2C3 α (.09), C1C2 θ (.09)	C1C3 α (.12), C2C4 θ (.11), C1C2 α (.10), C3C4 θ (.09), C2C3 θ (.09)
	Periodic Alpha	.59	.59	.59	.59	log_PA_ch3 (.13), PA_ch3 (.13), log_PA_mean (.10), log_PA_ch1 (.10), PA_ch1 (.10)	PA_ch3 (.13), log_PA_ch3 (.13), log_PA_mean (.10), PA_ch1 (.10), log_PA_ch1 (.10)
	All Features	.76	.76	.90	.90	θ/β _Ch4 (.09), TBR_Ch4 (.08), TBR_mean (.07), θ_{rel} _Ch4 (.06), β_{rel} _Ch4 (.05)	θ/β _Ch4 (.07), TBR_Ch4 (.06), TBR_mean (.06), α_{rel} _Ch3 (.05), θ_{rel} _Ch4 (.05)
Multi-class	Spectral Ratios	.72	.66	.69	.60	Ch4 θ/β (.08), Ch4 θ_{rel} (.08), Ch3 α_{rel} (.06), Ch1 α_{rel} (.06), Ch3 θ_{rel} (.05)	Ch4 θ_{rel} (.08), Ch4 θ/β (.07), Ch1 α_{rel} (.07), Ch3 α_{rel} (.07), Ch4 β_{rel} (.05)
	Absolute Power	.76	.74	.54	.46	Ch3 θ (.14), Ch3 β (.11), Ch1 θ (.10), Ch3 α (.09), Ch1 β (.09)	Ch3 θ (.12), Ch1 α (.11), Ch1 θ (.10), Ch3 β (.10), Ch3 α (.10)
	TBR+EI	.76	.70	.63	.55	TBR_ch4 (.15), TBR_mean (.11), TBR_std (.10), EI_ch4 (.09), TBR_ch3 (.09)	TBR_ch4 (.14), TBR_mean (.12), TBR_ch3 (.10), EI_ch3 (.10), TBR_std (.09)
	Coherence	.76	.67	.69	.61	C1C2 α (.11), C1C3 α (.10), C2C4 θ (.10), C2C3 α (.08), C1C4 α (.08)	C1C3 α (.10), C2C4 θ (.09), C1C2 α (.09), C2C3 θ (.09), C2C3 α (.09)
	Periodic Alpha	.45	.33	.50	.41	log_PA_ch3 (.12), log_PA_mean (.11), PA_ch3 (.11), PA_ch1 (.10), log_PA_ch4 (.10)	PA_ch3 (.13), log_PA_ch3 (.12), log_PA_ch1 (.10), PA_ch1 (.10), log_PA_ch2 (.10)
	All Features	.83	.78	.66	.56	θ/β_Ch4 (.05), TBR_Ch4 (.04), TBR_mean (.03), α_{rel}_Ch3 (.03), θ_{rel}_Ch4 (.03)	θ/β_Ch4 (.04), TBR_mean (.04), TBR_Ch4 (.04), α_{rel}_Ch3 (.03), θ_{rel}_Ch4 (.03)

generalization of relative band dynamics. In contrast, absolute power (LOSO: Acc: .76) and periodic alpha features (LOSO: Acc: .59) provide weaker discrimination. For multiclass classification, spectral and coherence features obtain similar performance (LOSO: Acc: .69), and other groups perform lower. The combined feature set consistently performs better for all classification tasks (binary LOSO: Acc: .90; multiclass LOSO: Acc: .66), indicating complementary information across feature groups.

6.4 Correlation between posture and task-induced cognitive load

We analyze the relationship between task-induced cognitive load and posture on EEG biomarkers using linear mixed-effects models [41] in Table 11. The resting neutral posture (*D1*) serves as the baseline ($\beta = 3.209$, $SE = .090$, $p < .001$). Periodic alpha power shows a significant decrease

Table 11. Linear mixed-effects model results for periodic and relative alpha power. All coefficients are fixed-effect estimates relative to the resting, upright (D1) baseline.

Metric	Factor	Contrast	β	Std. Error	p
Periodic Alpha	Baseline	Resting & Neutral (D1)	3.209	.090	< .001*
	Cognitive Activity	Numerical vs. Rest	-0.455	.130	< .001*
	Cognitive Activity	Stroop vs. Rest	-0.084	.131	.521
	Posture	Mild flexion (D2)	-0.139	.104	.179
	Posture	Deep flexion (D3)	-0.138	.104	.184
Relative Alpha	Interaction	Stroop $\times$ Mild flexion (D2)	-0.031	.013	.018*
	Interaction	Stroop $\times$ Deep flexion (D3)	-0.020	.013	.122

* $p < .05$. Negative coefficients indicate reduced alpha power.

during the numerical task ($\beta = -0.455$, $p < .001$), and relative alpha power shows a significant reduction during the Stroop task under mild flexion (D2) ($\beta = -.031$, $p = .018$). No other significant interactions between task-induced cognitive load and posture are observed. As shown in Figure 7, slouched postures exhibit reduced alpha band power relative to neutral, indicating lower relaxation levels [102], and increased beta and theta activity in D2 and D3 suggests elevated cognitive effort and mental strain [69, 96].

Key Takeaway: *ErgoAssist's* EEG biomarkers are driven mainly by task-induced cognitive load but bad posture also plays a secondary role to increase cognitive strain in high demanding task.

6.5 Perceived Discomfort Classification

We measure perceived cognitive workload using NASA-TLX and physical discomfort using CMDQ on 7-point scales, both showing good internal consistency (TLX: $\alpha = .76$; CMDQ: $\alpha = .89$). TLX scores indicate moderate workload ($M = 3.67$, $SD = 1.01$) and CMDQ scores reflect mild to moderate physical discomfort ($M = 2.63$, $SD = 1.54$). We split TLX scores at the median (3.50) into low- and high-discomfort trials (Figure 8a). A Kruskal–Wallis test confirms that posture significantly affects TLX ($H = 23.76$, $p < .001$), with deeper neck flexion (D3) associated with higher perceived strain (Figures 8b, 8c).

Unimodal Perceived Discomfort Classification. For IMU-based models, RF achieves the highest within-subject accuracy (.79) under 80/20 split, and GB provides the most robust cross-participant accuracy (.65) under LOSO evaluation. The linear kernel SVM achieves .57 under 80/20 and .57 under LOSO evaluation. For EEG-based classification, RF achieves the strongest performance (.71 under 80/20 and .64 under LOSO evaluation), with SVM and GB performing worse. The linear kernel SVM achieves .43 under 80/20 and .55 under LOSO evaluation.

Multimodal Perceived Discomfort Classification. We combine EEG and IMU features in a multimodal setting (Section 4.3). Under 80/20 evaluation, RF and GB both achieve the strongest within-subject accuracy (.75), and under LOSO evaluation RF provides the best cross-participant accuracy (.58), followed by GB (.56) and SVM (.53). The multimodal linear kernel SVM achieves .67 under 80/20 evaluation but drops to .42 under LOSO evaluation. All results are summarized in Table 12.

Perceived Discomfort Classification Features. Feature rankings are summarized in Tables 13 and 14. For IMU-only recognition, the most influential features are pitch mean and dominant

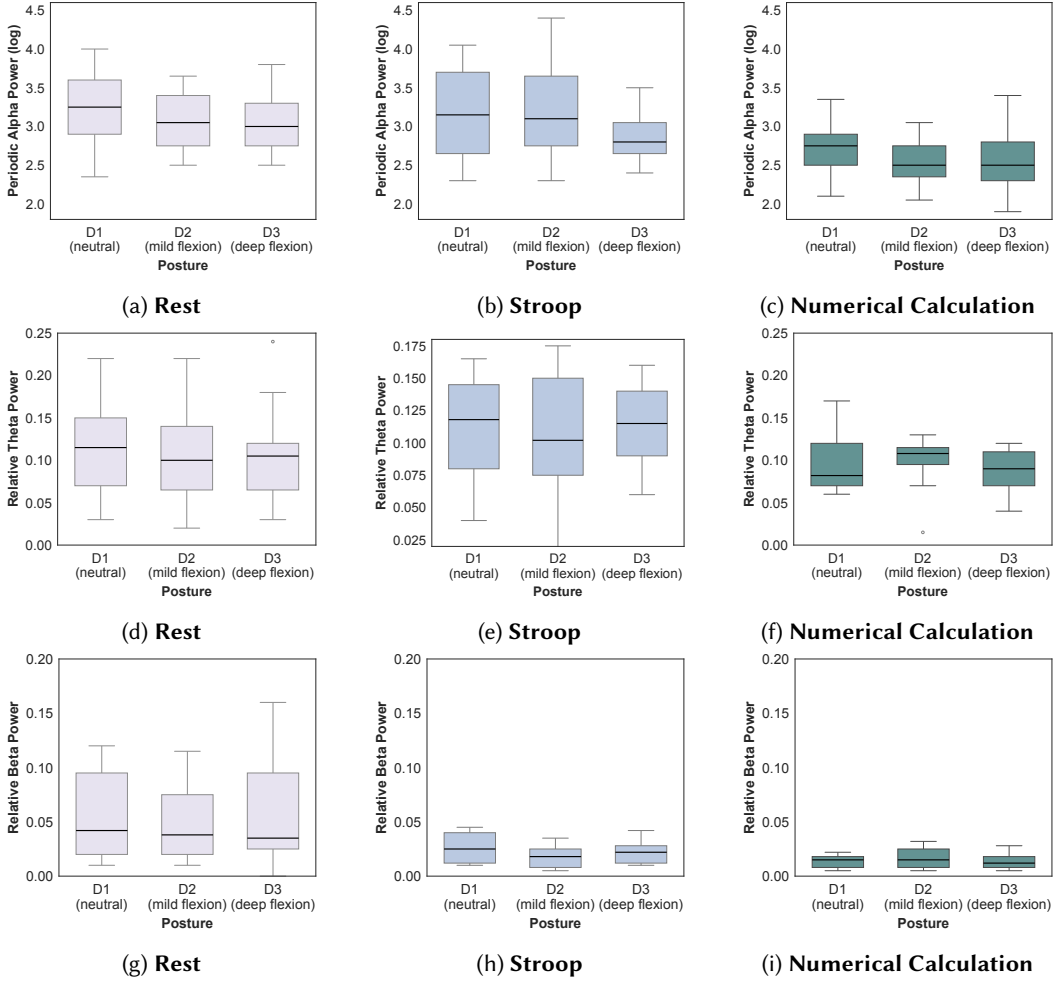


Fig. 7. EEG spectral features across diverse tasks and postures.

frequency, followed by skewness, spectral entropy, and kurtosis, suggesting that perceived discomfort is associated with sustained head flexion magnitude and postural adjustment dynamics. For EEG-only recognition, dominant features include relative beta and alpha power, theta band coherence, Hjorth complexity, and higher-order statistical descriptors. Under LOSO evaluation, the EEG feature set shifts toward waveform shape statistics, Hjorth measures, and band relative power terms. In the multimodal setting, the strongest predictors combine frontal EEG descriptors such as relative theta power, TBR, kurtosis, Hjorth mobility, and `pitch_mean_corr`, consistent across both RF and Linear SVM.

Key Takeaway: Both posture and task-induced cognitive load contribute to user-perceived discomfort, as reflected in NASA-TLX and CMDQ scores. Multimodal (IMU + EEG) models outperform unimodal approaches, achieving better accuracy and generalization across users.

All ML evaluations establish the technical validity of ErgoAssist components and alert policies (Sec. 4). Specifically, posture classification verifies the reliability of detecting non-neutral head configurations, while EEG-based inference confirms the feasibility of estimating task-induced

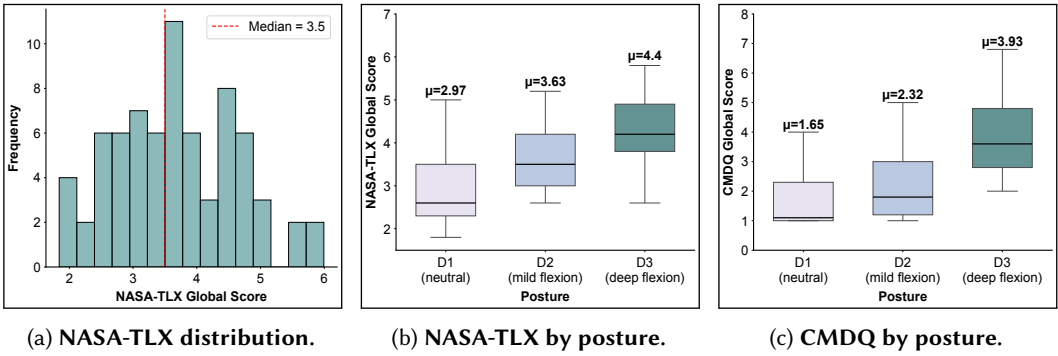


Fig. 8. Subjective workload and discomfort measures across postures. (a) Global NASA-TLX score distribution. (b) NASA-TLX global scores by posture. (c) CMDQ scores indicating increased physical discomfort during deeper neck flexion states.

Table 12. Discomfort classification performance across modalities and classifiers.

Modality	Classifier	Method	Accuracy	Precision	Recall	F1-score
IMU only	RF	80/20	.786	.857	.750	.800
	RF	LOSO	.652	.743	.634	.684
	SVM	80/20	.571	.667	.500	.571
	SVM	LOSO	.551	.647	.537	.587
	GB	80/20	.643	.800	.500	.615
	GB	LOSO	.652	.730	.659	.692
	Linear SVM	80/20	.571	.625	.604	.562
	Linear SVM	LOSO	.565	.582	.583	.565
EEG only	RF	80/20	.714	.750	.750	.750
	RF	LOSO	.638	.750	.585	.658
	SVM	80/20	.357	.429	.375	.400
	SVM	LOSO	.551	.647	.537	.587
	GB	80/20	.429	.500	.500	.500
	GB	LOSO	.536	.655	.463	.543
	Linear SVM	80/20	.429	.438	.438	.429
	Linear SVM	LOSO	.551	.552	.554	.547
Multimodal	RF	80/20	.750	.824	.824	.824
	RF	LOSO	.579	.625	.625	.625
	SVM	80/20	.583	.737	.737	.737
	SVM	LOSO	.526	.658	.658	.658
	GB	80/20	.750	.824	.824	.824
	GB	LOSO	.561	.648	.648	.648
	Linear SVM	80/20	.667	.657	.657	.657
	Linear SVM	LOSO	.421	.369	.393	.372

cognitive load. Further, the observed correlation between posture deviation and cognitive state provides empirical support for our core design assumption. Together, these results justify the integration of posture sensing and cognitive awareness in enabling context-aware alert scheduling.

Table 13. Top unimodal features for perceived discomfort classification under 80/20 and LOSO evaluation. For RF and GB, values denote model based feature importance scores. For Linear SVM, values denote normalized absolute coefficient based importance scores.

Modality	Model	Eval	Top Features (importance)
IMU only	RF	80/20	pitch_mean (.29), pitch_dom_freq (.15), pitch_skew (.14), pitch_spec_entropy (.10), pitch_hjorth_complexity (.07)
	GB	80/20	pitch_mean (.27), pitch_dom_freq (.26), pitch_std (.13), pitch_velocity_mean (.13), pitch_spec_entropy (.06)
	Linear SVM	80/20	pitch_spec_entropy (1.00), pitch_mean (.83), pitch_dom_freq (.62), pitch_hjorth_mobility (.38), pitch_hjorth_complexity (.34)
	RF	LOSO	pitch_mean (.25), pitch_dom_freq (.16), pitch_skew (.14), pitch_spec_entropy (.09), pitch_kurt (.08)
	GB	LOSO	pitch_mean (.29), pitch_dom_freq (.23), pitch_skew (.10), pitch_spec_entropy (.09), pitch_kurt (.08)
	Linear SVM	LOSO	pitch_mean (.95), pitch_dom_freq (.55), pitch_skew (.48), pitch_kurt (.43), pitch_std (.35)
EEG only	RF	80/20	ch3_hjorth_complexity_mean (.18), ch4_mean_mean (.14), ch5_beta_rel_mean (.13), ch3_beta_rel_mean (.13), ch1_ch3_coh_theta_std (.13)
	GB	80/20	ch3_beta_rel_mean (.27), ch2_hjorth_complexity_min (.16), ch3_hjorth_complexity_mean (.15), ch5_beta_rel_mean (.13), ch5_spec_entropy_max (.11)
	Linear SVM	80/20	ch3_hjorth_complexity_mean (1.00), ch5_beta_rel_mean (.96), ch5_alpha_rel_std (.93), ch5_hjorth_complexity_mean (.67), ch3_beta_rel_mean (.60)
	RF	LOSO	ch4_mean_mean (.15), ch2_hjorth_mobility_std (.13), ch1_ch3_coh_theta_std (.13), ch3_hjorth_complexity_mean (.12), ch5_kurt_mean (.12)
	GB	LOSO	ch1_ch3_coh_theta_std (.19), ch4_mean_mean (.19), ch5_kurt_mean (.16), ch4_hjorth_complexity_mean (.13), ch2_alpha_rel_std (.13)
	Linear SVM	LOSO	ch3_skew_max (.58), ch4_mean_mean (.58), ch2_beta_rel_min (.51), ch2_hjorth_mobility_std (.49), ch2_tbr_mean (.44)

6.6 Energy Consumption

We quantify on-device computational cost by measuring each model's serialized size and inference latency. RF models achieve the highest classification accuracy but incur higher computational overhead, with inference latency of 13–15 ms and memory footprints up to 0.7 MB. In contrast, SVM requires only 0.03 ms per inference and 5.5 KB of memory, approximately 400× lower latency than RF. These results indicate that lightweight classifiers are well suited for long-term on-device operation, and our prototype can support extended user monitoring with modest energy consumption (Table 15).

Key Takeaway: All models run with low latency (13–15 ms) and modest memory footprint, supporting real-time on-device deployment on resource-constrained wearable platforms.

7 ErgoAssist in the Wild: A Preliminary Real-Time Comparative Study

We conduct a preliminary exploratory study to compare two real-time alert policies: posture-only (P-only) and cognition-aware (P+C), as described in Section 4.5. We recruit thirteen participants,

Table 14. Top multimodal features for discomfort classification under 80/20 random split and LOSO evaluation. Values denote model specific feature importance measures. For RF, SVM, and GB, values denote model based feature importance scores. For Linear SVM, values denote normalized absolute coefficient based importance scores.

Setting	Model	Top Features (importance)
80/20	RF	OTE_L-FpZ_theta_rel_mean (.17), RF-FpZ_tbr_min (.14), OTE_L-FpZ_kurtosis_min (.14), OTE_R-FpZ_hjorth_complexity_max (.10), RF-FpZ_tbr_std (.08)
	SVM	OTE_L-FpZ_theta_rel_mean (.01), OTE_R-FpZ_hjorth_mobility_max (.01), pitch_mean_corr (.003), RF-FpZ_tbr_min (.00), OTE_L-FpZ_kurtosis_min (-.07)
	GB	OTE_L-FpZ_theta_rel_mean (.20), pitch_mean_corr (.18), RF-FpZ_tbr_min (.13), OTE_L-FpZ_kurtosis_min (.11), OTE_L-FpZ_tbr_std (.10)
	Linear SVM	OTE_L-FpZ_tbr_std (1.00), pitch_mean_corr (.97), OTE_L-FpZ_hjorth_mobility_max (.90), OTE_L-FpZ_theta_rel_mean (.48), RF-FpZ_tbr_std (.42)
LOSO	RF	pitch_mean_corr (.23), OTE_L-FpZ_theta_rel_mean (.21), LF-FpZ_skewness_max (.17), OTE_L-FpZ_skewness_max (.15), OTE_L-FpZ_kurtosis_mean (.12)
	SVM	OTE_R-FpZ_theta_rel_mean (.06), RF-FpZ_skewness_std (.02), OTE_L-FpZ_theta_rel_mean (.02), OTE_L-FpZ_kurtosis_mean (.01), coh_OTE_LR_alpha_mean (.01)
	GB	pitch_mean_corr (.30), OTE_L-FpZ_theta_rel_mean (.25), LF-FpZ_skewness_max (.17), RF-FpZ_skewness_std (.16), OTE_L-FpZ_abr_min (.16)
	Linear SVM	OTE_L-FpZ_kurtosis_mean (.77), OTE_L-FpZ_kurtosis_min (.71), OTE_R-FpZ_hjorth_complexity_max (.44), OTE_R-FpZ_tbr_std (.42), RF-FpZ_theta_rel_mean (.41)

including five from the earlier data collection and eight newly recruited. We explicitly evaluate system behavior on participants whose data was not used during model development (P6–P13) to assess cross-user robustness [16, 116]. Each participant completes both conditions in a randomized order to control for order effects. Each session begins with a 1-minute baseline phase for calibration, followed by a 3-minute numerical calculation task. Calibration is repeated before every session because participants may change posture or move between tasks, making a continuous baseline unreliable [29]. Participants receive visual and haptic alerts and their task scores are logged. When an alert is triggered, a corrective message and guided stretching video¹ are sent. Participants adjust their posture freely at their own discretion. At the end of each condition, participants complete a modified System Usability Scale (SUS) and an interruptibility survey, both rated on a 7-point Likert scale. The full questionnaire items are in Appendix B.

Experimental Results. From Table 16 and Figures 9a and 9b, alerts decrease from 9.6 (P-only) to 1.8 (P+C) across $N = 13$ participants, corresponding to an 81% reduction consistent across known and unseen groups. For example, alerts decrease from 18 to 2 for P6 and from 16 to 1 for P9. P+C alerts consistently exceed the pitch deviation threshold, confirming that the reduction reflects improved timing rather than relaxed criteria. P13 shows zero P+C alerts, indicating selective suppression when interruptibility conditions are not met. The mean rest probability ranges from .50 to .77, confirming that alerts are delivered during lower cognitive load periods. Consistent trends for unseen participants (P6–P13) confirm generalizability beyond participant-specific fitting.

¹<https://www.youtube.com/watch?v=Q12nlfVCpdU>

Table 15. Model footprint and runtime metrics for on-device inference. We report mean inference latency, 95th percentile latency (p95), and serialized model size for IMU, EEG, and multimodal classifiers under 80/20 and LOSO evaluation.

Modality	Classifier	Method	Latency (ms)	p95 (ms)	Size (KB)
IMU	RF	80/20	13.916	14.964	67.08
IMU	RF	LOSO	12.885	13.590	523.53
IMU	SVM	80/20	.035	.062	5.31
IMU	SVM	LOSO	.031	.035	5.89
IMU	GB	80/20	.074	.082	116.96
IMU	GB	LOSO	.074	.093	113.44
EEG	RF	80/20	13.834	14.757	66.45
EEG	RF	LOSO	13.190	13.674	516.03
EEG	SVM	80/20	.032	.032	5.11
EEG	SVM	LOSO	.039	.104	5.50
EEG	GB	80/20	.074	.105	131.87
EEG	GB	LOSO	.076	.097	128.50
Multimodal	RF	80/20	12.976	13.581	719.75
Multimodal	RF	LOSO	14.116	15.679	141.22
Multimodal	SVM	80/20	.028	.032	5.58
Multimodal	SVM	LOSO	.028	.029	5.48
Multimodal	GB	80/20	.094	.098	490.02
Multimodal	GB	LOSO	.091	.092	322.03

Table 16. Per-participant alert statistics for posture-only (P-only) and posture + cognition (P+C) policies. # **Alerts**: total alerts per session. **Mean sustain time**: average duration (s) from alert onset to user dismissal. **Mean rest probability**: estimated probability of low cognitive load at alert delivery (P+C only). P13 receives no P+C alerts because sustained bad posture does not coincide with an interruptible cognitive state.

Participant	P-only		P+C Policy		
	# Alerts	Mean sustain time (s)	# Alerts	Mean sustain time (s)	Mean rest probability
P1 (known)	15	2.20	2	2.83	.61
P2 (known)	5	5.41	1	5.72	.76
P3 (known)	4	2.47	1	2.97	.70
P4 (known)	4	3.36	1	5.73	.77
P5 (known)	18	1.96	2	3.25	.72
P6 (unseen)	18	3.25	2	5.42	.70
P7 (unseen)	7	2.72	1	7.96	.50
P8 (unseen)	3	4.97	2	2.87	.75
P9 (unseen)	16	1.58	1	3.01	.75
P10 (unseen)	14	1.96	7	2.19	.67
P11 (unseen)	1	2.13	2	2.57	.75
P12 (unseen)	14	2.26	1	4.35	.52
P13 (unseen)	6	2.30	0	–	–

Tradeoff between interruptibility and ergonomic correction. Across all 13 participants, P+C issues 23 alerts compared with 125 under P-only, corresponding to an 81.6% reduction in interruption

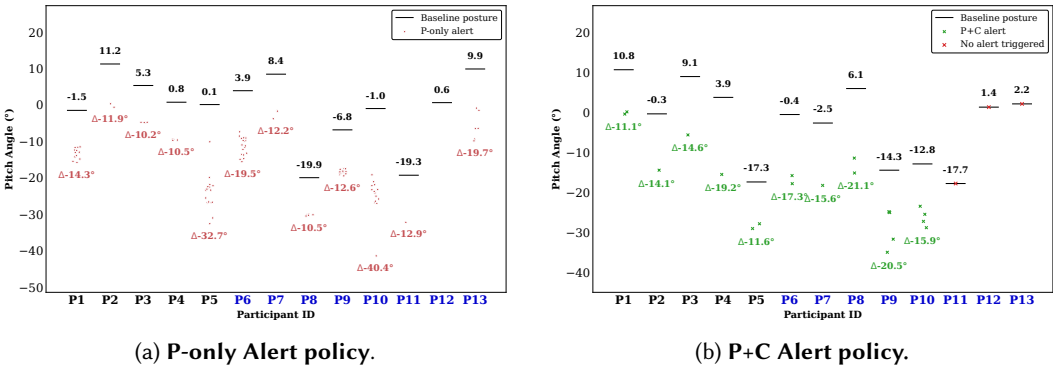


Fig. 9. Comparative alert and postural deviation analysis. The P+C policy triggers significantly fewer alerts by considering cognitive load compared to the P-only alert policy. P6 to P13 are unseen participants.

Table 17. Participant-level alert exposure and normalized posture response outcomes for five follow-up participants (P9–P13). **Response rate** is the normalized posture response, computed as corrected alerts divided by total issued alerts. **Mean pitch deviation at alert** is the mean absolute deviation of head pitch from each participant’s calibrated neutral baseline at alert onset, $|\text{pitch}_{\text{before}} - \text{baseline}|$. Alert counts differ across participants and policies because P-only triggers whenever sustained poor posture crosses the threshold, whereas P+C additionally requires alignment with a low-cognitive-load window or the 30 s override. Response rate normalises for these unequal exposures. Full per-alert pitch and recovery records are in Appendix C.

Participant	P-only				P+C Policy			
	# Alerts	# Corrected	Response rate	Mean pitch dev. (°)	# Alerts	# Corrected	Response rate	Mean pitch dev. (°)
P9	16	16	100.0%	11.44	1	1	100.0%	10.73
P10	14	1	7.1%	24.08	7	7	100.0%	11.30
P11	1	1	100.0%	12.92	2	2	100.0%	12.59
P12	14	5	35.7%	16.00	1	1	100.0%	17.46
P13	6	4	66.7%	15.62	0	0	N/A [†]	–
Participant-level Mean (N=5)	10.2	5.4	61.9%	16.01	2.2	2.2	100.0%[†]	13.02[†]

[†] P13 receives no P+C alert during the session and is excluded from P+C response-rate and mean pitch-deviation averages.

frequency (Table 16). To examine whether this reduction preserves ergonomic benefit, we conduct a post-hoc study for five unseen participants (P9–P13). Table 17 reports participant-level alert counts, normalized response rates, and pitch deviations, along with full per-alert pitch and recovery records in Appendix C. Despite heterogeneous alert counts, a consistent pattern of response rate shows that P+C substantially reduces alert overall while maintaining ergonomic benefits.

Participant-Level Response Rate Analysis: To assess ergonomic benefit and interruptibility, we use the normalized mean response rate as the primary proxy metric. Here, response rate is computed as corrected alerts divided by total issued alerts for each participant. From Table 17, P+C reduces the mean participant-level alert count from 10.2 to 2.2 per session, corresponding to a 78.4% reduction in alert exposure. The mean participant-level response rate increases from 61.9% under P-only to 100.0% under P+C. Across this subset, P+C consistently achieves **equal or higher response rates** per alert while delivering substantially fewer interruptions. This pattern

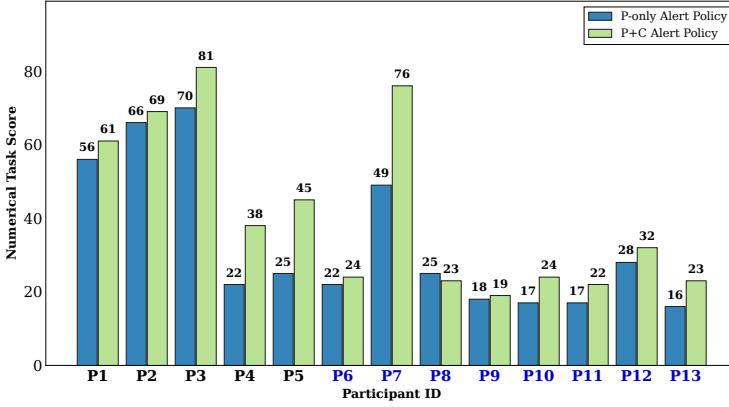


Fig. 10. Comparison of arithmetic task performance between the posture-only (P-only) and cognition-aware (P+C) alert policy.

holds across different participant profiles. Participants with low P-only responsiveness benefit most. P10 responds to only 1 of 14 P-only alerts (7.1%) but to all 7 P+C alerts (100.0%), consistent with alert fatigue under posture-threshold triggering. Similarly, P12 improves from 5 of 14 P-only alerts (35.7%) to 1 of 1 P+C alerts (100.0%). Participants who already respond consistently under P-only also benefit through reduced alert burden. P9 and P11 already achieve 100.0% response under P-only. In these cases, P+C provides no additional gain in response rate, but it reduces alert exposure; for example, P9's alerts decrease from 16 to 1 with no loss in correction outcome. P11 receives two P+C alerts, both triggered by the 30s override, confirming that the override remains active when posture does not recover within the permitted duration. P13 receives no P+C alert because the participant's poor-posture episodes did not persist long enough to trigger the override and did not coincide with a low-cognitive-load interruptible window.

Variation in Alert Exposure and Participant Response: In Table 17, alert counts and response rates differ across participants for three reasons. First, participants differ in neutral posture, sitting habits, pain sensitivity, musculature, and tolerance for sustained flexion. These individual differences determine how often each participant enters poor posture and how long each participant remains in it. Identical alert criteria therefore expose participants to different alert frequencies. Second, the two policies trigger alerts under different conditions. P-only issues an alert upon detecting sustained non-neutral posture. P+C applies the same posture criterion together with an evaluation of estimated task-induced cognitive load. P+C delays alerts during high-load periods unless poor posture persists for $T_{\text{cont}} = 30$ s. A participant sustaining high cognitive load throughout the task therefore receives fewer P+C alerts than a participant with moderate or variable load, provided the override never triggers. The fixed override also interacts with individual posture behavior. Participants recovering from poor posture before the 30 s window may receive few or no P+C alerts. Participants remaining in poor posture beyond 30 s receive a P+C alert even during high cognitive load, because the override limits prolonged uncorrected posture. Third, participants respond to alerts differently. Some participants ignore frequent alerts under P-only. This pattern is consistent with alert fatigue [6]. The same participants respond reliably to the less frequent and better-timed alerts under P+C. Other participants correct posture consistently under both policies.

Override Rationale: Our choice of a fixed continuous bad-posture override parameter ($T_{\text{cont}} = 30$ s) is motivated by ergonomics and musculoskeletal research linking risk to posture angle, exposure duration, repetition, and recovery profile, rather than brief isolated deviations [10, 30,

100, 107, 109]. EMG-based studies similarly show that muscle fatigue manifests as gradual changes in EMG features over sustained postures, on time scales from 120 s [55] to 15 minutes [45]. We therefore treat 30 s as a practical design bound that limits uncorrected exposure rather than a clinical cutoff, and it can be tightened for users with weaker musculature or pre-existing pain. In our design, cognition-aware alerts preserve focus during high-engagement periods, and the timeout override ensures that prolonged poor posture still triggers intervention when no interruptible state occurs, providing a bounded tradeoff between interruption reduction and ergonomic safety.

In summary, alert count alone is an insufficient basis for comparing policies because it does not indicate whether alerts actually lead to ergonomic corrections. Therefore, we use the normalized response rate to capture this trade-off. For our P+C policy, the normalized response rate reveals a consistent trend of maintaining or improving posture responses relative to the P-only policy, while also reducing alert fatigue.

User experience and interruptibility analysis. Our preliminary within-subjects evaluation ($N = 13$) shows improved user experience under the P+C policy. The mean System Usability Scale (SUS) increases to 74.7 ± 9.1 compared to 52.3 ± 13.3 for P-only ($t(12) = 7.56, p < .001$). Perceived interruptibility also improves, with lower disruption scores under P+C (5.28 ± 0.85) than P-only (2.82 ± 0.75), yielding a large effect size (Cohen's $d = 2.07$). These results indicate that context-aware alerting reduces perceived interruption while maintaining task engagement. Given the limited sample size and short session duration, we treat these findings as preliminary. Qualitative feedback further suggests that participants found the P+C system better aligned with task demands and less intrusive during high cognitive load.

Effectiveness Observation. Figure 10 shows improved task performance under the P+C alert policy compared to the P-only baseline. Out of 13 participants, 12 achieved higher arithmetic task scores under the P+C condition. The mean score increases from 33.2 ($SD = 18.9$) under P-only to 41.3 ($SD = 21.8$) under P+C, corresponding to an average improvement of 8.2 points. These results indicate that cognition-aware alert scheduling supports task performance by reducing poorly timed interruptions. These observations are directionally consistent with prior work on intelligent interruption in HCI, suggesting that context-aware alert scheduling can improve intervention timing and reduce perceived disruption during active tasks [18, 44]. Our results motivate future longitudinal studies to examine long-term adherence, broader user preference, and the design space of cognition-aware intervention strategies across more diverse real-world settings.

Key Takeaway: Cognition-aware alerting significantly reduces alert frequency while improving user experience compared to posture-only feedback. By deferring alerts during high cognitive load and intervening at receptive moments, the system delivers fewer but better-timed interventions without compromising posture correction. Participants consistently perceived cognition-aware alerts as less disruptive, more acceptable, and better aligned with sustained productivity.

8 Discussion

Our work focuses on intelligent, adaptive interventions, an often overlooked aspect of posture assistants. Existing systems typically trigger corrections based on minor posture deviations, which can be disruptive and impractical for long-term use. We address this by introducing a cognition-aware alert policy that considers task-induced cognitive load in addition to posture.

Necessity of Multimodal Sensing: IMU data reliably captures posture and correlates with physical discomfort, while EEG provides insight into task-induced cognitive load. *ErgoAssist* combines these modalities by using IMU to detect posture deviations and EEG to time alerts during less disruptive moments. Our real-time study shows that this integration improves user experience by reducing

alert burden and increasing perceived usability. Our findings align with prior work on posture-aware systems [9], which report that frequent threshold-based alerts can lead to alert fatigue and reduced compliance. We show that incorporating cognitive context significantly reduces unnecessary interruptions while maintaining ergonomic benefits.

Enhancing User Experience and Productivity: The proposed policy reduces alerts by 81% while maintaining ergonomic effectiveness. This suggests that avoiding interruptions during high-focus periods helps preserve productivity. The model generalizes well to unseen participants, indicating that alerts are both timely and meaningful. These results highlight that ergonomic feedback should be interruptibility-aware rather than purely threshold-based. Instead of enforcing constant correction, our approach focuses on preventing prolonged harmful posture, making the system more practical and user-friendly. This is consistent with prior findings in interruptibility research [2, 87], implying that interventions should be delivered during periods of low cognitive load, when users are more receptive to behavioral correction.

8.1 Practicality and Limitations

While our results demonstrate the feasibility of cognition-aware ergonomic assistance, several limitations merit discussion. First, long-term wearability, social acceptability, and signal robustness of consumer-grade wearable EEG in unconstrained environments remain open challenges. The Frenz Brainband is well suited for detecting broad task-induced cognitive load contrasts such as Rest versus Active [63], but fine-grained cognitive state distinctions and source-level analyses typically require higher-density channels [50]. Our system is better characterized as suitable for interruptibility-aware alerting rather than precise neurological assessment. Second, our study focuses on short-term controlled tasks and future work is needed to evaluate long-term behavioral change and adherence in real-world deployments. Third, controlled sessions are sufficient to elicit EEG-detectable cognitive load [85, 123] and postural degradation [55], but gradual postural drift during longer unstructured work sessions may follow different dynamics. Future longitudinal studies should examine how cognitive load and posture co-evolve over extended naturalistic work periods. Finally, our task-induced cognitive load model simplifies mental workload into coarse categories, sufficient for interruptibility-aware alerting, but richer cognitive models may further improve personalization at the cost of additional complexity and privacy concerns.

8.2 Future Directions

Future development of *ErgoAssist* will incorporate recommendations from end users and ergonomics specialists, including exploring smart glasses or helmet-based prototypes for robust skin contact and improved signal fidelity. Our current reliance on traditional machine learning presents limitations as feature selection is labor-intensive and requires domain expertise. We will therefore investigate deep learning architectures for automated feature extraction from raw sensor streams, better suited to scaling with larger datasets and capturing subtle temporal patterns. Future iterations may also support more nuanced activity recognition using large-scale representation learning to distinguish task types, adapt alert preferences, and account for collaborative work sessions. Our pilot evaluation in Section 7 is a preliminary case study rather than a definitive user study. The small sample ($N=13$), short session duration, and single baseline comparison limit the conclusions that can be drawn about generalizability and long-term effectiveness. Additionally, longer naturalistic work sessions may exhibit slower and more variable posture drift than our shortened session for each posture. In addition, the P+C policy exposes a limitation of the current fixed-threshold design: a single set of cooldown and override parameters cannot adapt to individual postural habits or cognitive engagement patterns. Future work should therefore investigate personalized timeout and deferral parameters—such as user-specific override thresholds and alert deferral windows—learned

from longitudinal observations of each user's posture patterns, work environment, and behavioral responses to feedback. Such personalization may enable the system to better balance ergonomic protection and interruptibility across users with different sitting habits, musculature, pain sensitivity, and tolerance for sustained flexion. Longer-term, in-the-wild studies are needed to examine how cognitive load, posture adaptation, and user responses evolve over extended real-world use, and to establish more robust evidence for the effectiveness of cognition-aware alerting policies. We also plan to monitor correct neck exercise performance [111] and extend focus beyond musculoskeletal health to other tech-related risks such as eye strain. Current solutions such as Apple's Screen Distance² issue disruptive alerts that are often ignored. We aim to develop cognition-aware and context-sensitive interventions to mitigate eye strain risk and integrate additional sensing modalities such as ambient sensors, cameras, and audio to capture environmental factors including workspace lighting and viewing distance.

9 Conclusion

In this work, we present *ErgoAssist*, a head-worn wearable system that integrates IMU-based posture sensing with EEG-based task-induced cognitive load estimation to enable cognition-aware ergonomic interventions. Our results show that incorporating timing into alert delivery is more effective than threshold-based approaches, reducing alert frequency by 81% while improving posture correction by 38%. These findings shows that modeling holistic user state enables fewer, better-timed interventions without sacrificing ergonomic effectiveness. Future work will evaluate long-term adherence and real-world performance through longitudinal deployments. The code and supplementary materials for this work are available at <https://github.com/umassos/ErgoAssist>.

Acknowledgments

We thank the anonymous reviewers for their suggestions for improving the paper. This research was supported in part by NSF grants 2211302, 2211888, 2213636, 2325956, and 2105494, and by U.S. Army contract W911NF-17-2-0196. Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the funding agencies.

References

- [1] 2019. Wearable EEG: Past, present and future. *Electronics* (2019).
- [2] Piotr D Adamczyk and Brian P Bailey. 2004. If not now, when? The effects of interruption at different moments within task execution. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 271–278.
- [3] Soyed Tuhin Ahmed, Michael Hefenbrock, and Mehdi B Tahoori. 2024. Tiny deep ensemble: Uncertainty estimation in edge ai accelerators via ensembling normalization layers with shared weights. In *Proceedings of the 43rd IEEE/ACM International Conference on Computer-Aided Design*. 1–9.
- [4] Hafeez Ullah Amin, Aamir Saeed Malik, Rana Fayyaz Ahmad, Nasreen Badruddin, Nidal Kamel, Muhammad Hussain, and Weng-Tink Chooi. 2015. Feature extraction and classification for EEG signals using wavelet transform and machine learning techniques. *Australasian physical & engineering sciences in medicine* 38, 1 (2015), 139–149.
- [5] Ushtar Amin, Fábio A Nascimento, Ioannis Karakis, Donald Schomer, and Selim R Benbadis. 2023. Normal variants and artifacts: importance in EEG interpretation. *Epileptic disorders* 25, 5 (2023), 591–648.
- [6] Jessica S Ancker, Alison Edwards, Sarah Nosal, Diane Hauser, Elizabeth Mauer, Rainu Kaushal, and With the HITEC Investigators. 2017. Effects of workload, work complexity, and repeated alerts on alert fatigue in a clinical decision support system. *BMC medical informatics and decision making* 17, 1 (2017), 36.
- [7] Ayesha Arooj, Aamir Aziz, Fariha Khalid, Muhammad Hussain Iqbal, and Huda Binte Ashfaq. 2022. Forward head posture in young adults: a systematic review: forward head posture in young adults. *THE THERAPIST (Journal of Therapies & Rehabilitation Sciences)* (2022), 32–35.

²<https://support.apple.com/en-us/105007>

- [8] Hugues Ghislain Atakla, Maguette Mbaye, Thomas Kofi Dakurah, Marie Kisito Quenum, Lounceny Fatoumata Barry, Daouda Wague, Cyrus Igor Sossa, Dismand Stephan Houinato, Alioune Badara Thiam, and Momar Code Ba. 2023. Tech Neck Syndrome: A global epidemic of the modern era among students at the University of Abomey Calavi in Benin. *Interdisciplinary Neurosurgery* 34 (2023), 101812.
- [9] Jason L Baer, Anita Vasavada, and Rajal G Cohen. 2022. Posture biofeedback increases cognitive load. *Psychological research* 86, 6 (2022), 1892–1903.
- [10] Bruce P Bernard. 1997. Musculoskeletal disorders and workplace factors: a critical review of epidemiologic evidence for work-related musculoskeletal disorders of the neck, upper extremity, and low back. (1997).
- [11] Hongliang Bi, Wenbo Zhang, Shuaihao Li, Yanjiao Chen, Chaoyang Zhou, and Tang Zhou. 2024. SmartSit: Sitting posture recognition through acoustic sensing on smartphones. *IEEE Transactions on Multimedia* 26 (2024), 8119–8130.
- [12] Tasneem Borhany, Erum Shahid, Wasim Ahmed Siddique, and Hussain Ali. 2018. Musculoskeletal problems in frequent computer and internet users. *Journal of family medicine and primary care* 7, 2 (2018), 337–339.
- [13] Geoffrey Brookshire, Jake Kasper, Nicholas M Blauch, Yunan Charles Wu, Ryan Glatt, David A Merrill, Spencer Gerrol, Keith J Yoder, Colin Quirk, and Ché Lucero. 2024. Data leakage in deep learning studies of translational EEG. *Frontiers in neuroscience* 18 (2024), 1373515.
- [14] Anne-Marie Brouwer, Maarten A Hogervorst, Jan BF Van Erp, Tobias Heffelaar, Patrick H Zimmerman, and Robert Oostenveld. 2012. Estimating workload using EEG spectral power and ERPs in the n-back task. *Journal of neural engineering* 9, 4 (2012), 045008.
- [15] Diogo Caixeiro, Tomás Cordeiro, Leandro Constantino, João Carreira, Rui Mendes, Cândida G Silva, and Maria António Castro. 2025. Effectiveness of Wearable Devices for Posture Correction: A Systematic Review of Evidence from Randomized and Quasi-Experimental Studies. *Applied Sciences* 16, 1 (2025), 81.
- [16] Guanting Cen, Chongyang Wang, Temitayo A Olugbade, Amanda C de C Williams, and Nadia Bianchi-Berthouze. 2022. Exploring multimodal fusion for continuous protective behavior detection. In *2022 10th International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, 1–8.
- [17] Kehan Chen. 2019. Sitting posture recognition based on openpose. In *IOP Conference Series: Materials Science and Engineering*, Vol. 677. IOP Publishing, 032057.
- [18] Kuan-Wen Chen, Yung-Ju Chang, and Liwei Chan. 2022. Predicting opportune moments to deliver notifications in virtual reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–18.
- [19] Bhawana Chhagani and Alan Seefeldt. 2024. NeckCare: Preventing Tech Neck using Hearable-based Multimodal Sensing. *arXiv preprint arXiv:2412.13579* (2024).
- [20] Samy Chikhi, Nadine Matton, and Sophie Blanchet. 2022. EEG power spectral measures of cognitive workload: A meta-analysis. *Psychophysiology* 59, 6 (2022), e14009.
- [21] Henrik W Christensen and Niels Nilsson. 1998. Natural variation of cervical range of motion: a one-way repeated-measures design. *Journal of manipulative and physiological therapeutics* 21, 6 (1998), 383–387.
- [22] Federico Cirett Galán and Carole R Beal. 2012. EEG estimates of engagement and cognitive workload predict math problem solving outcomes. In *International conference on user modeling, adaptation, and personalization*. Springer, 51–62.
- [23] Francesco Daghero, Alessio Burrello, Enrico Macii, Paolo Montuschi, Massimo Poncino, and Daniele Jahier Pagliari. 2023. Dynamic decision tree ensembles for energy-efficient inference on IoT edge nodes. *IEEE Internet of Things Journal* 11, 1 (2023), 742–757.
- [24] Minh Duc Dang, Duy Phuoc Luong, Christopher Napier, and Lawrence H Kim. 2025. Co-Design & Evaluation of Visual Interventions for Head Posture Correction in Virtual Reality Games. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–15.
- [25] Frédéric Dehais, Alex Lafont, Raphaëlle Roy, and Stephen Fairclough. 2020. A neuroergonomics approach to mental workload, engagement and human performance. *Frontiers in neuroscience* 14 (2020), 268.
- [26] Berken Utku Demirel, Adnan Harun Dogan, Juliete Rossie, Max Moebus, and Christian Holz. 2025. Beyond Subjectivity: Continuous Cybersickness Detection Using EEG-based Multitaper Spectrum Estimation. *IEEE Transactions on Visualization and Computer Graphics* 31, 5 (2025), 1234–1245.
- [27] Khumukcham A Devi and Sudhakar K Singh. 2023. The hazards of excessive screen time: Impacts on physical health, mental health, and overall well-being. *Journal of Education and Health Promotion* 12, 1 (2023), 413.
- [28] Yi Ding, Yaqin Cao, Qingxing Qu, and Vincent G Duffy. 2020. An exploratory study using electroencephalography (EEG) to measure the smartphone user experience in the short term. *International Journal of Human-Computer Interaction* 36, 11 (2020), 1008–1021.
- [29] Byron Donaldson, Helen Bayne, and Neil E Bezodis. 2021. Within-subject repeatability and between-subject variability in posture during calibration of an inertial measurement unit system. *ISBS Proceedings Archive* 39, 1 (2021), 224.
- [30] Yidan Dong, Ping Jiang, Xu Jin, Nanyu Jiang, Wenchu Huang, Yu Peng, Yuhong Shen, Lihua He, Mikael Forsman, and Liyun Yang. 2022. Association between long-term static postures exposure and musculoskeletal disorders among

- university employees: A viewpoint of inflammatory pathways. *Frontiers in public health* 10 (2022), 1055374.
- [31] Ellen Carolina dos Santos Assef, Alessandra Gotuzo Seabra Capovilla, and Fernando Cesar Capovilla. 2007. Computerized Stroop test to assess selective attention in children with attention deficit hyperactivity disorder. *The Spanish journal of psychology* 10, 1 (2007), 33–40.
 - [32] Saeid Edriss, Cristian Romagnoli, Lucio Caprioli, Vincenzo Bonaiuto, Elvira Padua, and Giuseppe Annino. 2025. Commercial vision sensors and AI-based pose estimation frameworks for markerless motion analysis in sports and exercises: a mini review. *Frontiers in Physiology* 16 (2025), 1649330.
 - [33] Alexandra Lauren Elliott. 2019. The Upright Go Wearable Posture Device: An Evaluation of Postural Health, Improvement of Posture, and Salivary Cortisol Fluctuations in College Students. (2019).
 - [34] Mohamed Emish and Sean D Young. 2024. Remote wearable neuroimaging devices for health monitoring and neurophenotyping: A scoping review. *Biomimetics* 9, 4 (2024), 237.
 - [35] Lin Feng, Ziyi Li, and Chen Liu. 2019. Are you sitting right?-sitting posture recognition using RF signals. In *2019 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM)*. IEEE, 1–6.
 - [36] Justyna Fercho, Michał Krakowiak, Rami Yuser, Tomasz Szmuda, Piotr Zieliński, Dariusz Szarek, and Grzegorz Miękisiak. 2023. Kinematic analysis of the forward head posture associated with smartphone use. *Symmetry* 15, 3 (2023), 667.
 - [37] Andrea Ferlini, Alessandro Montanari, Cecilia Mascolo, and Robert Harle. 2019. Head motion tracking through in-ear wearables. In *Proceedings of the 1st International Workshop on Earable Computing*. 8–13.
 - [38] Thalía Fernández, Thalía Harmony, Mario Rodríguez, Alfonso Reyes, Erzsébet Marosi, and Jorge Bernal. 1993. Test-retest reliability of EEG spectral parameters during cognitive tasks: I absolute and relative power. *International Journal of Neuroscience* 68, 3-4 (1993), 255–261.
 - [39] Frenz. 2023. FRENZ Brainband. <https://frenzband.com/>. [Online; accessed 20-July-2023].
 - [40] Jerome H Friedman. 2001. Greedy function approximation: a gradient boosting machine. *Annals of statistics* (2001), 1189–1232.
 - [41] Andrzej Galecki and Tomasz Burzykowski. 2012. Linear mixed-effects model. In *Linear mixed-effects models using R: a step-by-step approach*. Springer, 245–273.
 - [42] Diana E Gherman, Laurens R Krol, Marius Klug, and Thorsten O Zander. 2025. An investigation of a passive BCI's performance for different body postures and presentation modalities. *Biomedical Physics & Engineering Express* 11, 2 (2025), 025052.
 - [43] Rajdeep Ghosh, Nabamita Deb, Kaushik Sengupta, Anurag Phukan, Nitin Choudhury, Sreshtha Kashyap, Souvik Phadikar, Ramesh Saha, Pranesh Das, Nidul Sinha, et al. 2022. SAM 40: Dataset of 40 subject EEG recordings to monitor the induced-stress while performing Stroop color-word test, arithmetic task, and mirror image recognition task. *Data in Brief* 40 (2022), 107772.
 - [44] Nitesh Goyal and Susan R Fussell. 2017. Intelligent interruption management using electro dermal activity based physiological sensor for collaborative sensemaking. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies* 1, 3 (2017), 1–21.
 - [45] Xinliang Guo, Lei Lu, Mark Robinson, Ying Tan, Kusal Goonewardena, and Denny Oetomo. 2021. A weak monotonicity based muscle fatigue detection algorithm for a short-duration poor posture using semg measurements. In *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2238–2241.
 - [46] Gyu-hyun Han, Chung-hwi Yi, Seo-hyun Kim, Su-bin Kim, and One-bin Lim. 2023. Test–retest Reliability and Concurrent Validity of a Headphone and Necklace Posture Correction System Developed for Office Workers. *Physical Therapy Korea* 30, 3 (2023), 174–183.
 - [47] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*. Vol. 52. Elsevier, 139–183.
 - [48] Chanchan He, Xiru Yu, Yuhe Zhang, Yuanning Li, and Nan Jiang. 2026. Wearable EEG devices in the detection of mild cognitive impairment: a systematic review. *npj Digital Medicine* (2026).
 - [49] Alan Hedge, Singe Morimoto, and Daniel McCrobie. 1999. Cornell musculoskeletal discomfort questionnaire. *Ergonomics* (1999).
 - [50] Daniel P J Heib, Caroline Pape, and Thomas Schultz. 2021. Cognitive Load Monitoring With Wearables—Lessons Learned From a Machine Learning Challenge. In *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)*.
 - [51] Elizabeth Hellier and Diane L Seger Partners HealthCare. 2010. A review of human factors principles for the design and implementation of medication safety alerts in clinical information systems. (2010).
 - [52] Mohammad-Parsa Hosseini, Amin Hosseini, and Kiarash Ahi. 2021. A Review on Machine Learning for EEG Signal Processing in Bioengineering. *IEEE Reviews in Biomedical Engineering* 14 (2021), 204–218. doi:10.1109/RBME.2020.2969915
 - [53] Chih-Wei Hsu, Chih-Chung Chang, Chih-Jen Lin, et al. 2003. A practical guide to support vector classification. (2003).

- [54] Yi-Fang Hsu and Jarmo A Hämäläinen. 2022. Load-dependent alpha suppression is related to working memory capacity for numbers. *Brain Research* 1791 (2022), 147994.
- [55] Go Igarashi, Chieko Karashima, and Minoru Hoshiyama. 2016. Effect of cognitive load on seating posture in children. *Occupational therapy international* 23, 1 (2016), 48–56.
- [56] Shamsi T Iqbal and Eric Horvitz. 2007. Disruption and recovery of computing tasks: field study, analysis, and directions. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 677–686.
- [57] David Jiménez-Grande, S Farokh Atashzar, Valter Devecchi, Eduardo Martínez-Valdes, and Deborah Falla. 2022. A machine learning approach for the identification of kinematic biomarkers of chronic neck pain during single-and dual-task gait. *Gait & Posture* 96 (2022), 81–86.
- [58] Ju-Yeon Jung, Yeong-Bae Lee, and Chang-Ki Kang. 2024. Effect of forward head posture on resting state brain function. In *Healthcare*, Vol. 12. MDPI, 1162.
- [59] Manasa Kalanadhabhatta, Chulhong Min, Alessandro Montanari, and Fahim Kawsar. 2021. FatigueSet: A multi-modal dataset for modeling mental fatigue and fatigability. In *International Conference on Pervasive Computing Technologies for Healthcare*. Springer, 204–217.
- [60] Rushil Khurana, Elena Marinelli, Tulika Saraf, and Shan Li. 2014. NeckGraffe: a postural awareness system. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems*. 227–232.
- [61] Ae-Ryeong Kim, Ju-Hyun Park, Si-Hyun Kim, Kwang Bok Kim, and Kyue-Nam Park. 2022. The validity of wireless earbud-type wearable sensors for head angle estimation and the relationships of head with trunk, pelvis, hip, and knee during workouts. *Sensors* 22, 2 (2022), 597.
- [62] Andre Kingma, Wido La Heij, Luciano Fasotti, and Paul Eling. 1996. Stroop interference and disorders of selective attention. *Neuropsychologia* 34, 4 (1996), 273–281.
- [63] Michael T Knierim, Christian Zimny, Gabriel Ivucic, and Tobias Röddiger. 2025. Advancing wearable BCI: Headphone EEG for cognitive load detection in lab and field. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 9, 1 (2025), 1–26.
- [64] Rabia Koca and Yavuz Bahadır Koca. 2025. Anatomy-Based assessment of spinal posture using IMU sensors and machine learning. *Sensors* 25, 19 (2025), 5963.
- [65] Chalindu Kodikara, Sapumal Wijekoon, and Lakmal Meegahapola. 2025. FatigueSense: Multi-device and multimodal wearable sensing for detecting mental fatigue. *ACM Transactions on Computing for Healthcare* 6, 2 (2025), 1–36.
- [66] Nataliya Kosmyna, Caitlin Morris, Utkarsh Sarawgi, and Pattie Maes. 2019. Attentivu: a biofeedback system for real-time monitoring and improvement of engagement. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–2.
- [67] Yong-Hwan Kwon, Seungjin Kim, and Dongwon Lee. 2021. Real-Time EEG-Based Cognitive Workload Monitoring on Wearable Devices. *IEEE Access* 9 (2021), 85356–85366.
- [68] Daria La Rocca, Patrizio Campisi, Balazs Vegso, Peter Cserti, György Kozmann, Fabio Babiloni, and F De Vico Fallani. 2014. Human brain distinctiveness based on EEG spectral coherence connectivity. *IEEE transactions on Biomedical Engineering* 61, 9 (2014), 2406–2412.
- [69] Marieke M Lansbergen, Martijn Arns, Martine van Dongen-Boomsma, Desirée Spronk, and Jan K Buitelaar. 2011. The increase in theta/beta ratio on resting-state EEG in boys with attention-deficit/hyperactivity disorder is mediated by slow alpha peak frequency. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* 35, 1 (2011), 47–52.
- [70] Oscar D. Lara and Miguel A. Labrador. 2013. A Survey on Human Activity Recognition using Wearable Sensors. *IEEE Communications Surveys & Tutorials* 15, 3 (2013), 1192–1209. doi:10.1109/SURV.2012.110112.00192
- [71] Peter Le, Emily HL Mills, Charles A Weisenbach, and Kermit G Davis. 2024. Neck muscle coactivation response to varied levels of mental workload during simulated flight tasks. *Human Factors* 66, 8 (2024), 2041–2056.
- [72] Jaebong Lee, Eunji Cho, Minjae Kim, Yongmin Yoon, and Seungmoon Choi. 2014. [D06] PreventFHP: Detection and warning system for Forward Head Posture. In *2014 IEEE Haptics Symposium (HAPTICS)*. IEEE, 1–1.
- [73] Jiye Lee and Hanbyul Joo. 2024. Mocap everyone everywhere: Lightweight motion capture with smartwatches and a head-mounted camera. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 1091–1100.
- [74] Miaoyu Li, Zhuohan Jiang, Yutong Liu, Shuheng Chen, Marcin Wozniak, Rafal Scherer, Robertas Damasevicius, Wei Wei, Ziyi Li, and Zuxin Li. 2021. Sitsen: Passive sitting posture sensing based on wireless devices. *International Journal of Distributed Sensor Networks* 17, 7 (2021), 15501477211024846.
- [75] Yingxin Liu, Yang Yu, Zeqi Ye, Ming Li, Yifan Zhang, Zongtan Zhou, Dewen Hu, and Ling-Li Zeng. 2023. Fusion of spatial, temporal, and spectral EEG signatures improves multilevel cognitive load prediction. *IEEE Transactions on Human-Machine Systems* 53, 2 (2023), 357–366.
- [76] Mikael Lundqvist, Earl K Miller, Jonatan Nordmark, Johan Liljefors, and Pawel Herman. 2024. Beta: bursts of cognition. *Trends in cognitive sciences* 28, 7 (2024), 662–676.

- [77] Hu Luo, Tianhao Jin, Yu Zhang, Bohao Tian, Yuru Zhang, and Dangxiao Wang. 2023. A skin-integrated device for neck posture monitoring and correction. *Microsystems & Nanoengineering* 9, 1 (2023), 150.
- [78] Yuhan Luo, Bongshin Lee, Donghee Yvette Wohn, Amanda L Rebar, David E Conroy, and Eun Kyoung Choe. 2018. Time for break: understanding information workers' sedentary behavior through a break prompting system. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–14.
- [79] Ranjana K Mehta and Raja Parasuraman. 2013. Neuroergonomics: a review of applications to physical and cognitive work. *Frontiers in human neuroscience* 7 (2013), 889.
- [80] Vojkan Mihajlović, Bernard Grundlehner, Ruud Vullers, and Julien Penders. 2014. Wearable, wireless EEG solutions in daily life applications: what are we missing? *IEEE journal of biomedical and health informatics* 19, 1 (2014), 6–21.
- [81] Tim Mullen, Christian Kothe, Yu Mike Chi, Alejandro Ojeda, Trevor Kerth, Scott Makeig, Gert Cauwenberghs, and Tzyy-Ping Jung. 2013. Real-time modeling and 3D visualization of source dynamics and connectivity using wearable EEG. In *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2184–2187. doi:10.1109/EMBC.2013.6609968
- [82] Bilge Mutlu, Andreas Krause, Jodi Forlizzi, Carlos Guestrin, and Jessica Hodgins. 2007. Robust, low-cost, non-intrusive sensing and recognition of seated postures. In *Proceedings of the 20th annual ACM symposium on User interface software and technology*. 149–158.
- [83] Sunil Neupane, U Ali, and A Mathew. 2017. Text neck syndrome-systematic review. *Imperial journal of interdisciplinary research* 3, 7 (2017), 141–148.
- [84] Alex Nica, Siddhartha Banerjee, Rohan Ghosh, et al. 2022. Assessing Fatigue with Multimodal Wearable Sensors and Machine Learning. *arXiv preprint arXiv:2205.00287* (2022).
- [85] Ali Nirabi, Faridah Abd Rahman, Mohamed Hadi Habaebi, Khairul Azami Sidek, and Siti Yusoff. 2025. Cognitive load assessment through EEG: A dataset from arithmetic and Stroop tasks. *Data in Brief* 60 (2025), 111477.
- [86] Joseph K Nuamah and Younho Seong. 2018. Support vector machine (SVM) classification of cognitive tasks based on electroencephalography (EEG) engagement index. *Brain-Computer Interfaces* 5, 1 (2018), 1–12.
- [87] Mikio Obuchi, Wataru Sasaki, Tadashi Okoshi, Jin Nakazawa, and Hideyuki Tokuda. 2016. Investigating interruptibility at activity breakpoints using smartphone activity recognition API. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*. 1602–1607.
- [88] Heiko Oppel and Michael Munz. 2021. Analysis of Feature Dimension Reduction Techniques Applied on the Prediction of Impact Force in Sports Climbing Based on IMU Data. *AI* 2, 4 (2021), 662–683.
- [89] Simon O'Regan, Stephen Faul, and William Marnane. 2012. Automatic detection of EEG artefacts arising from head movements using EEG and gyroscope signals. *Medical engineering & physics* 35, 7 (2012), 867.
- [90] Huiming Pan, Hong Wang, Dongxuan Li, Kezhe Zhu, Yuxiang Gao, Ruiqing Yin, and Peter B Shull. 2024. Automated, IMU-based spine angle estimation and IMU location identification for telerehabilitation. *Journal of NeuroEngineering and Rehabilitation* 21, 1 (2024), 96.
- [91] Harsh Patel. 2025. *Developing a Multimodal Model for Activity Recognition Using Electroencephalography (EEG) and Human Activity Recognition (HAR) Technologies: A Pilot Study Utilizing High Intensity Interval Training (HIIT)*. Ph.D. Dissertation.
- [92] Ilia Pershin, Gian Candrian, Marionna Mürger, Gian-Marco Baschera, Maryam Rostami, Dominique Eich, and Andreas Müller. 2023. Vigilance described by the time-on-task effect in EEG activity during a cued Go/NoGo task. *International Journal of Psychophysiology* 183 (2023), 92–102.
- [93] Anastasios Petropoulos, Dimitrios Sikeridis, and Theodore Antonakopoulos. 2017. SPoMo: IMU-based real-time sitting posture monitoring. In *2017 IEEE 7th International Conference on Consumer Electronics-Berlin (ICCE-Berlin)*. IEEE, 5–9.
- [94] Sébastien Puma, Nadine Matton, Pierre-V Paubel, Éric Raufaste, and Radouane El-Yagoubi. 2018. Using theta and alpha band power to assess cognitive workload in multitasking environments. *International Journal of Psychophysiology* 123 (2018), 111–120.
- [95] Yanxu Qu, Wei Gao, and Chao Liu. 2023. SitPAA: Sitting posture and action recognition using acoustic sensing. *Electronics* 13, 1 (2023), 40.
- [96] Bujar Raufi and Luca Longo. 2022. An evaluation of the EEG alpha-to-theta and theta-to-alpha band ratios as indexes of mental workload. *Frontiers in Neuroinformatics* 16 (2022), 861967.
- [97] Brian J Roach and Daniel H Mathalon. 2008. Event-related EEG time-frequency analysis: an overview of measures and an analysis of early gamma band phase locking in schizophrenia. *Schizophrenia bulletin* 34, 5 (2008), 907–926.
- [98] S. Ltd. 2023. Superware Head Tracker. <https://superware.co.uk/headtracker-overview>. [Online; accessed 20-July-2023].
- [99] Aislinn Sandre and Sonya V Troller-Renfree. 2026. Power struggles: Absolute vs. relative EEG power in developmental neuroscience. *Developmental Cognitive Neuroscience* (2026), 101698.

- [100] Weiner Santos, Carmen Rojas, Rui Isidoro, Alejandro Lorente, Ana Dias, Gonzalo Mariscal, María Benlloch, and Rafael Lorente. 2025. Efficacy of ergonomic interventions on work-related musculoskeletal pain: A systematic review and meta-analysis. *Journal of clinical medicine* 14, 9 (2025), 3034.
- [101] Qijia Shao. 2024. Weaving Physical and Physiological Sensing with Computational Fabrics. In *Proceedings of the 22nd Annual International Conference on Mobile Systems, Applications and Services*. 748–750.
- [102] Akanksha Sharma and Mandeep Singh. 2015. Assessing alpha activity in attention and relaxed state: An EEG analysis. In *2015 1st International Conference on Next Generation Computing Technologies (NGCT)*. IEEE, 508–513.
- [103] Aditi Singh. 2019. *Shape-changing break reminders for people with repetitive strain injury*. Ph.D. Dissertation. Carleton University.
- [104] Martin Spüler, Carina Walter, Wolfgang Rosenstiel, Peter Gerjets, Korbinian Moeller, and Elise Klein. 2016. EEG-based prediction of cognitive workload induced by arithmetic: a step towards online adaptation in numerical learning. *Zdm* 48, 3 (2016), 267–278.
- [105] Martijn S Stenneberg, Herman Ten Berge, Bart de Leeuw, Erik Cattrysse, Rob de Bie, Henrica de Vet, and Gwen-dolijne GM Scholten-Peeters. 2023. Variability of active cervical range of motion within and between days in healthy participants: A prospective observational study. *Journal of Manipulative and Physiological Therapeutics* 46, 2 (2023), 125–131.
- [106] A Stipacek, Roland H Grabner, Christa Neuper, Andreas Fink, and AC Neubauer. 2003. Sensitivity of human EEG alpha band desynchronization to different working memory components and increasing levels of memory load. *Neuroscience letters* 353, 3 (2003), 193–196.
- [107] Leon M Straker, Peter B O’Sullivan, Anne J Smith, and Mark C Perry. 2009. Relationships between prolonged neck/shoulder pain and sitting spinal posture in male and female adolescents. *Manual therapy* 14, 3 (2009), 321–329.
- [108] Jian Su and Shijie Chen. 2025. Dynamic sitting posture recognition system using passive RFID tags in Internet of Things. *Chinese Journal of Electronics* 34, 2 (2025), 586–599.
- [109] Grace PY Szeto, Leon Straker, and Sally Raine. 2002. A field comparison of neck and shoulder postures in symptomatic and asymptomatic office workers. *Applied ergonomics* 33, 1 (2002), 75–84.
- [110] Hao-Yuan Tang, Shih-Hua Tan, Ting-Yu Su, Chang-Jung Chiang, and Hsiang-Ho Chen. 2021. Upper body posture recognition using inertial sensors and recurrent neural networks. *Applied Sciences* 11, 24 (2021), 12101.
- [111] Brett Taylor, Max Birk, Regan L Mandryk, and Zenja Ivkovic. 2013. Posture training with real-time visual feedback. In *CHI’13 Extended Abstracts on Human Factors in Computing Systems*. 3135–3138.
- [112] Total Health Chiropractic. n.d.. *Are Smartwatch Posture Alerts Helpful or Harmful?* <https://totalhealthchiropractic.com.au/smartwatches-posture-alerts-helpful-or-harmful/> Accessed: 2026-01-28.
- [113] Ming-Chih Tsai, Edward T-H Chu, and Chia-Rong Lee. 2023. An automated sitting posture recognition system utilizing pressure sensors. *Sensors* 23, 13 (2023), 5894.
- [114] Ronca Vincenzo, Cecchetti Marianna, Capotorto Rossella, Di Flumeri Gianluca, Giorgi Andrea, Germano Daniele, Borghini Gianluca, Babiloni Fabio, and Aricò Pietro. 2026. Beyond the lab: real-world benchmarking of wearable EEGs for passive brain-computer interfaces. *Brain Informatics* 13, 1 (2026), 3.
- [115] Sian Virtue-Griffiths, Alex Fornito, Sarah CH Thompson, Mana Biabani, Jeggan Tiego, Tribikram Thapa, Neil W Bailey, and Nigel C Rogasch. 2025. Task-related changes in aperiodic activity are related to visual working memory capacity independent of event-related potentials and alpha oscillations. *Imaging Neuroscience* 3 (2025), IMAG-a.
- [116] Chongyang Wang, Yuan Feng, Lingxiao Zhong, Siyi Zhu, Chi Zhang, Siqi Zheng, Chen Liang, Yuntao Wang, Chengqi He, Chun Yu, et al. 2024. Ubiphysio: Support daily functioning, fitness, and rehabilitation with action understanding and feedback in natural language. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 1 (2024), 1–27.
- [117] Sophie L Wang, Gene Civillico, Wesley Niswander, and Kimberly L Kontson. 2022. Comparison of motion analysis systems in tracking upper body movement of myoelectric bypass prosthesis users. *Sensors* 22, 8 (2022), 2953.
- [118] Yao Wang, Estate M Sokhadze, Ayman S El-Baz, Xiaoli Li, Lonnie Sears, Manuel F Casanova, and Allan Tasman. 2016. Relative power of specific EEG bands and their ratios during neurofeedback training in children with autism spectrum disorder. *Frontiers in human neuroscience* 9 (2016), 723.
- [119] Zishan Wang, Ruqiang Huang, Ye Yan, Zhiguo Luo, Shaokai Zhao, Bei Wang, Jing Jin, Liang Xie, and Erwei Yin. 2023. An improved canonical correlation analysis for EEG inter-band correlation extraction. *Bioengineering* 10, 10 (2023), 1200.
- [120] Li Wanying, Lou Shuwei, and Sun Qi. 2024. Robot-assisted upper limb rehabilitation training pose capture based on optical motion capture. *The International Journal of Advanced Manufacturing Technology* (2024), 1–12.
- [121] Wen Wen, Shrey Grover, Douglas Hazel, Peyton Berning, Frederik Baumgardt, Vighnesh Viswanathan, Olivia Tween, and Robert MG Reinhart. 2024. Beta-band neural variability reveals age-related dissociations in human working memory maintenance and deletion. *Plos Biology* 22, 9 (2024), e3002784.

- [122] Andreas Widmann, Erich Schröger, and Burkhard Maess. 2015. Digital filter design for electrophysiological data—a practical approach. *Journal of neuroscience methods* 250 (2015), 34–46.
- [123] Ronglong Xiong, Fanmeng Kong, Xuehong Yang, Guangyuan Liu, and Wanhui Wen. 2020. Pattern recognition of cognitive load using EEG and ECG signals. *Sensors* 20, 18 (2020), 5122.
- [124] Jammisetty Yedukondalu, Kalyani Sunkara, Vankayalapati Radhika, Sivakrishna Kondaveeti, Murali Anumothu, and Yadavavalli Murali Krishna. 2025. Cognitive load detection through EEG lead wise feature optimization and ensemble classification. *Scientific Reports* 15, 1 (2025), 842.
- [125] Matteo Zago, Matteo Luzzago, Tommaso Marangoni, Mariolino De Cecco, Marco Tarabini, and Manuela Galli. 2020. 3D tracking of human motion using visual skeletonization and stereoscopic vision. *Frontiers in bioengineering and biotechnology* 8 (2020), 181.
- [126] Jiadong Zhu, Rubén San-Segundo, and José M Pardo. 2017. Feature extraction for robust physical activity recognition. *Human-centric Computing and Information Sciences* 7, 1 (2017), 16.
- [127] Makia Zmitri, Hassen Fourati, and Nicolas Vuillerme. 2019. Human activities and postures recognition: From inertial measurements to quaternion-based approaches. *Sensors* 19, 19 (2019), 4058.

Table 18. Participant demographics.

Participant	Age	Gender	Laptop use	Other electronics use	Vision range
P1	18 to 25	Female	> 8 hrs	3 to 5 hrs	-4.00, -3.75
P2	26 to 35	Male	> 8 hrs	5 to 8 hrs	6/6
P3	26 to 35	Male	> 8 hrs	1 to 3 hrs	6/6
P4	26 to 35	Female	> 8 hrs	1 to 3 hrs	6/6
P5	36 to 45	Male	1 to 3 hrs	1 to 3 hrs	6/6
P6	26 to 35	Female	> 8 hrs	3 to 5 hrs	6/6
P7	26 to 35	Male	> 8 hrs	1 to 3 hrs	6/6
P8	26 to 35	Male	> 8 hrs	3 to 5 hrs	-0.25, -0.25
P9	26 to 35	Male	5 to 8 hrs	1 to 3 hrs	6/6
P10	26 to 35	Male	> 8 hrs	> 8 hrs	6/6
P11	26 to 35	Female	3 to 5 hrs	5 to 8 hrs	-2.5, -2.5
P12	26 to 35	Female	5 to 8 hrs	1 to 3 hrs	6/6
P13	18 to 25	Female	5 to 8 hrs	3 to 5 hrs	6/6
P14	18 to 25	Female	> 8 hrs	1 to 3 hrs	6/6
P15	36 to 45	Male	5 to 8 hrs	1 to 3 hrs	2/2.5
P16	26 to 35	Male	5 to 8 hrs	1 to 3 hrs	-3/-2
P17	18 to 25	Male	5 to 8 hrs	3 to 5 hrs	6/6
P18	18 to 25	Male	1 to 3 hrs	5 to 8 hrs	-2/-2
P19	26 to 35	Male	> 8 hrs	3 to 5 hrs	6/6
P20	26 to 35	Male	> 8 hrs	1 to 3 hrs	6/6
P21	26 to 35	Male	> 8 hrs	1 to 3 hrs	-3/-3
P22	26 to 35	Male	5 to 8 hrs	5 to 8 hrs	0.5/0.5
P23	26 to 35	Male	3 to 5 hrs	5 to 8 hrs	4/4
P24	26 to 35	Male	> 8 hrs	1 to 3 hrs	6/6

A Participant Demographics

Table 18 summarizes the demographic details of all 24 participants including age, gender, device usage patterns, and vision range.

B Questionnaire Items

Subjective Workload: NASA-TLX. Participants rate the following dimensions from 1 (very low) to 7 (very high):

- **Mental Demand:** How much mental and perceptual activity was required?
- **Physical Demand:** How much physical activity was required?
- **Temporal Demand:** How much time pressure was felt?
- **Performance:** How successful were you in accomplishing the task goals?
- **Effort:** How hard did you work to achieve your level of performance?
- **Frustration:** How irritated or stressed did you feel during the task?

Musculoskeletal Discomfort: CMDQ. Participants rate discomfort from 1 (no discomfort) to 7 (extreme discomfort) for the following body regions: Neck, Upper back, Arms, Eyes.

Modified System Usability Scale (SUS).

- (1) I think that I would like to use this app frequently for my work.
- (2) I found the app unnecessarily complex.
- (3) I thought the alerts were easy to understand and follow.

- (4) I think that I would need the support of a technical person to be able to use this app.
- (5) I thought there was too much inconsistency in the alert timing.
- (6) I would imagine that most people would learn to use the app very quickly.
- (7) I found the alerts very awkward to use during my tasks.
- (8) I felt very confident using the app to manage my posture.
- (9) I needed to learn a lot of things before I could get going with this app.

Table 19. Per-alert posture correction for unseen participants (P9–P13). **Baseline pitch:** neutral pitch from calibration. **Pitch before/after alert:** head pitch at alert onset and after correction. **Recovery time:** seconds from alert to correction; 0.00 = immediate, dashes = no correction within 30 s. **Total corrected:** alerts resulting in correction out of total issued. Baseline pitch values differ across P-only and P+C sessions because personalized calibration is recomputed independently before each session.

Participant	Condition	Alert #	Baseline pitch (°)	Pitch before alert (°)	Pitch after alert (°)	Recovery time (s)	Total corrected
P9	P-only	1		-17.65	-16.65	0.00	
	P-only	2		-19.44	-16.76	22.33	
	P-only	3		-18.69	-16.76	12.27	
	P-only	4		-17.93	-16.76	1.50	
	P-only	5		-17.46	-16.76	7.15	
	P-only	6		-18.55	-16.62	0.00	
	P-only	7		-17.57	-16.59	20.20	
	P-only	8	-6.83	-18.32	-16.59	9.77	16/16
	P-only	9		-17.60	-16.59	0.00	
	P-only	10		-18.41	-16.62	12.55	
	P-only	11		-19.11	-16.62	2.36	
	P-only	12		-18.66	-16.73	10.96	
	P-only	13		-18.88	-16.73	0.44	
	P-only	14		-18.16	-16.79	10.70	
	P-only	15		-18.35	-16.79	0.69	
	P-only	16		-17.49	-16.76	0.00	
	P+C	1	4.39	-6.34	-1.71	3.02	1/1
P10	P-only	1		-27.00	-10.94	2.13	
	P-only	2		-41.41	–	–	
	P-only	3		-26.77	–	–	
	P-only	4		-24.12	–	–	
	P-only	5		-21.04	–	–	
	P-only	6		-19.16	–	–	
	P-only	7	-0.98	-25.79	–	–	1/14
	P-only	8		-26.41	–	–	
	P-only	9		-25.01	–	–	
	P-only	10		-23.72	–	–	
	P-only	11		-23.19	–	–	
	P-only	12		-23.25	–	–	
	P-only	13		-21.88	–	–	
	P-only	14		-22.13	–	–	
	P+C	1		4.03	5.99	5.23	
	P+C	2		3.72	9.12	12.19	
	P+C	3		4.64	9.12	2.19	
	P+C	4	15.83	5.51	6.01	1.48	7/7
	P+C	5		4.28	6.01	2.89	
	P+C	6		3.92	15.86	14.41	
	P+C	7		5.64	11.86	4.40	
P11	P-only	1	-19.28	-32.20	-29.26	0.00	1/1
	P+C	1		-15.67	-10.66	2.99	
	P+C	2	-1.26	-12.03	-5.74	2.18	2/2
P12	P-only	1		-0.67	-0.50	0.00	
	P-only	2		-13.04	–	–	
	P-only	3		-10.52	–	–	
	P-only	4		-10.83	–	–	
	P-only	5		-5.79	–	–	
	P-only	6		-5.32	–	–	
	P-only	7		-4.90	-0.56	29.52	
	P-only	8	9.43	-5.06	-0.56	19.80	5/14
	P-only	9		-4.11	-0.56	9.05	
	P-only	10		-2.83	-0.56	0.00	
	P-only	11		-8.90	–	–	
	P-only	12		-7.27	–	–	
	P-only	13		-7.27	–	–	
	P-only	14		-5.51	–	–	
	P+C	1	-1.96	-19.42	-7.64	4.36	1/1
P13	P-only	1		-1.48	-0.11	0.00	
	P-only	2		-0.87	–	–	
	P-only	3		-9.37	–	–	
	P-only	4	9.88	-6.46	0.06	22.84	4/6
	P-only	5		-6.43	0.06	13.12	
	P-only	6		-9.85	0.06	3.09	
	P+C	0	6.21	–	–	–	0/0

C Detailed Per-Alert Posture Correction Records