J Musculoskelet Neuronal Interact 2026; 26(3):317-338

Journal of Musculoskeletal and Neuronal Interactions JMNI

Original Article

Increased Gait Variability while Texting and Walking with Added Cognitive Loads in Challenging Environments

Chi-Whan Choi[1], **Simone V. Gill**[1]

[1]Sargent College of Health and Rehabilitation Sciences, Boston University, USA

Abstract

Objectives: This study investigated (1) whether additional cognitive loads impair dual-task walking, especially when crossing obstacles of varying heights, and (2) if extreme gradient boosting (XGBoost) could identify key features of walking while texting. **Methods**: Twenty participants completed the following conditions: baseline walking, Task (walking + Typing: copying sentences, or Texting: answering questions), Obstacle (walking + crossing obstacles of three heights: low, medium, high), and combined (Task + Obstacle). For texting performance, we computed the rate, accuracy, and dual-task cost (DTC). Gait measures included mean spatiotemporal parameters, coefficient of variation (%CV) for each parameter, and mediolateral trunk acceleration. Four machine learning models were trained for evaluation: K-Nearest Neighbor, Support Vector Machines, and Random Forest, XGBoost. **Results**: XGBoost achieved the best performance, with a test weighted F1 score of 0.848, and identified rate, %CV for step and stride time, and DTC of rate as the most important features for classification of conditions. Task × Obstacle interaction was found for %CV of step time with greater variability during high obstacle crossing while Texting (β=1.742, p=0.029). **Conclusions**: Findings demonstrate that mobile phone use increases walking variability, particularly in complex environments. The result may lend weight to growing concerns about mobile phone use while walking in young adults.

Keywords: Dual Task Walking, Machine Learning, Obstacle Crossing, Texting While Walking, XGBoost

Introduction

Mobile phone use is prevalent among young adults, especially during postural and locomotor tasks in the face of various environmental conditions[1,2]. However, mobile phone use induces cognitive distraction, reduces visual attention to the environment, and alters motor skills via decreased arm motion and head mobility[3].

Accumulating evidence demonstrates that texting while walking could have a negative impact on postural and locomotor performance during both flat ground walking and obstacle crossing[1,2,4–6]. For example, previous research reported that walking with texting led to a reduction in stride length, cadence, gait velocity as well as increased double limb support time[6]. Texting while walking also increases the relative variability of walking, a surrogate marker associated with greater fall risk[7]. Another study showed that the medio-lateral center of mass displacement in the frontal plane significantly increased during stride in the texting condition during obstacle crossing among young adults[5]. Recent studies have shown that young adults prioritize texting over walking performance; when texting while walking, they reduce their gait speed and zigzag back and forth on a walking path[8,9], which leads to pedestrian injury[2].

Even though several studies indicate that texting could affect the kinematics of gait, other studies report that young adults can safely adapt their gait to use a mobile phone and to negotiate obstacles. For example, one study observed that healthy young adults walking on a split-belt treadmill had no significant effects on their gait and could maintain texting performance[10]. Similarly, another study reported that healthy young adults could successfully text while walking on flat ground and while negotiating obstacles[11]. These inconsistent results are likely due to differences in

The authors have no conflict of interest.

Corresponding author: Dr. Chi-Whan Choi, 111 Cummington Mall, Boston MA, 02215, USA
E-mail: cwchoi@bu.edu

Edited by: G. Lyritis
Accepted 13 April 2026

doi: 10.22540/JMNI-26-317

cognitive task difficulty. Some studies asked participants to text to answer questions[5,6,8,12,13] and others asked them to copy or type to text[10,14–17]. Each type of texting task imposes a different level of cognitive load. For example, copying sentences primarily relies on visual-spatial processing and sustained attention, whereas answering questions engages working memory and syntactic recall[18]. For simpler cognitive tasks such as copying text without generating content, young healthy adults likely have sufficient cognitive resources to control their balance and maintain their cognitive performance. In contrast, tasks requiring content generation, such as answering questions, may place greater cognitive demands and more significantly disrupt walking performance. However, it remains unclear how young adults manage walking and texting simultaneously when faced with varying cognitive demands in more complex walking environments including negotiating different obstacle heights.

Machine learning (ML) techniques have emerged as a potential way to discriminate the clinically relevant gait features to form a predictive model[19–22]. Among various ML algorithms including support vector machine, random forest, k-nearest neighbor, and extreme gradient boosting (XGBoost), XGBoost is an enhanced algorithm based on gradient boosting decision trees, characterized by internal variable selection followed by parallel construction of boosting trees[23]. Since XGBoost can create models with high accuracy and high performance generalization, it is a widely used method in the medical field[21,24,25]. For instance, XGBoost was used to classify high and low fall risk levels using gait spatiotemporal parameters[22] and to determine the most important features for detecting low muscle mass using gait kinematic and kinetic data in older adults[21].

The use of XGBoost instead of using kinematic or kinetic data alone is advantageous for two main reasons. First, kinematic and kinetic datasets often have missing values due to sensor malfunctions, data collection errors, or subjects not completing certain tasks, and the datasets usually include a large number of gait features. XGBoost can easily combat missing values with its automatic feature selection[25]. Second, XGBoost can be used to extract gait parameters of interest. A meta-analysis showed that texting while walking negatively affects spatiotemporal gait parameters linked with increased fall risk[3]. Therefore, XGBoost is expected to be sensitive enough to capture key features that reflect the impact of texting while walking. If the model successfully identifies gait features consistent with findings in the existing literature, it could be applied to other paradigms to explore more complex interactions in gait analysis.

The objective of this study was to evaluate how young adults walk while completing texting tasks during over-ground and obstacle conditions. To examine this, we quantified texting and walking performance as participants walked with or without varying cognitive demands (copying or answering questions), walked with or without obstacles of different heights, and performed combined texting and obstacle-crossing tasks. Varying obstacle heights were included to reflect increasing levels of physical challenge, as higher obstacles require greater motor planning, dynamic balance, and postural control, which may further strain attentional resources during dual-task walking. We hypothesized that the texting condition, obstacle condition, and their interaction would degrade texting and/or gait performance. We also hypothesized that XGBoost would identify the most important features with greater performance compared to other machine learning algorithms.

Table 1. Participant characteristics (N = 20) and familiarity with texting while walking (FTW).

	Mean or %	(SD)
Sex, female	60 %	
Age, years,	21.5	(2.8)
Height, cm,	170.7	(9.9)
Weight, kg,	80.9	(25.1)
Leg length, cm	89.08	(5.36)
Familiarity with texting while walking (FTW)	4	(0.86)
1	0 %	
2	0 %	
3	35 %	
4	30 %	
5	35 %	

Note: SD: Standard Deviation; Level of FTW: 1 – not at all familiar; 2 – Slightly familiar; 3 – Somewhat familiar; 4 – Moderately familiar; 5 – Extremely familiar.

Materials and Methods

Participants

Twenty participants were recruited from the Boston University community and the Greater Boston area (Table 1). Study eligibility included being between 18 and 35 years old, having no significant cardiovascular, vestibular, or other neurologic disorders, having no hip, knee, or foot pain on most days during the past 90 days, and having the ability to walk independently. These criteria were confirmed via participant reports and experimenters' observations. The study was approved by the Boston University Institutional Review Board (# 6835) and conformed to the Declaration of Helsinki. Informed consent was obtained before testing began.

Procedures and equipment

Participants were asked to answer how familiar they were with texting while walking (FTW) on a scale from 1 (not

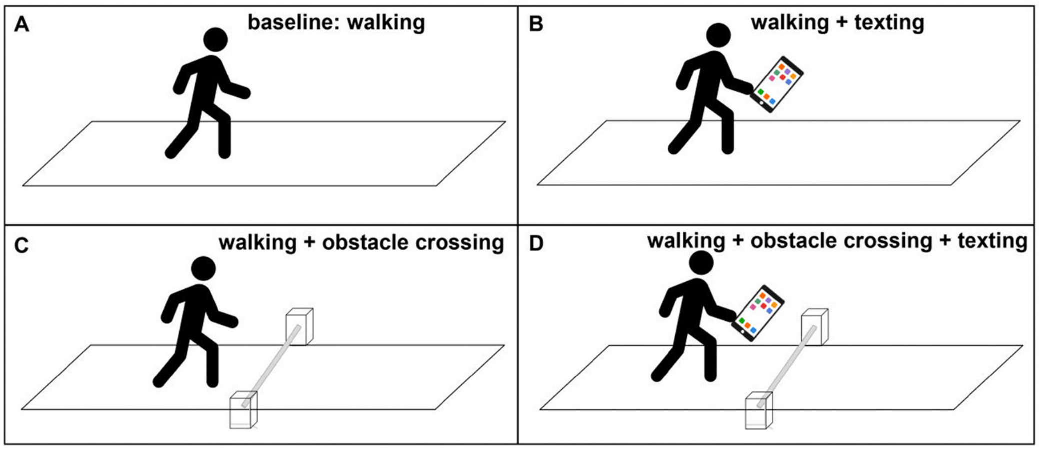


Figure 1. (A) Baseline motor task (walking). (B) Typing or Texting condition (walking + copying sentence; walking + answer the questions via text). (C) Obstacle condition (walking + crossing obstacles of low, medium, and high heights). (D) Combined condition (Walking + Typing or Texting with obstacle crossing of the 3 heights). Obstacles were created using a wooden dowel (121 cm long, 2.5 cm diameter) and 2 rectangular towers (9 cm x 10 cm x 22 cm) with holes drilled at 4, 8, and 15 cm (low, medium, and high). Towers were placed halfway down the walking path, creating visually salient obstacles (natural wood finish, high contrast against flooring) that required step-over negotiation.

at all familiar) to 5 (extremely familiar). Then participants performed two different texting tasks while seated: copying a given sentence (Typing) and answering a given question (Texting) with their own smartphone to assess their baseline texting. Autocorrect and autofill functions were turned off. We recorded texting response speed (rate: number of characters per second) and accuracy (number of correct responses/total number of answers provided). Questions were open-ended, validated and selected from prior texting while walking research[26,27]. Accuracy was operationalized as proportion of characters entered in correct sequential word order. Participants used their own smartphones with no pausing/stopping permitted during trials. Identical question sets were randomly provided across conditions with texting. Each response was sent to the researcher.

Participants walked at a self-selected pace across twelve conditions, including: (1) a baseline motor task (flat ground walking, Figure 1A); (2) two texting conditions—Typing and Texting while walking (Figure 1B); (3) three obstacle conditions involving low, medium, and high obstacle heights (Figure 1C); and (4) six conditions combining each texting task (Typing or Texting) with each obstacle height (low, medium, high; Figure 1D). Obstacle heights were set at 4 cm (low), 8 cm (medium), and 15 cm (high), representing common environmental challenges such as door thresholds, small steps, or tall steps[28]. For the obstacle trials, participants stepped over obstacles placed at the midpoint of the walkway 10m long and continued to the end, repeating each condition three times. Condition order was randomized, except for the baseline trials, which were always completed first.

For walking performance, spatiotemporal gait parameters were measured and analyzed using a gait carpet (Zeno Walkway, Protokinetics, Havertown, PA, USA; 6.10 m long $\times$ 0.89 m wide; 576 pressure sensors per square foot; sampling rate 120 Hz; spatial resolution 0.5 by 0.5 inches; temporal resolution 1/120 second). The gait carpet data was processed with Protokinetics Software (PKMAS), which uses the x and y coordinates of each step to calculate walking parameters. Spatiotemporal gait parameters include step length, stride length, stride width, step time, stride time, total double limbs support time, each coefficient of variation (%CV) of the parameters, cadence, and velocity were measured. Consecutive steps recorded on the gait carpet in each trial contributed to the calculation of spatiotemporal and variability measures. No additional trimming of the first or last step was performed since all steps detected within the gait mat placed in the middle of 10m walkway. %CV was calculated separately within each individual trial using the mean and standard deviation of the step level values for that trial. The mean number of steps were used to calculate %CV was 7.2 ± 0.6 per trial (range: 6–8). In addition, postural stability (trunk movement) was measured with inertial measurement units (MotionNode, Motion Workshop, Seattle, WA, USA) attached with stretchable Velcro straps to participants' thighs, shins, and around their waists. We collected data at a sampling frequency of 100 Hz and smoothed data using a 2nd-order Butterworth low-pass filter with a cut-off frequency of 6 Hz[29]. We synchronized IMUs and pressure walkway data using a heel-strike event recorded simultaneously by both systems. The heel-strike event was detected with each shank's IMU (prominent minimum in sagittal angular velocity)[30]. Mean mediolateral acceleration (ML_A), which successfully distinguished postural stability differences in texting while walking with or without obstacle crossing[26], was calculated from the mediolateral acceleration recorded by the waist-mounted IMU for each trial. For texting performance, response speed (Rate; higher=better) and accuracy (Accuracy; higher=better) of Typing or Texting, and their dual task costs (DTC), were calculated using Equation (1). A negative DTC indicates worse performance.

DTC = [(dual task performance - baseline performance)/ baseline performance] (1)

Data processing and statistical analysis

The ML model was trained on data calculated at the trial level. A total of 720 (20 subjects x 36 trials) gait trials were recorded. We excluded the trials resulting from (1) gait carpet sensor dropouts during obstacle negotiation, (2) insufficient step counts (<6 steps detected per trial), and (3) boundary effects where feet landed outside sensor range. For 12 Task (0, 1, 2) × Obstacle (0, 1, 2, 3) conditions (C), final N samples per condition were C_{00} (N=59), C_{01} (N=55), C_{02} (N=53), C_{03} (N=53), C_{10} (N=53), C_{11} (N=53), C_{12} (N=52), C_{13} (N=53), C_{20} (N=49), C_{21} (N=49), C_{22} (N=55), and C_{23} (N=53). To improve model stability given the modest sample size, the 'Obstacle' variable (low, medium, high obstacle) was recoded as a binary column: No Obstacle = 0 and Obstacle = 1. This binary recoding was implemented for the classification framework, as multi-class targets (12 conditions) can lead to unstable models with small samples[31]. A supplementary sensitivity analysis using the original 12-condition labeling was also implemented to match ML interpretation to the experimental manipulation. Each ML model including XGBoost, k-nearest neighbor (KNN), random forest, and support vector machine (SVM) was applied to discriminate six labels (L) with Equation (2): Task (No Task=0, Typing=1, Texting=2) and Obstacle variables (No Obstacle = 0, Obstacle = 1), using the Leave-One-Subject-Out (LOSO) cross-validation for the outer loop. Each label contained L_{00} (N=59), L_{01} (N=161), L_{10} (N=53), L^{11} (N=158), L_{20} (N=49), and L_{21} (N=157) trials. Variability metrics (%CV) used all consecutive steps within each trial (average 7.2 steps/trial). All trials from held-out participants (i.e., for the testing set) were completely excluded from both training and hyperparameter tuning during each outer fold of the nested cross-validation, ensuring no data leakage between training and testing sets. All preprocessing trials were embedded within a nested cross-validation framework to prevent data leakage. Standardization was applied only for SVM, KNN, XGBoost within training folds, while Random Forest was used without scaling. No explicit feature selection or correlation filtering was performed; all predefined gait and cognitive features were retained to preserve interpretability and ensure fair model comparison. Within each outer loop, hyperparameter tuning was performed using random search within a nested cross-validation framework, where the inner loop utilized Stratified Group K-Fold (K=4), and the weighted F1 score with Equation (3) was used as the primary performance metric. No class-weighting was applied during model training. Instead, class imbalance was addressed at the evaluation level using Balanced Accuracy, Macro F1, and Weighted F1 (primary metric). The procedure was repeated for 100 random-seed iterations. For each model and random seed, predictions from all LOSO folds were pooled to compute overall performance metrics.

Based on this procedure, XGBoost was selected as the final model with the highest average weighted F1 score in performance. The most important features were subsequently identified using SHapley Additive exPlanations (SHAP) with additional optimization across 100 random-seed iterations. All machine learning algorithms were implemented in Python (version 3.10) using Scikit-Learn for data preprocessing, model evaluation, and group-aware cross-validation, XGBoost for gradient boosting classification, SciPy for hyperparameter sampling, and SHAP for model interpretability (Supplementary Figure 1).

$$L=\begin{bmatrix} L_{00} & L_{01} \\ L_{01} & L_{11} \\ L_{10} & L_{21} \end{bmatrix}, \text{Task} \in \{0,1,2\}, \text{Obstacle} \in \{0,1\},$$

where L denotes the label (2)

k= 2 ·Task + Obstacle +1, $k \in \{1,2,3,4,5,6\}$.

$$F1_{weighted}=\sum_{k=1}^{6} \frac{n_k}{N} \cdot \frac{2\cdot \text{True Positive}_k}{2\cdot \text{True Positive}_k+\text{False Positive}_k+\text{False Negative}_k} \quad (3),$$

where n_k *denotes the number of samples in class k,and N denotes the number of all samples*.

To investigate the interaction between Task and Obstacle, we applied random effects linear mixed models (LMMs) in R software (version 4.4.0) with lme4 and lmerTest packages to compute Satterthwaite degrees of freedom and p-values, with models estimated via restricted maximum likelihood. Since age, sex, FTW and leg length (defined as the distance from anterior superior iliac spine to lateral malleolus) may be confounding factors, they were also added in the model. Finally, each Task (factor; 0, 1, 2), Obstacle (0, 1), age, sex, FTW, leg length and interaction between Task and Obstacle was modelled as a fixed effect while each participant was modelled as a random effect on the important features identified with SHAP. When significant interactions between Task and Obstacle were found, we examined how each task condition (Typing=1 or Texting=2) influenced response variables within each

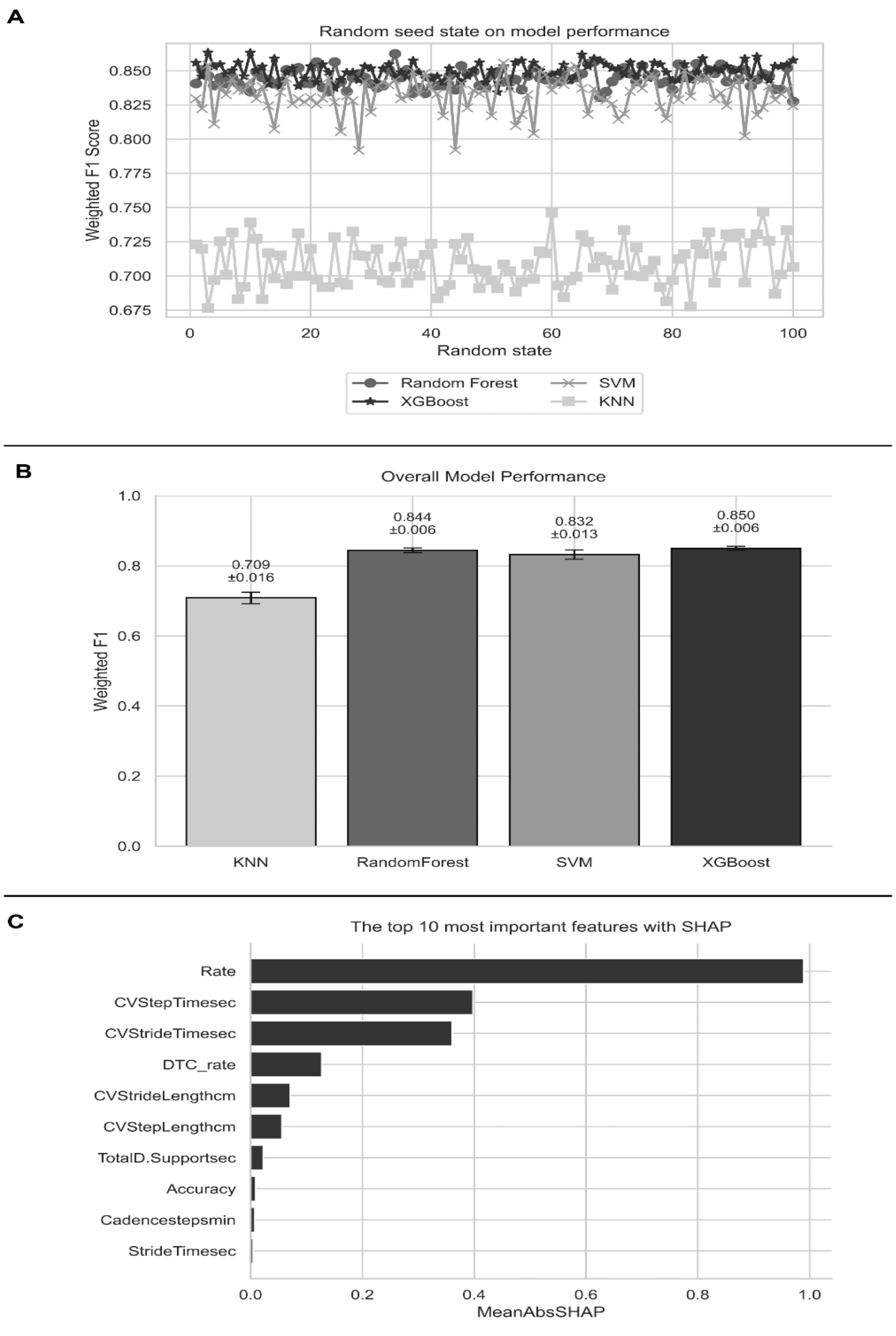


Figure 2. (A) Effect of random seed state on model performance across 100 repetitions. (B) Overall model performance comparison among k-nearest neighbor (KNN), random forest, support vector machine (SVM) and extreme gradient boosting (XGBoost). Each Weighted F1 score was averaged across 100 repetitions with incrementing random seed. Error bar indicates standard deviation between 100 random states. (C) The top ten most important features, ranked by their mean absolute SHAP values (MeanAbsSHAP) for XGBoost classifier. The features are shown from top to bottom in order of importance. Note: SHAP – SHapley Additive exPlanations; Rate – Rate of response (characters/sec); CVStepTimesec – Coefficient of variation (%CV) of step time; CVStrideTimesec – Coefficient of variation (%CV) of stride time; DTC_rate – Dual task cost of rate; CVStrideLengthcm – Coefficient of variation (%CV) of stride length; CVStepLengthcm – Coefficient of variation (%CV) of step length; TotalD.Supportsec – total double support time; Accuracy – Response accuracy (number of correct responses/total number of answers provided); Cadencestepmin – cadence (steps/minute); StrideTimesec – stride time.

Table 2. The effects of Task (No Task or Typing or Texting) and Obstacle (No obstacle or Obstacle) on selected features from the extreme gradient boosting (XGBoost).

	Source	*β*	*p*	*95%CI for β*
Rate	(Intercept)	-0.212	0.960	-8.262 – 7.839
	Obstacle	-0.517	0.046*	-1.022 – -0.011
	Task	-1.908	<.001**	-2.186 – -1.637
	Age	0.018	0.811	-0.128 – 0.164
	Sex	0.304	0.477	-0.512 – 1.119
	leg length	0.057	0.093	-0.005 – 0.119
	FTW	0.255	0.292	-0.202 – 0.711
	Obstacle*Task	0.172	0.293	-0.148 – 0.492
Conditional R^2	0.698			
Marginal R^2	0.484			
DTC_rate	(Intercept)	-0.613	0.464	-2.211 – 0.986
	Obstacle	-0.098	0.261	-0.269 – 0.073
	Task	0.006	0.903	-0.089 – 0.101
	Age	0.002	0.886	-0.027 – 0.031
	Sex	-0.008	0.806	-0.169 – 0.153
	leg length	0.005	0.472	-0.007 – 0.017
	FTW	0.036	0.445	-0.054 – 0.126
	Obstacle*Task	0.034	0.543	-0.075 – 0.143
Conditional R^2	0.205			
Marginal R^2	0.032			
%CV for step time	(Intercept)	2.344	0.736	-11.04 – 15.73
	Obstacle	8.136	<.001**	7.424 – 8.848
	Task	0.535	0.036*	0.037 – 1.034
	Age	0.203	0.122	-0.040 – 0.445
	Sex	-0.254	0.719	-1.612 – 1.104
	leg length	-0.060	0.276	-0.163 – 0.044
	FTW	0.357	0.372	-0.404 – 1.117
	Obstacle*Task	1.168	<.001**	0.564 – 1.771
Conditional R^2	0.674			
Marginal R^2	0.646			
%CV for stride time	(Intercept)	5.387	0.459	-8.498 – 19.28
	Obstacle	5.609	<.001**	5.016 – 6.201
	Task	0.446	0.036*	0.031 – 0.861
	Age	0.075	0.568	-0.176 – 0.326
	Sex	-0.638	0.389	-2.165 – 0.771
	leg length	-0.058	0.303	-0.165 – 0.049
	FTW	-0.005	0.990	-0.794 – 0.784
	Obstacle*Task	1.111	<.001**	0.608 – 1.613
Conditional R^2	0.631			
Marginal R^2	0.575			

*Note: No Task – walking without texting; Typing – copying sentences; Texting – answering questions; Rate – Rate of response (characters/sec); DTC_rate – Dual task cost of rate; %CV for step time – coefficient of variation (%CV) of step time; %CV stride time – coefficient of variation (%CV) of stride time; Conditional R^2 – the variance explained by both fixed and random effects in the model; Marginal R^2 – the variance explained by the fixed effects alone; Familiarity with texting while walking – FTW; * – Significance with $p < 0.05$; ** – Significance with $p < 0.001$.*

Table 3. The effects of Task (Typing or Texting) and Obstacle (low, medium, or high obstacle) on %CV for step time and %CV stride time.

	Source	β	p	95%CI for β
%CV for step time	(Intercept)	9.345	0.404	-11.98 – 30.67
	Low Obstacle	Ref.	Ref.	Ref.
	Medium Obstacle	2.941	0.021*	0.464 – 5.417
	High Obstacle	0.786	0.529	-1.656 – 3.228
	Task	0.296	0.600	-0.811 – 1.404
	Age	0.191	0.345	-0.194 – 0.577
	Sex	-0.072	0.948	-2.233 – 2.088
	leg length	-0.059	0.492	-0.223 – 0.105
	FTW	0.789	0.221	-0.421 – 1.999
	Medium Obstacle*Task	-0.692	0.387	-2.256 – 0.872
	High Obstacle*Task	1.742	0.029*	0.186 – 3.297
Conditional R^2	0.373			
Marginal R^2	0.197			
%CV for stride time	(Intercept)	16.99	0.139	-4.352 – 38.34
	Low Obstacle	Ref.	Ref.	Ref.
	Medium Obstacle	2.357	0.047*	0.045 – 4.668
	High Obstacle	2.410	0.039*	0.131 – 4.690
	Task	0.343	0.516	-0.691 – 1.376
	Age	-0.048	0.810	-0.434 – 0.337
	Sex	-0.915	0.420	-3.079 – 1.248
	leg length	-0.094	0.277	-0.259 – 0.070
	FTW	0.167	0.790	-1.044 – 1.379
	Medium Obstacle*Task	-0.608	0.360	-2.144 – 0.777
	High Obstacle*Task	0.283	0.703	-1.169 – 1.735
Conditional R^2	0.354			
Marginal R^2	0.141			

*Note: Typing – copying sentences; Texting – answering questions; %CV for step time – coefficient of variation (%CV) of step time; %CV stride time – coefficient of variation (%CV) of stride time; Conditional R^2 – the variance explained by both fixed and random effects in the model; Marginal R^2 – the variance explained by the fixed effects alone; FTW – Familiarity with texting while walking; * – Significance with p <0.05; ** – Significance with p <0.001; Ref. – Reference.*

obstacle height (factor; Low=1, Medium=2, High=3). Obstacle levels were analyzed using dummy variables to assess their interactions with key variables identified in prior analyses. Post-hoc t-tests with Benjamini-Hochberg corrections were conducted following significant Task × Obstacle interactions between all task conditions. Bootstrapping (B=1000 resamples) was employed to obtain 95% confidence intervals (95% CI) for fixed-effect estimates (the 2.5th and 97.5th percentiles). Statistical significance was defined as $p < 0.05$.

Results

Model performance comparison and important features by the best model

The XGBoost showed the highest performance (e.g., average performance in weighted F1 score with a very low standard deviation) compared to other machine learning algorithms with 100 iterations, indicating stable and reproducible performance (Figure 2A & 2B). Figure 2C presents the most important features for XGBoost classifier, ranked by their mean absolute SHAP values (SHAP based importance). The top four most important parameters with SHAP importance > 0.1 (features indicating their significant contribution to the model's predictions) were Rate (SHAP importance = 0.990), %CV for step time (SHAP importance = 0.398), %CV for stride time (SHAP importance = 0.360), and DTC of rate (SHAP importance = 0.127). Per-class performance metrics (precision, recall, F1) and confusion matrix are detailed in Supplementary Table 1, Figure 2.

Random effects linear mixed models result

LMM diagnostic plots (Supplementary Figure 3-6) confirmed normality of residuals (Q-Q plot) and

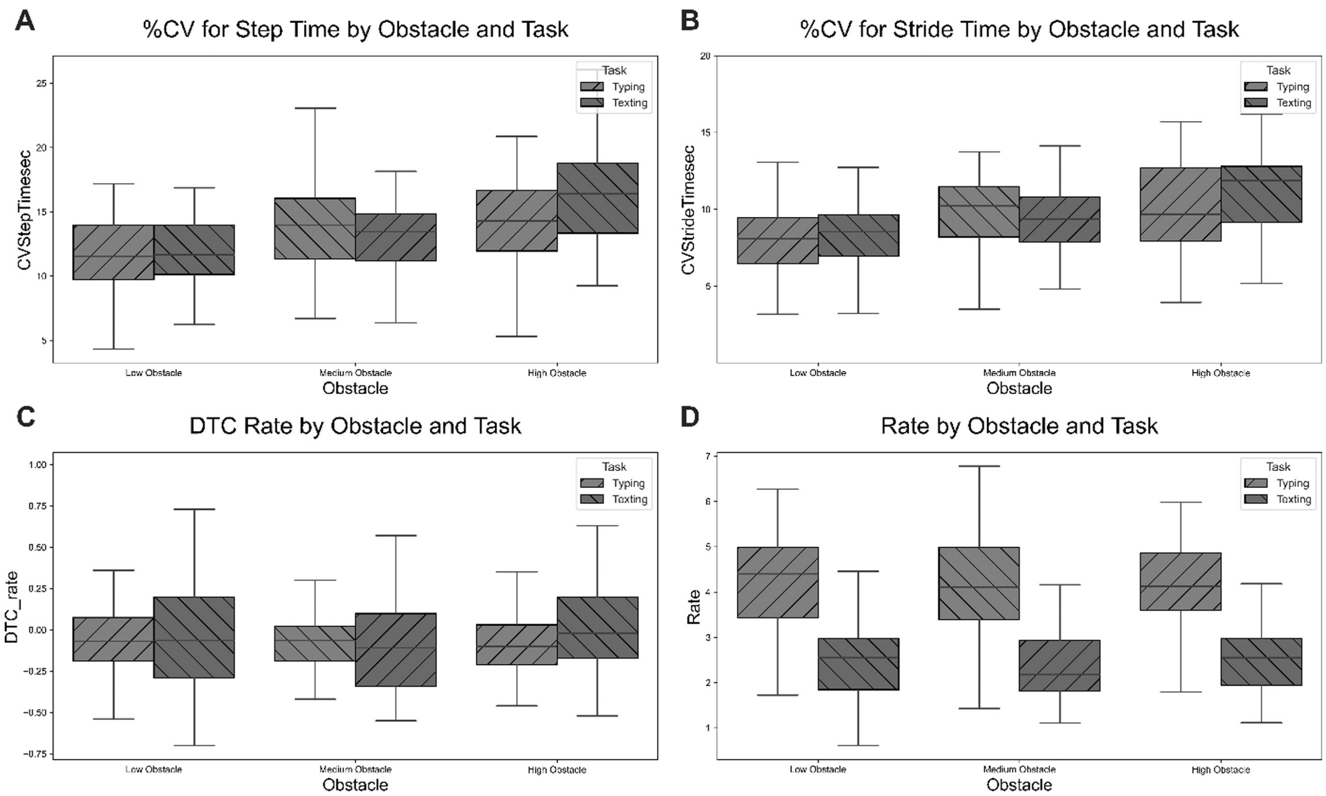


Figure 3. Task (Typing or Texting) and Obstacle (Low, Medium, or High obstacle) on (A) Coefficient of variation (%CV) for step time, (B) Coefficient of variation (%CV) for stride time, (C) Dual task cost (DTC) of rate, (D) Rate. Each forward-slash (/) pattern indicates Typing in three different heights of obstacles. Each backward-slash (\ ) pattern indicates Texting in three different heights of obstacles.

homoscedasticity (residuals vs fitted) for all SHAP-selected features. Task and Obstacle (β=-1.908, p<0.001; β=-0.517, p=0.046) significantly decreased Rate, but there was no significant impact on DTC of rate. Task, Obstacle and the interaction (β=0.535, p=0.036; β=8.136, p<0.001; β=1.168, p<0.001) significantly increased %CV for step time. Task, Obstacle and their interaction (β=0.446, p=0.036; β=5.609, p<0.001; β=1.111, p<0.001) significantly increased %CV for stride time. Age, sex, FTW, and leg length had no impact on Rate, DTC of rate, %CV for step time, and %CV for stride time (Table 2).

When Low Obstacle was treated as the reference level, Medium Obstacle significantly increased %CV of step time (β=2.941, p=0.021) and %CV of stride time (β=2.357, p=0.047). High Obstacle increased the %CV for stride time (β=2.410, p=0.039). A significant interaction was found between Task (Typing or Texting) and the High Obstacle condition for the %CV of step time (β=1.742, p=0.029), with the Low Obstacle condition used as the reference level (Table 3 and Figure 3A). The same trend was observed in %CV for stride time, DTC of rate, and Rate (Figure 3B, 3C, and 3D).

Based on post hoc analyses, Texting significantly increased %CV for step time compared to both No Task (p=0.023, Cohen's d=0.45) and Typing (p=0.005 Cohen's d=0.56) during the High Obstacle condition. There was no significant difference between Typing and Texting in other obstacle conditions (Table 4 and Figure 4).

Discussion

The primary aim of this study was to investigate how texting combined with varying cognitive loads would impact gait performance in young adults when crossing obstacles with different heights. Additionally, we investigated how well the ML algorithm XGBoost discriminates against the features most important to texting while walking. To our knowledge, this study extends existing research about the effect of texting while walking on flat ground and during obstacle crossing. We found that the XGBoost algorithm

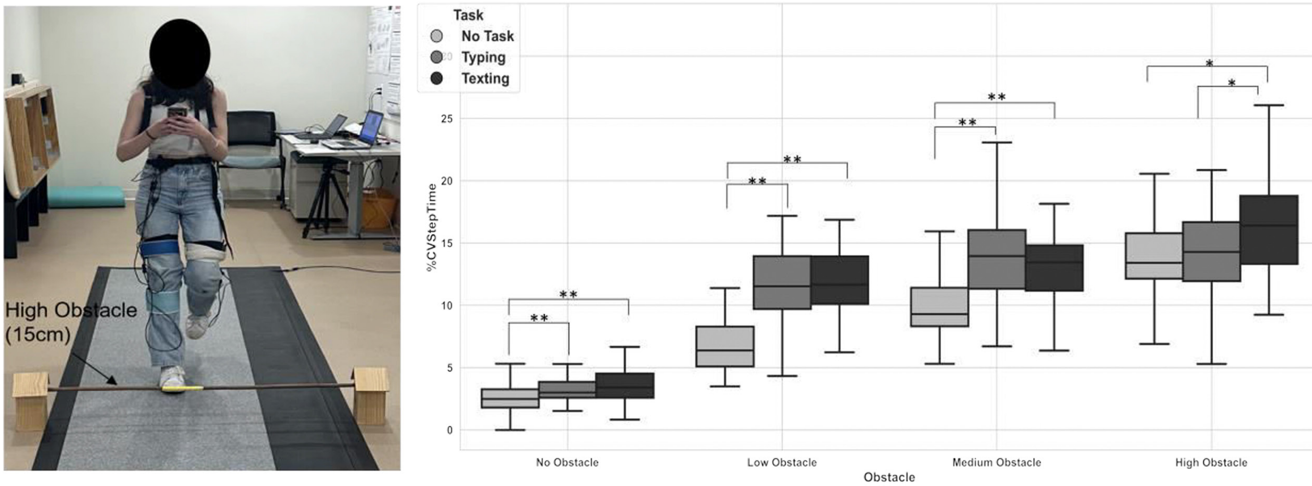


Figure 4. The post hoc analyses of coefficient of variation (%CV) of step time. Left column showed walking with texting (generating content) when crossing high obstacles. Right column showed comparison of coefficient of variation (%CV) for step time for Task (No Task or Typing or Texting) within Obstacle (No, Low, Medium, or High Obstacle). Each light gray bar indicates No Task, each medium gray bar indicates Typing and each dark gray bar indicates Texting in no obstacle and three different heights of obstacles. Note: * – Significance with p <0.05; ** – Significance with p <0.001 from post hoc analyses.

Table 4. The pairwise comparisons of coefficient of variation (%CV) of step time across three tasks—No Task, Typing, and Texting—within each obstacle condition (No, Low, Medium, and High).

Obstacle	*Task 1*	*Mean (SD)*	*Task 2*	*Mean (SD)*	*raw p-value*	*Cohen's d*	*95% CI for Cohen's d*	*FDR Corrected*
No Obstacle	No Task	2.64 (1.12)	Typing	3.32 (1.25)	<0.001	0.58	0.258 – 0.898	**<0.001****
	No Task	2.64 (1.12)	Texting	3.69 (1.77)	<0.001	0.76	0.438 – 1.086	**<0.001****
	Typing	3.32 (1.25)	Texting	3.69 (1.77)	0.184	0.24	0.119 – 0.607	0.246
Low Obstacle	No Task	7.07 (3.23)	Typing	11.65 (3.01)	<0.001	1.46	1.054 – 1.875	**<0.001****
	No Task	7.07 (3.23)	Texting	11.92 (3.11)	<0.001	1.53	1.105 – 1.949	**<0.001****
	Typing	11.65 (3.01)	Texting	11.92 (3.11)	0.640	0.09	-0.284 – 0.459	0.698
Medium Obstacle	No Task	9.96 (2.71)	Typing	13.84 (3.75)	<0.001	1.19	0.787 – 1.593	**<0.001****
	No Task	9.96 (2.71)	Texting	13.47 (3.13)	<0.001	1.20	0.802 – 1.591	**<0.001****
	Typing	13.84 (3.75)	Texting	13.47 (3.13)	0.572	0.11	-0.477 – 0.236	0.686
High Obstacle	No Task	14.31 (4.18)	Typing	14.14 (3.23)	0.805	0.05	-0.318 – 0.408	0.805
	No Task	14.31 (4.18)	Texting	16.15 (3.97)	0.016	0.45	0.082 – 0.821	**0.023***
	Typing	14.14 (3.23)	Texting	16.15 (3.97)	0.003	0.56	0.183 – 0.929	**0.005***

*Note: No Task – walking without texting; Typing – copying sentences; Texting – answering questions; SD – Standard Deviation; FDR – false discovery rate control; CI – Confidence interval for Cohen's d; * – Significance with p <0.05; ** – Significance with p <0.001.*

could effectively identify the important texting and gait features under varying cognitive loads with or without obstacles. In addition, we revealed that added cognitive loads impact gait variability more during crossing high obstacles. The findings highlight potential safety risks associated with multitasking during walking, reinforcing the need for public awareness initiatives to promote safer behaviors[2,32,33].

This study found that an XGBoost decision tree model could identify key indicators of texting and gait performance and demonstrated higher performance compared to other machine learning algorithms. This result is in agreement with previous work classifying low or high muscle mass using gait parameters in older adults that showed poorer discriminability when using the support vector machine, random forest, and k-nearest neighbor algorithms compared to the XGBoost model[21]. Even though the small gap between XGBoost and Random Forest suggests that both models are strong, XGBoost is slightly better in the average weighted F1 score throughout 100 iterations. SVM showed a larger difference in performance across the iterations, while KNN exhibited the worst performance. This result was consistent with other performance metrics such as Accuracy, Balanced Accuracy, Cohen's Kappa, Macro F1, for which XGBoost achieved the highest scores (Supplementary Table 2). Among the top ten features from SHAP, Rate, the %CV for step time, the %CV for stride time, and DTC of rate were deemed to be the most distinctive features, which is in agreement with other dual task walking studies[9,12,39,40,15,18,19,34–38]. 12-class sensitivity analysis confirmed SHAP feature stability, with identical top 4 predictors despite reduced performance, underscoring robust signals (Supplementary Figure 7). These findings were corroborated by effect sizes ranked by marginal R^2 in LMMs (Supplementary Table 3). Means and standard deviations for each outcome variable are reported in Supplementary Table 4.

The XGBoost model may pose several advantages over other models. First, the XGBoost model supports the use of SHAP for interpretability by calculating the individual contribution of each feature to model predictions based on game-theoretic principles[41]. This allows the model to identify the most relevant or sensitive variables, focusing on those likely to significantly affect classification outcomes and enhancing the interpretability of the results[23,42]. On top of that, it could handle high-dimensional data, manage noise, deal with multicollinearity, prevent overfitting, and ensure generalizability which is an ideal approach for the complex nature of gait studies[21,43]. Prior studies have commonly relied on relatively simple classifications, such as fallers versus non-fallers, which may oversimplify the complex interactions underlying gait performance. An important implication of XGBoost is its potential to extend beyond binary classification to multidimensional stratification. Unlike prior simplified groupings such as fallers/non-fallers[22], XGBoost may capture nonlinear interactions among spatiotemporal, kinetic, and variability features, enabling personalized rehabilitation strategies. Four-class stratification (intervention responder–fallers, intervention responder–non-fallers, intervention non-responder–fallers, intervention non-responder–non-fallers) could distinguish fallers requiring continued intervention from non-responder fallers needing alternative interventions. Combined with wearable sensors, XGBoost models could maximize predictive value for clinical gait studies. Nevertheless, algorithm selection procedure among the candidate classifiers may be necessary, as their performance can vary depending on the dataset.

In this study, young adults demonstrated the ability to modify their texting (e.g. Typing: texting without generating content) and walking behavior with respect to different environmental demands. However, the addition of a cognitive load such as working memory to answer questions (e.g. Texting: texting with generating content) may lend credibility to growing concerns about mobile phone use[9,44,45]. To be specific, the additional cognitive load needed to answer the questions significantly increased variability for step time while crossing high obstacles. Two main factors may contribute to the findings of the present study. First, cognitive distraction due to the additional of cognitive load can affect walking. The Texting condition likely required participants to manage multiple cognitive processes to compose meaningful messages. This complexity may have led participants to prioritize texting performance over walking. In contrast, the Typing condition, which involved entering a predetermined text without the need to write an original reply, would have posed a lower cognitive load. This result is consistent with other work that pure cognitive tasks such as mathematical calculations do not change walking performance in young adults[46]. The reasoning behind this is that tasks like mathematical calculations or basic typing primarily engage cognitive resources without requiring complex multitasking or decision-making, allowing participants to maintain adequate motor function. Second, environmental factors including obstacles may impact walking performance as well as texting performance[27]. Accumulated evidence demonstrates that walking is typically affected by the surrounding physical environment, requiring executive function, attention, and the ability to adapt to environmental complexity[47]. In the current study, walking with different obstacle heights may have caused participants to divide their attention from walking to texting performance: a similar occurrence in the real world. Prioritizing visual attention toward a mobile phone instead of an obstacle when performing both tasks simultaneously may pose additional risks to pedestrian safety.

Although there was no significant difference in other outcome variables, a gait performance-texting performance trade off occurred. The %CV for stride time increased when crossing the high obstacle in Texting condition, and the DTC of rate was close to zero when crossing the high obstacle in Texting condition, which means almost no dual task cost for Rate. Rate was also improved when crossing the high obstacle in the Texting condition. This may be closely related to pedestrian safety due to the fact that young adults may prioritize their texting activity even when they are facing risky circumstances in the real-world. Since gait variability reflects the ability to optimally control gait from one stride to the next[7,48], increased gait variability has been linked to postural instability[49].

Several limitations could have affected the results in the present work. First, our sample was relatively small. While internal validation suggested stable model performance in XGBoost, generalizability to broader populations cannot be assumed. Nevertheless, XGBoost may be particularly suited to small-to-moderate datasets due to their ability to handle nonlinearity, regularization mechanisms, and robustness to multicollinearity. In addition, habitual activity level, typical smartphone/texting use, and vision correction status were not collected, despite their potential influence on dual-task behavior and obstacle negotiation. Future studies should incorporate these variables to better account for inter-individual differences in dual-task performance and obstacle crossing. Lastly, participants in this study may mostly rely on the autocorrect and autofill operations. Since previous studies indicated that autocorrect and autofill functions may have made the texting task easier[14], participants were asked to turn off those functions before the experiment to acquire objective measures of texting performance. This lack of access to those functions may have affected their texting accuracy. This represents a potential limitation of ecological validity, suggesting that future studies may benefit from comparing performance with and without such assistive features to better capture real-world texting behavior.

Conclusion

The current study is the first to compare the impact of different texting tasks on walking when crossing obstacles of varying heights in healthy adults. Our results demonstrate that walking performance was not compromised when participants texted without generating content as obstacle height increased. However, participants walked with greater gait variability when crossing high obstacles compared to lower obstacles when they texted to generate content, while maintaining better texting performance. These results highlight a trade-off between gait and cognitive performance, showing how physical and cognitive demands interact to influence movement in complex environment.

To enhance pedestrian safety during texting while walking, we propose leveraging smartphone sensors (e.g., accelerometers, GPS, and Bluetooth) which can detect walking activity[50]. A potential safety feature could involve adding an "I'm Not Walking" button—similar to the "I'm Not Driving" mode—that appears when users attempt to text while walking with their smartphone. Further research is warranted to evaluate the effectiveness of such interventions on pedestrian safety during texting while walking.

Ethics Approval

This study was approved by the Boston University institutional review board (#6835) and was conducted in accordance with the Declaration of Helsinki (1964) and its later amendments.

Consent to Participate

Written informed consent was obtained from all participants.

Authors' Contributions

CC, and SG: Conceptualization, Methodology, Data collection. CC: Methodology, Data curation and cleaning, Planning data analyses, Formal analysis. CC: Writing – Original draft preparation. CC, and SG: Writing – Reviewing and editing. All authors approved the final manuscript before submission. Revision and final approval of the manuscript were done by all authors.

Funding

This study is independent research and was funded by: Sargent College Faculty Research, Boston University.

Acknowledgements

The authors acknowledge all participants for their time and effort. We thank members of the Master of Science in Statistical Practice Program at Boston University for assistance with data analysis and processing. This work was presented as a podium at the 2024 American Society of Biomechanics (ASB) Conference.

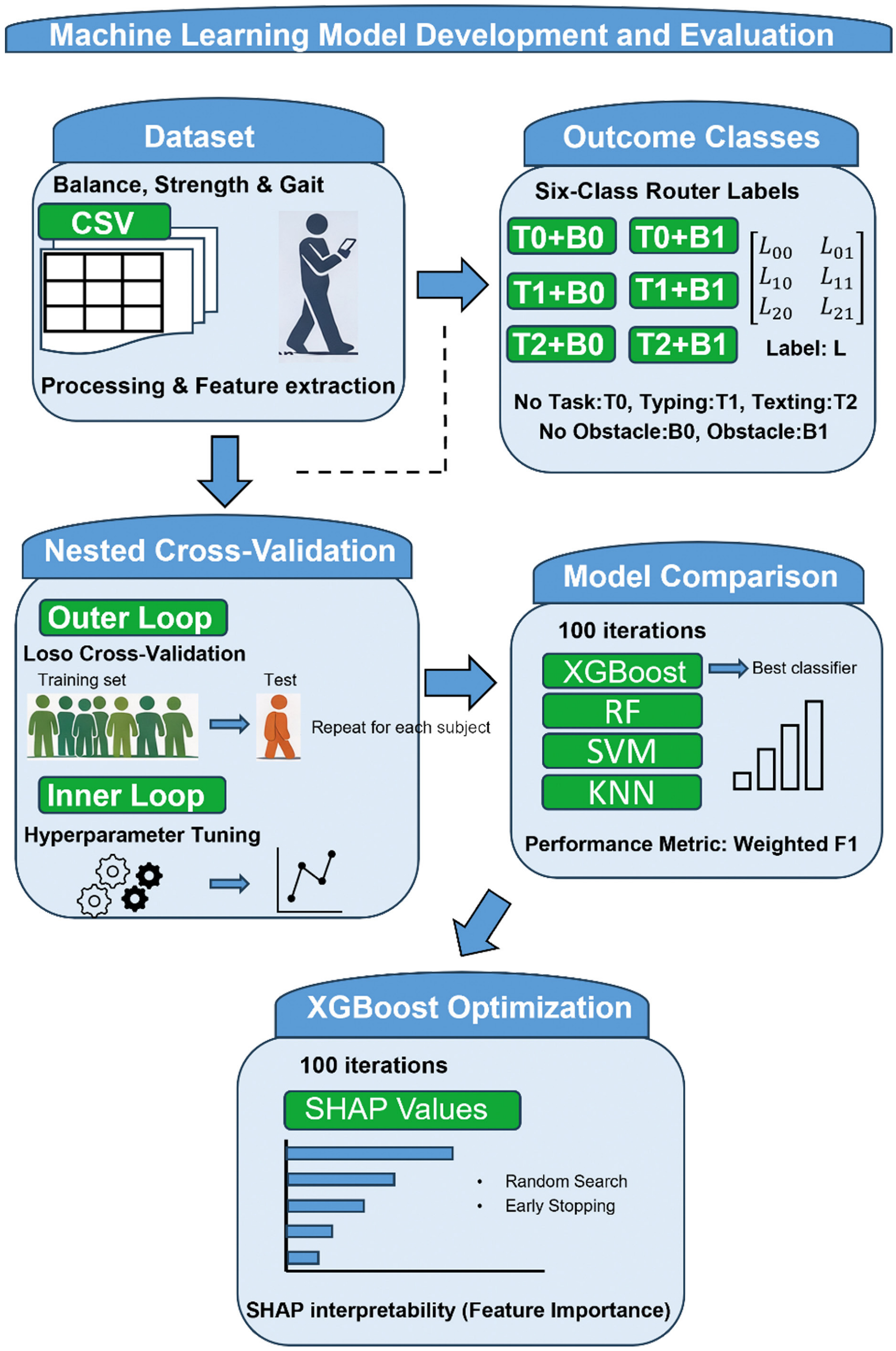
References

1. Bruyneel AV, Duclos NC. Effects of the use of mobile phone on postural and locomotor tasks: a scoping review. Gait Posture. 2020;82:233-241.
2. Gary CS, Lakhiani C, Defazio M V., Masden DL, Song DH. Smartphone use during ambulation and pedestrian trauma: A public health concern. J Trauma Acute Care Surg. 2018;85(6):1092-1101.
3. Bruyneel AV, Reinmann A, Gafner SC, Sandoz JD, Duclos NC. Does texting while walking affect spatiotemporal gait parameters in healthy adults, older people, and persons with motor or cognitive disorders? A systematic review and meta-analysis. Gait Posture. 2023;100(September 2022):284-301.
4. Al-yahya E, Dawes H, Smith L, Dennis A, Howells K, Cockburn J. Cognitive motor interference while walking: A systematic review and meta-analysis. Neurosci Biobehav Rev. 2011;35(3):715-728.
5. Chen SH, Lo OY, Kay T, Chou LS. Concurrent phone texting alters crossing behavior and induces gait imbalance during obstacle crossing. Gait Posture. 2018;62:422-425.
6. Crowley P, Madeleine P, Vuillerme N. The effects of mobile phone use on walking: A dual task study. BMC Res Notes. 2019;12(1):1-6.
7. Hausdorff JM, Rios DA, Edelberg HK. Gait variability and fall risk in community-living older adults: A 1-year prospective study. Arch Phys Med Rehabil. 2001;82(8):1050-1056.
8. Prupetkaew P, Lugade V, Kamnardsiri T, Silsupadol P. Cognitive and visual demands, but not gross motor demand, of concurrent smartphone use affect laboratory and free-living gait among young and older adults. Gait Posture. 2019;68(October 2018):30-36.

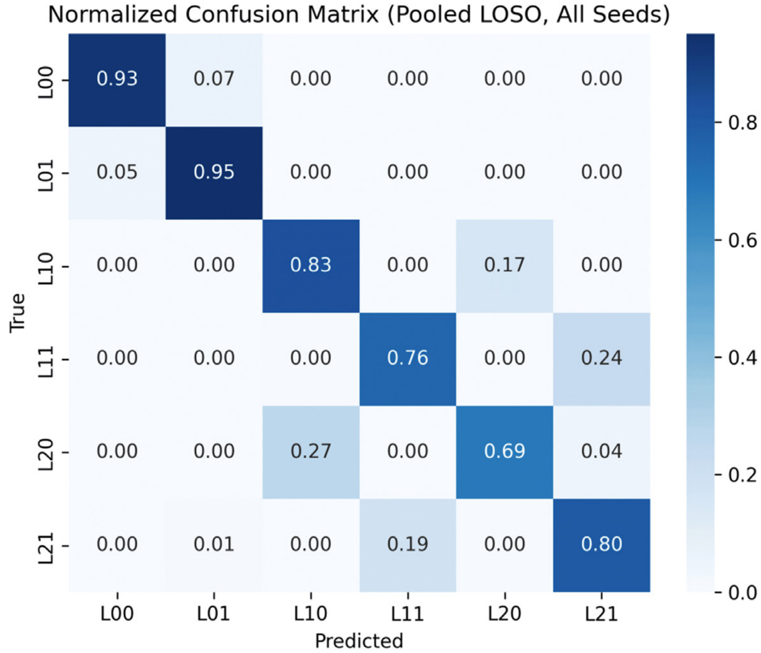
9. Hyman Jr IE, Boss SM, Wise BM, McKenzie KE, Caggiano JM. Did You See the Unicycling Clown? InattentionalBlindness while Walking and Talking on a Cell Phone. Appl Cogn Psychol. 2010;24(October 2009):597-607.
10. Hinton DC, Cheng YY, Paquette C. Everyday multitasking habits: University students seamlessly text and walk on a split-belt treadmill. Gait Posture. 2018;59(October 2017):168-173.
11. Timmis MA, Bijl H, Turner K, Basevitch I, Taylor MJD, Paridon KNV. The impact of mobile phone use on where we look and how we walk when negotiating floor based obstacles. PLoS One. 2017;12(6).
12. Agostini V, Lo Fermo F, Massazza G, Knaflitz M. Does texting while walking really affect gait in young adults? J Neuroeng Rehabil. 2015;12(1):1-10.
13. Licence S, Smith R, McGuigan MP, Earnest CP. Gait pattern alterations during walking, texting and walking and texting during cognitively distractive tasks while negotiating common pedestrian obstacles. PLoS One. 2015;10(7).
14. Plummer P, Apple S, Dowd C, Keith E. Texting and walking: Effect of environmental setting and task prioritization on dual-task interference in healthy young adults. Gait Posture. 2015;41(1):46-51.
15. Brennan AC, Breloff SP. The effect of various cell phone related activities on gait kinematics. J Musculoskelet Res. 2019;22(3-4):1-10.
16. Lim J, Kim J, Seo K, van Emmerik REA, Lee S. The effects of mobile texting and walking speed on gait characteristics of normal weight and obese adults. Motor Control. 2020;24(4):588-604.
17. Krasovsky T, Lanir J, Felberbaum Y, Kizony R. Mobile phone use during gait: The role of perceived prioritization and executive control. Int J Environ Res Public Health. 2021;18(16).
18. Ferraro FR, Holte AJ, Giesen D. Does texting make you slower? Curr Psychol. 2022;41(5):2980-2984.
19. Wahid F, Begg RK, Hass CJ, Halgamuge S, Ackland DC. Classification of Parkinson's disease gait using spatial-temporal gait features. IEEE J Biomed Heal Informatics. 2015;19(6):1794-1802.
20. Qiu H, Rehman RZU, Yu X, Xiong S. Application of Wearable Inertial Sensors and A New Test Battery for Distinguishing Retrospective Fallers from Non-fallers among Community-dwelling Older People. Sci Rep. 2018;8(1):1-10.
21. Fujita K, Hiyama T, Wada K, et al. Machine learning-based muscle mass estimation using gait parameters in community-dwelling older adults: A cross-sectional study. Arch Gerontol Geriatr. 2022;103(June):104793.
22. Noh B, Youm C, Goh E, et al. XGBoost based machine learning approach to predict the risk of fall in older adults using gait outcomes. Sci Rep. 2021;11(1):1-9.
23. Kavzoglu T, Teke A. Advanced hyperparameter optimization for improved spatial prediction of shallow landslides using extreme gradient boosting (XGBoost). Bull Eng Geol Environ. 2022;81(5).
24. Liu YH, Hung PCK, Iqbal F, Fung BCM. Automatic fall risk detection based on imbalanced data. IEEE Access. 2021;9:163594-163611.
25. Zhang X, Yan C, Gao C, Malin BA, Chen Y. Predicting Missing Values in Medical Data Via XGBoost Regression. J Healthc Informatics Res. 2020;4(4):383-394.
26. Tandon R, Javid P, Di Giulio I. Mobile phone use is detrimental for gait stability in young adults. Gait Posture. 2021;88(November 2019):37-41.
27. Demura S, Uchiyama M. Influence of cell phone email use on characteristics of gait. Eur J Sport Sci. 2009;9(5):303-309.
28. Gill S V. Effects of obesity class on flat ground walking and obstacle negotiation. J Musculoskelet Neuronal Interact. 2019;19(4):448-454.
29. Luo Y, Coppola SM, Dixon PC, Li S, Dennerlein JT, Hu B. A database of human gait performance on irregular and uneven surfaces collected by wearable sensors. Sci Data. 2020;7(1):1-9.
30. Greene BR, Foran TG, McGrath D, Doheny EP, Burns A, Caulfield B. A comparison of algorithms for body-worn sensor-based spatiotemporal gait parameters to the gaitrite electronic walkway. J Appl Biomech. 2012;28(3):349-355.
31. Approaches E based, Rocha A, Goldenstein SK, Member S. Multiclass from Binary : Expanding One-Versus-All, One-Versus-One and ECOC-Based Approaches. IEEE Trans Neural Networks Learn Syst. 2014;25(2):289-302.
32. Pelicioni PHS, Chan LLY, Shi S, et al. Impact of mobile phone use on accidental falls risk in young adult pedestrians. Heliyon. 2023;9(8):e18366.
33. Ha SY, Jung YJ, Shin DC. The effect of smartphone uses on gait and obstacle collision during walking. Med Hypotheses. 2020;141(March):109730.
34. Stöckel T, Mau-Moeller A. Cognitive control processes associated with successful gait performance in dual-task walking in healthy young adults. Psychol Res. 2020;84(6):1766-1776.
35. Hennah C, Doumas M. Dual-task walking on real-world surfaces: Adaptive changes in walking speed, step width and step height in young and older adults. Exp Gerontol. 2023;177(April):112200.
36. Postigo-Alonso B, Galvao-Carmona A, Conde-Gavilán C, et al. The effect of prioritization over cognitive-motor interference in people with relapsing-remitting multiple sclerosis and healthy controls. Akinwuntan AE, ed. PLoS One. 2019;14(12):e0226775.
37. Souza FHN de, Lisboa AKP, Maia MDFT, et al. Effects of dual task training on gait temporal-spatial parameters of children with autism. Man Ther Posturology Rehabil J. 2017;15:0-0.
38. Thomson D, Gupta A, Liston M. Does attention switching between multiple tasks affect gait stability and task performance differently between younger

and older adults? Exp Brain Res. 2020;238(12):2819-2831.
39. Hollman JH, Kovash FM, Kubik JJ, Linbo RA. Age-related differences in spatiotemporal markers of gait stability during dual task walking. Gait Posture. 2007;26(1):113-119.
40. Yogev-Seligmann G, Rotem-Galili Y, Mirelman A, Dickstein R, Giladi N, Hausdorff JM. How does explicit prioritization alter walking during dual-task performance? Effects of age and sex on gait speed and variability. Phys Ther. 2010;90(2):177-186.
41. Cai WM, Rui D, Zhao LB, et al. SHAP Analysis of Machine Learning-Based Carotid Atherosclerosis Prediction of Older Adults with OSA, a Multicenter Study. Nat Sci Sleep. 2025;17:2351-2367.
42. Gertz M, Große-Butenuth K, Junge W, et al. Using the XGBoost algorithm to classify neck and leg activity sensor data using on-farm health recordings for locomotor-associated diseases. Comput Electron Agric. 2020;173(January):105404.
43. Özgür S, Koçaslan Toran M, Toygar İ, Yalçın GY, Eraksoy M. A machine learning approach to determine the risk factors for fall in multiple sclerosis. BMC Med Inform Decis Mak. 2024;24(1).
44. Nasar J, Hecht P, Wener R. Mobile telephones, distracted attention, and pedestrian safety. Accid Anal Prev. 2008;40(1):69-75.
45. Hatfield J, Murphy S. The effects of mobile phone use on pedestrian crossing behaviour at signalised and unsignalised intersections. Accid Anal Prev. 2007;39(1):197-205.
46. Marone JR, Patel PB, Hurt CP, Grabiner MD. Frontal plane margin of stability is increased during texting while walking. Gait Posture. 2014;40(1):243-246.
47. Snijders AH, Warrenburg BP Van De, Giladi N, Bloem BR. Neurological gait disorders in elderly people : clinical approach and classification. 2007;6(January):63-74.
48. Magnani RM, Lehnen GC, Rodrigues FB, de Sá e Souza GS, de Oliveira Andrade A, Vieira MF. Local dynamic stability and gait variability during attentional tasks in young adults. Gait Posture. 2017;55(June 2016):105-108.
49. Nestico J, Novak A, Perry SD, Mansfield A. Does increased gait variability improve stability when faced with an expected balance perturbation during treadmill walking? Gait Posture. 2021;86(March):94-100.
50. Straczkiewicz M, James P, Onnela JP. A systematic review of smartphone-based human activity recognition methods for health research. npj Digit Med. 2021;4(1):1-15.

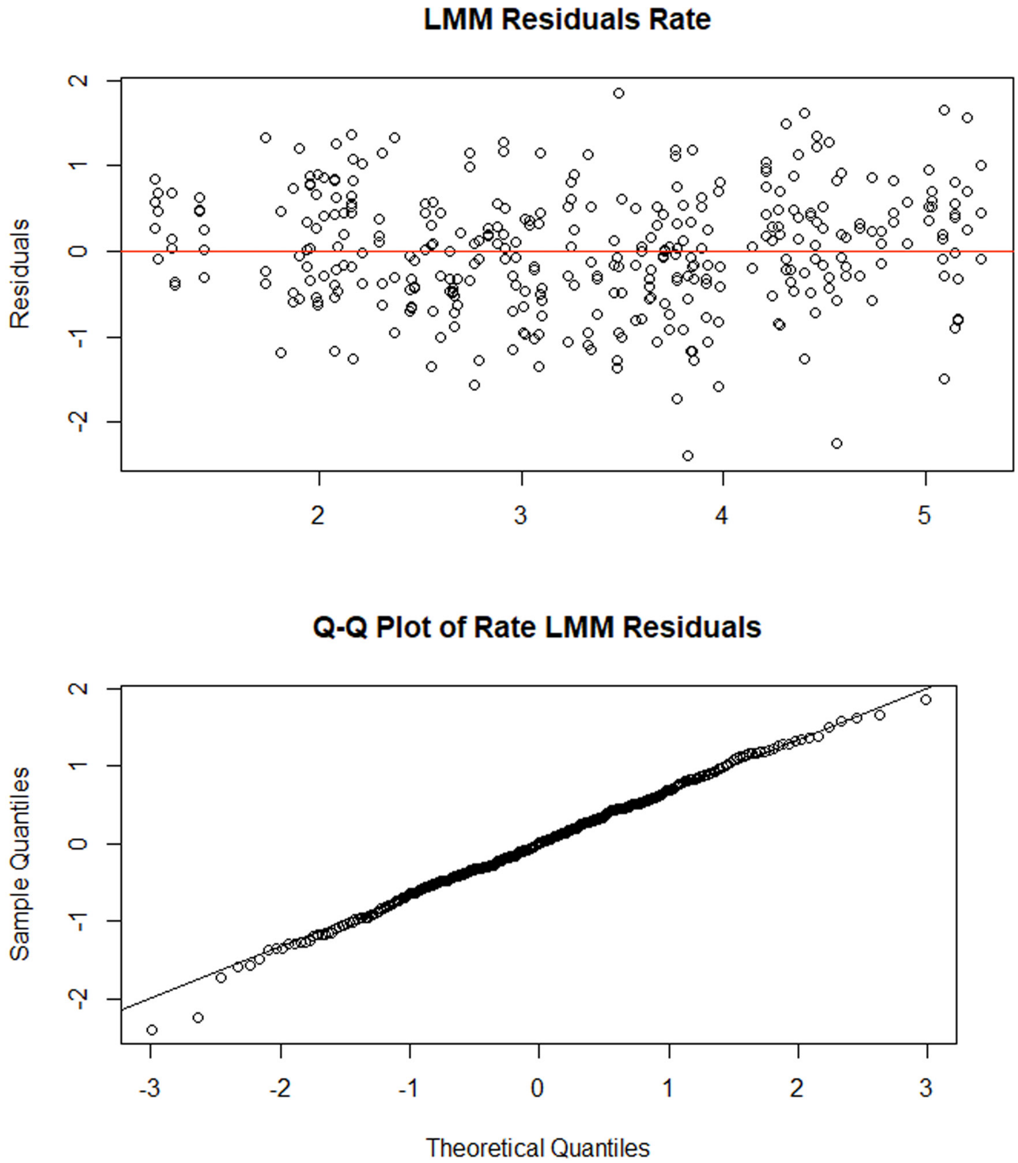
Supplementary Material



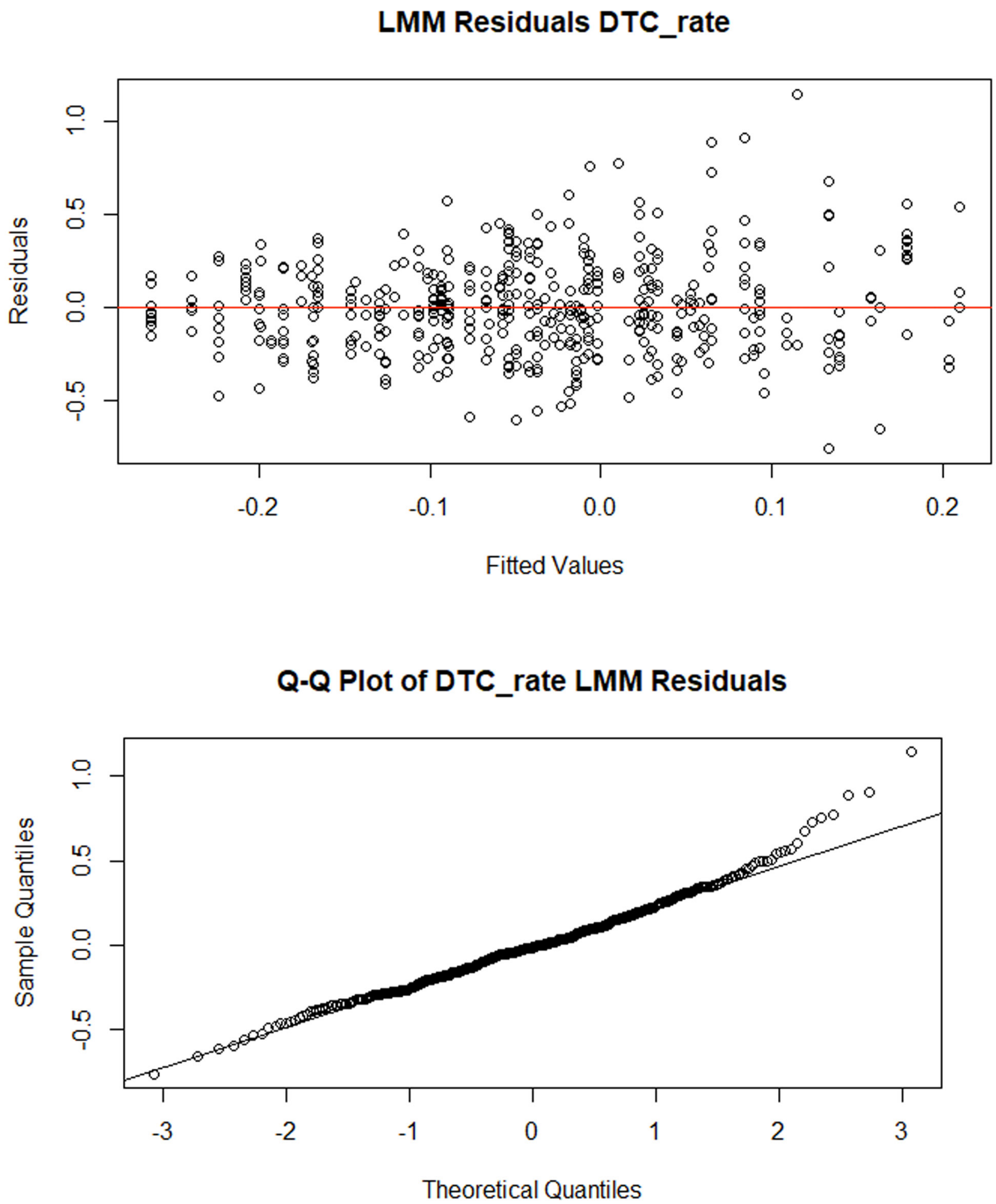
Supplementary Figure 1. Schematic of the machine learning framework for six-class classification. The proposed work used Leave-One-Subject-Out (LOSO) cross-validation as outer loop, in which one subject was iteratively held out as the test set, and the model was trained on the remaining subjects. For hyperparameter optimization of XGBoost, maximum depth of the tree was searched over [2, 10], learning rate over [0.01–0.20], subsampling ratio over [0.50–1.00], column subsampling ratio over [0.50–1.00], minimum child weight over [1–21], split regularization gamma over [0–5], L1 regularization alpha over [0–1], and L2 regularization lambda over [0.5–5.5]. Early stopping in preventing overfitting was applied during model refitting.



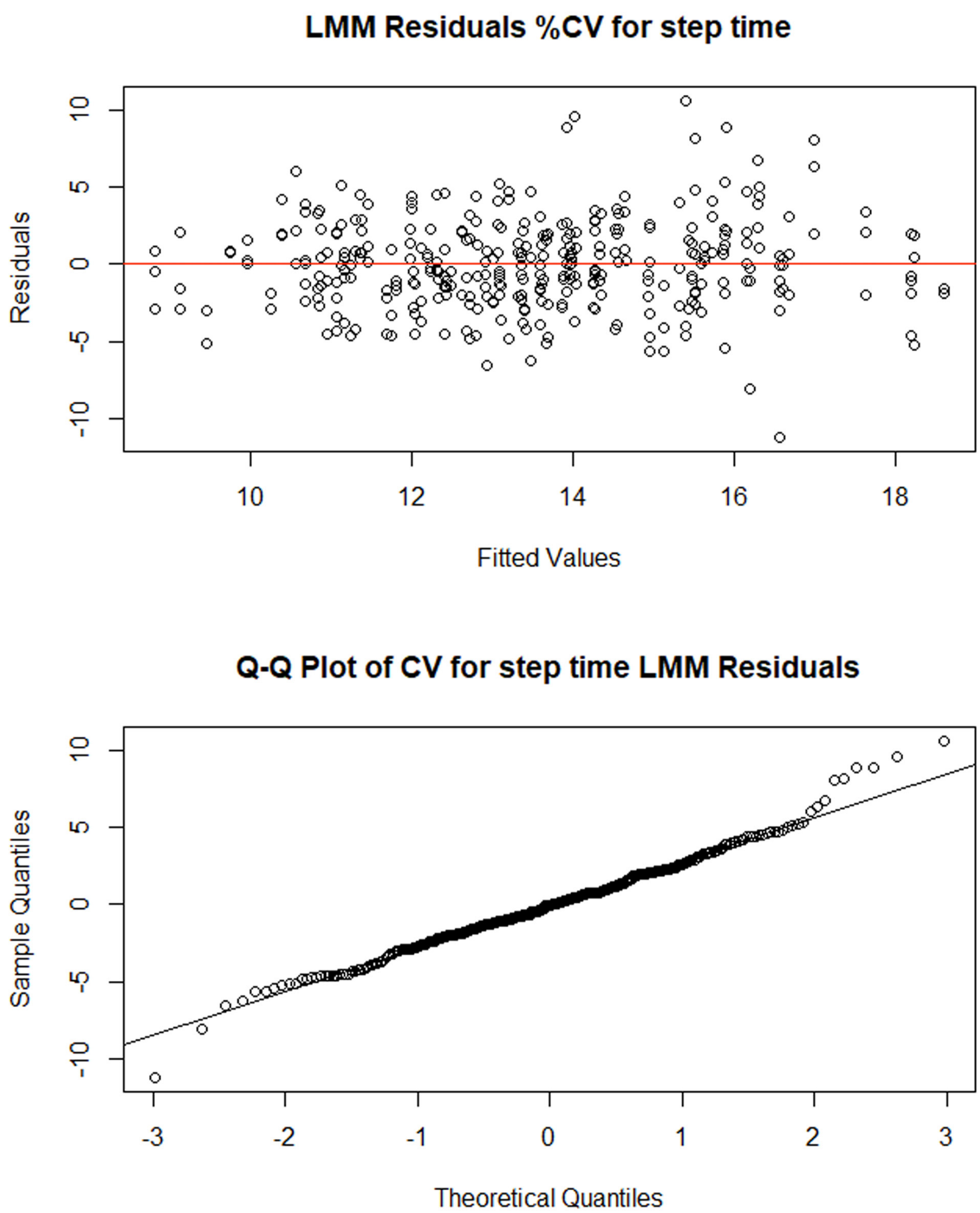
Supplementary Figure 2. XGBoost 6-Class Confusion Matrix (Pooled Leave-One-Subject-Out cross-validation, LOSO CV). Note: Task (No Task=0, Typing=1, Texting=2) and Obstacle variables (No Obstacle = 0, Obstacle = 1): L_{00}, L_{01}, L_{10}, L_{11}, L_{20}, and L_{21}.



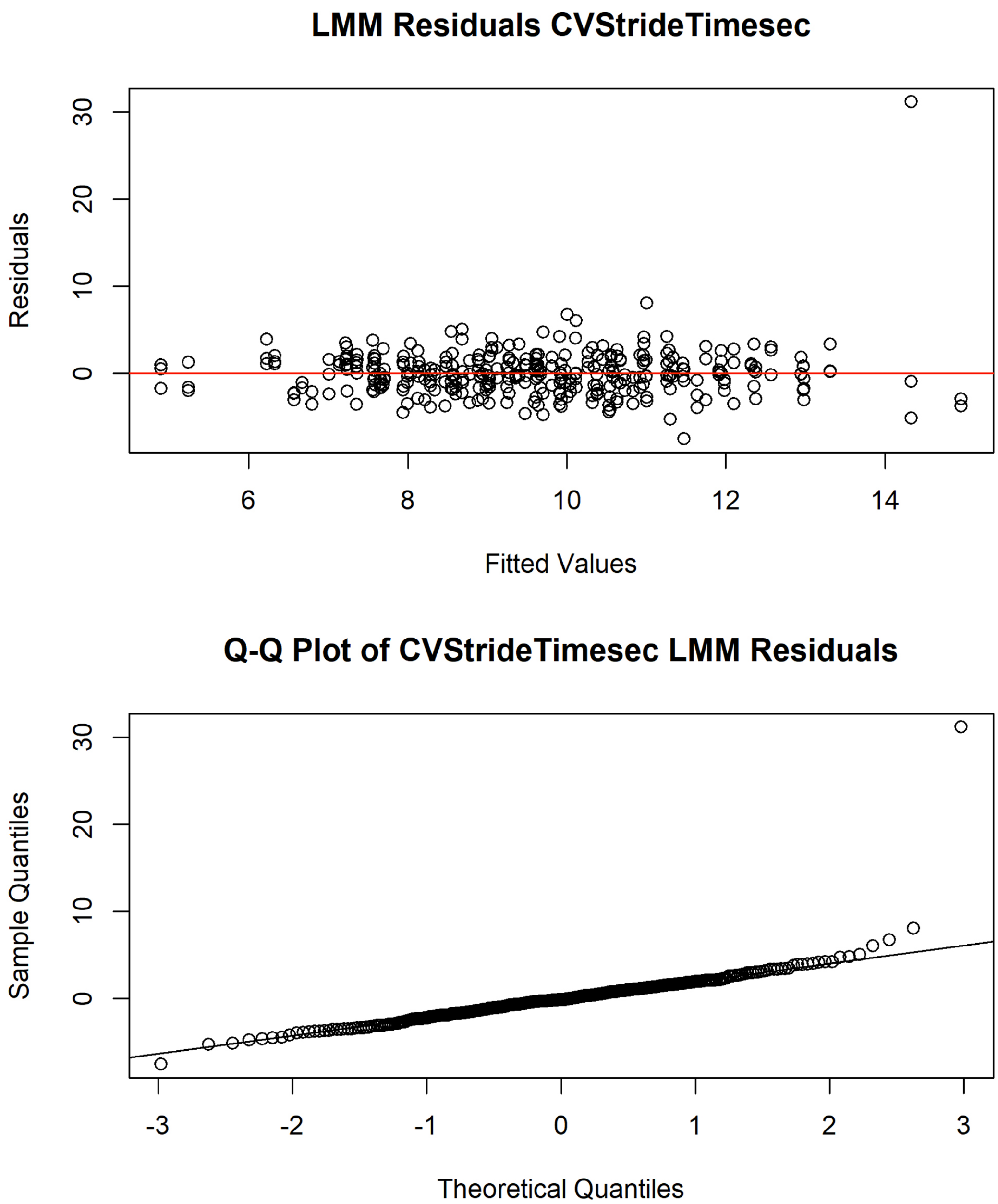
Supplementary Figure 3. Diagnostic plots for LMM Rate. The residuals vs. fitted plot for the "LMM Rate" indicates that the residuals are randomly scattered around zero with a consistent variance across the fitted values, which satisfies the assumptions of linearity and homoscedasticity. Additionally, the residuals closely follow a normal distribution, as evidenced by their alignment with the diagonal line in the Q-Q plot, with only minor deviations at the tails. These diagnostic results suggest that the model provides a good fit to the data, capturing the relationship between Rate and the predictors (Task, Obstacle, Age, average leg length, FTW, and Task:Obstacle), while effectively accounting for group-level variability through random effects.



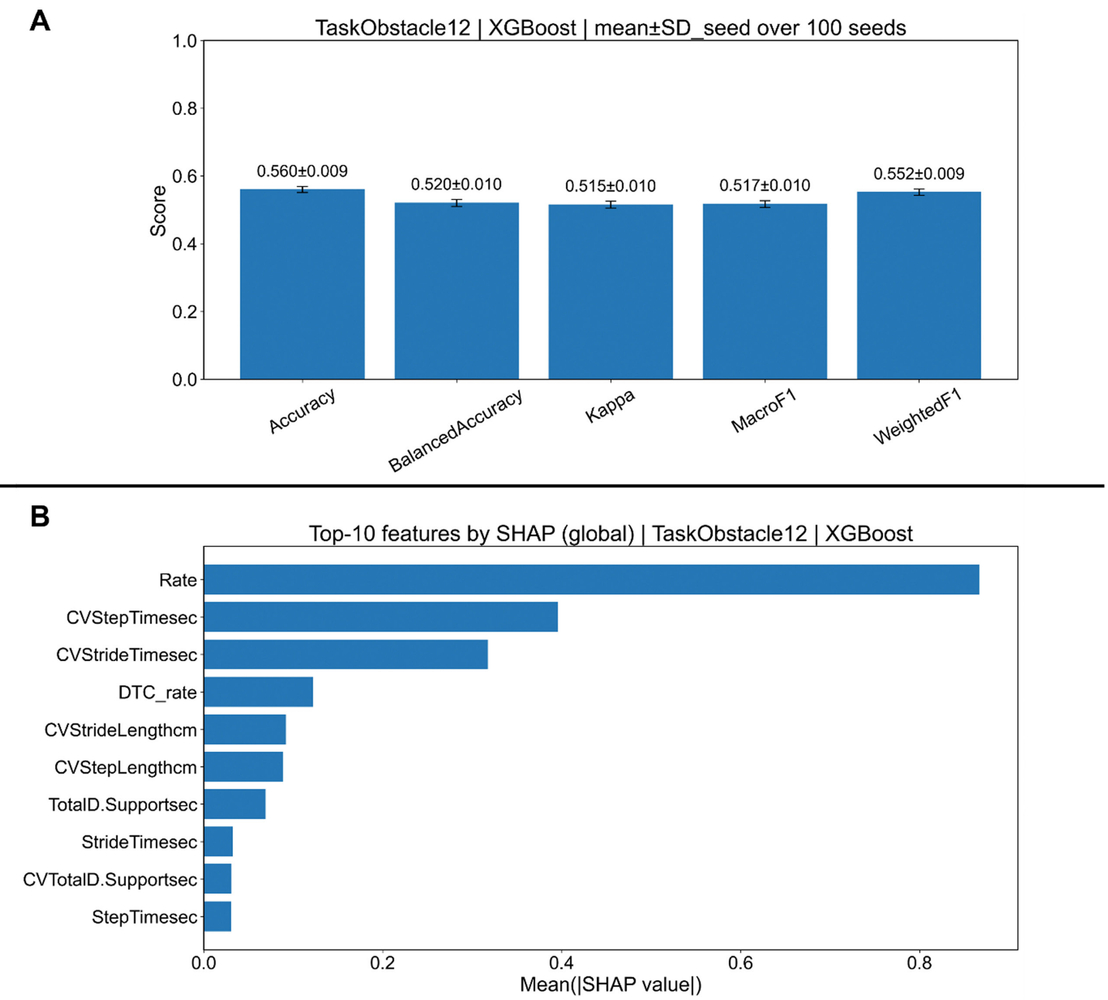
Supplementary Figure 4. Diagnostic plots for LMM DTC_rate. The residuals vs. fitted plot for the "LMM Residuals DTC_rate" shows random scatter around zero, indicating linearity and homoscedasticity. The Q-Q plot aligns well with the diagonal, suggesting approximate normality with minor tail deviations. Overall, these diagnostic plots suggest that the model effectively captures the relationship between DTC_rate and the predictors, making it a suitable choice for analysis. Minor improvements could involve addressing the slight tail deviations if the analysis demands higher precision or robustness.



Supplementary Figure 5. Diagnostic plots for LMM %CV for step time. The residuals vs. fitted plot for the "LMM %CV for step time" indicate that the residuals are randomly scattered around the horizontal zero indicate that the residuals are randomly scattered around the horizontal zero line in the residual plot, with no distinct pattern and consistent variance, meeting the assumptions of homoscedasticity and linearity. The Q-Q Plot demonstrates approximate normality, although there are minor deviations at the tails. These results suggest the model is reasonably well-specified.



Supplementary Figure 6. Diagnostic plots for LMM %CV for stride time. The residuals vs. fitted plot for the "LMM %CV for stride time" indicate that the residual plots show a random scatter of points, which suggests that the residuals are normally distributed and the models are fitting the data well without any obvious patterns or biases. The Q-Q Plot appear as straight lines with one minor tail outlier.



Supplementary Figure 7. (A) XGBoost 12-class classification performance. (B) The top ten most important features, ranked by their mean absolute SHAP values (MeanAbsSHAP) for the 12-class model. The features are shown from top to bottom in order of importance. Note: SHAP – SHapley Additive exPlanations; Rate – Rate of response (characters/sec); CVStepTimesec – Coefficient of variation (%CV) of step time; CVStrideTimesec – Coefficient of variation (%CV) of stride time; DTC_rate – Dual task cost of rate; CVStrideLengthcm – Coefficient of variation (%CV) of stride length; CVStepLengthcm – Coefficient of variation (%CV) of step length; TotalD.Supportsec – total double support time; StepTimesec – step time; ML_A – Mean mediolateral acceleration; Cadencestepmin – cadence (steps/minute).

Supplementary Table 1. The mean and standard deviation for precision, recall, and F1 per-6 class.

Label (Task)	Precision (Mean±SD)	Recall (Mean±SD)	F1 (Mean±SD)
LOO (No task)	0.929±0.006	0.932±0.008	0.930±0.005
L_{01} (Obstacle)	0.939±0.007	0.951±0.004	0.945±0.004
L_{10} (Typing)	0.761±0.015	0.834±0.017	0.796±0.013
L_{11} (Typing+Obstacle)	0.803±0.011	0.757±0.008	0.779±0.007
L_{20} (Texting)	0.778±0.020	0.685±0.021	0.728±0.017
L_{21} (Texting+Obstacle)	0.760±0.007	0.796±0.013	0.778±0.008

Note: No Task – walking without texting; Typing – copying sentences; Texting – answering questions; SD – Standard Deviation.

Supplementary Table 2. The variation in performance across XGBoost, random forest, support vector machine, and k-nearest neighbor.

	Accuracy (Mean±SD)	BalancedAcc (Mean±SD)	Kappa (Mean±SD)	Macro F1 (Mean±SD)	Weighted F1 (Mean±SD)
XGBoost	0.851±0.006	0.834±0.006	0.814±0.007	0.835±0.006	0.848±0.006
random forest	0.845±0.008	0.829±0.008	0.807±0.008	0.828±0.008	0.844±0.006
SVM	0.832±0.014	0.824±0.012	0.791±0.017	0.821±0.015	0.832±0.015
KNN	0.716±0.015	0.692±0.016	0.646±0.020	0.699±0.016	0.709±0.016

Note: Abbreviations, XGBoost – Extreme Gradient Boosting; SVM – Support Vector Machine with radial basis function kernel; KNN – K-Nearest Neighbor; BalancedAcc – Balanced Accuracy; Kappa – Cohen's Kappa Coefficient; Macro F1 – Macro averaged F1 Score; Weighted F1 – Weighted averaged F1 Score.

Supplementary Table 3. The effect size ranked by marginal R^2.

Rank	Response	Marginal R^2	Conditional R^2
1	CVStepTimesec	0.646	0.674
2	CVStrideTimesec	0.575	0.631
3	Rate	0.484	0.698
4	CVStrideLengthcm	0.444	0.502
5	CVStepLengthcm	0.440	0.507
6	Cadencestepsmin	0.419	0.704
7	StrideLengthcm	0.391	0.691
8	StepLengthcm	0.383	0.683
9	StepTimesec	0.351	0.769
10	StrideTimesec	0.300	0.752
11	Velocitycm.sec	0.299	0.744
12	StrideWidthcm	0.298	0.714
13	TotalD.Supportsec	0.194	0.344
14	CVTotalD.Supportsec	0.151	0.715
15	CVStrideWidthcm	0.102	0.749
16	ML_A	0.090	0.298
17	Accuracy	0.047	0.149
18	DTC_rate	0.032	0.205
19	DTC_accuracy	0.015	0

Note: Abbreviations, CVStepTimesec – Coefficient of variation (%CV) of step time; CVStrideTimesec – Coefficient of variation (%CV) of stride time; Rate – Rate of response (characters/sec); CVStrideLengthcm – Coefficient of variation (%CV) of stride length; CVStepLengthcm – Coefficient of variation (%CV) of step length; CVTotalD.Supportsec – coefficient of variation (%CV) of total double support time; ML_A – Mean mediolateral acceleration; Accuracy – Response accuracy (number of correct responses/total number of answers provided); Cadencestepmin – cadence (steps/minute); StrideTimesec – stride time; DTC_rate – Dual task cost of rate; DTC_accuracy – Dual task cost of accuracy; Marginal R^2 – the variance explained by the fixed effects alone Conditional R^2 – the variance explained by both fixed and random effects in the model.

Supplementary Table 4. The mean and standard deviation for each outcome variable.

Cond. Mean (SD)	Step Length (cm)	Stride Length (cm)	Stride Width (cm)	Step Time (sec)	Stride Time (sec)	TotalD. Support (sec)	Velocity (m/sec)	Cadence steps/ min	CV Step Length (cm)	CV Stride Length (cm)	CV Stride Width (cm)	CV Step Time (sec)	CV Stride Time (sec)	CV TotalD. Support (sec)	ML_A (m/ sec^2)	Rate	Accuracy	DTC_ rate	DTC_ accuracy
No task	69.7 (6.2)	136.4 (12.3)	10.4 (2.2)	0.56 (0.02)	1.11 (0.04)	0.32 (0.04)	1.26 (0.12)	106.6 (4.02)	2.65 (0.71)	1.50 (0.55)	17.39 (6.31)	2.64 (1.12)	1.39 (0.45)	3.75 (1.21)	0.10 (0.03)				
Low	73.4 (7.7)	147.2 (15.7)	10.4 (2.4)	0.56 (0.03)	1.14 (0.05)	0.28 (0.04)	1.30 (0.16)	106.5 (5.09)	6.07 (2.19)	4.08 (1.74)	18.28 (7.54)	7.07 (3.23)	4.40 (2.23)	6.67 (2.81)	0.11 (0.03)				
Medium	73.6 (7.2)	146.9 (14.5)	10.6 (2.4)	0.58 (0.02)	1.16 (0.05)	0.28 (0.03)	1.27 (0.14)	104.8 (4.37)	8.29 (3.05)	5.91 (2.37)	20.76 (9.37)	9.96 (2.71)	6.32 (2.09)	6.98 (2.64)	0.11 (0.02)				
High	74.3 (7.2)	148.4 (14.8)	11.4 (2.8)	0.6 (0.04)	1.21 (0.08)	0.28 (0.04)	1.23 (0.16)	101.5 (5.94)	7.84 (3.00)	5.45 (2.10)	19.71 (6.36)	14.31 (4.18)	9.06 (3.18)	6.84 (2.63)	0.12 (0.02)				
Typing	63.6 (4.8)	127.2 (9.4)	10.7 (3.2)	0.59 (0.04)	1.19 (0.08)	0.36 (0.05)	1.07 (0.12)	101.5 (6.62)	3.39 (0.89)	1.96 (0.85)	22.14 (10.6)	3.32 (1.25)	1.87 (0.54)	3.57 (1.27)	0.09 (0.17)	4.64 (1.1)	0.97 (0.03)	-0.02 (0.13)	-0.01 (0.03)
Typing Low	65.9 (5.2)	131.8 (10.1)	11.2 (3.1)	0.63 (0.06)	1.27 (0.11)	0.34 (0.05)	1.05 (0.13)	95.79 (8.58)	10.65 (3.72)	7.68 (2.58)	24.93 (8.56)	11.65 (3.01)	7.96 (1.84)	7.82 (2.77)	0.10 (0.15)	4.33 (1.0)	0.97 (0.03)	-0.08 (0.15)	-0.03 (0.08)
Typing Medium	65.6 (6.2)	131.2 (12.3)	11.1 (3.1)	0.64 (0.06)	1.28 (0.12)	0.34 (0.06)	1.04 (0.15)	95.72 (7.90)	10.13 (3.62)	7.38 (2.61)	23.61 (9.00)	13.84 (3.75)	8.32 (1.96)	7.10 (2.73)	0.10 (0.15)	2.48 (0.7)	0.99 (0.02)	-0.08 (0.27)	-0.02 (0.13)
Typing High	67.4 (6.3)	134.7 (12.8)	11.0 (3.3)	0.64 (0.06)	1.29 (0.13)	0.33 (0.05)	1.05 (0.14)	94.78 (8.60)	10.45 (3.24)	7.80 (2.10)	22.57 (8.13)	14.14 (3.23)	9.74 (2.29)	8.91 (5.09)	0.10 (0.19)	4.24 (1.0)	0.97 (0.03)	-0.09 (0.13)	-0.02 (0.08)
Texting	62.5 (5.4)	125.3 (10.4)	11.2 (3.2)	0.59 (0.04)	1.18 (0.08)	0.36 (0.04)	1.07 (0.12)	101.8 (6.29)	3.35 (1.34)	2.11 (1.02)	18.72 (7.23)	3.69 (1.77)	2.29 (1.06)	4.43 (1.99)	0.09 (0.19)	2.71 (0.7)	0.98 (0.03)	0 (0.28)	-0.02 (0.02)
Texting Low	67.13 (6.5)	134.5 (13.1)	11.4 (3.2)	0.65 (0.07)	1.31 (0.13)	0.33 (0.06)	1.04 (0.14)	94.13 (8.68)	10.55 (2.60)	7.71 (2.17)	22.08 (8.80)	11.92 (3.11)	9.34 (1.62)	7.91 (2.48)	0.11 (0.21)	2.40 (0.6)	0.98 (0.04)	-0.07 (0.15)	-0.02 (0.03)
Texting Medium	66.7 (7.1)	134.1 (14.1)	11.5 (3.4)	0.65 (0.06)	1.32 (0.15)	0.33 (0.06)	1.02 (0.15)	93.40 (8.58)	10.35 (4.02)	7.90 (3.68)	27.68 (11.5)	13.47 (3.13)	10.7 (3.42)	8.54 (2.90)	0.10 (0.14)	4.20 (0.9)	0.97 (0.02)	-0.10 (0.26)	-0.01 (0.03)
Texting High	66.4 (6.6)	132.8 (13.2)	11.3 (3.4)	0.66 (0.06)	1.32 (0.12)	0.33 (0.05)	1.01 (0.15)	93.13 (7.90)	10.59 (3.75)	7.70 (2.81)	27.02 (10.3)	16.15 (3.97)	11.27 (2.04)	7.65 (2.51)	0.10 (0.16)	2.63 (0.6)	0.98 (0.03)	-0.02 (0.22)	-0.01 (0.02)

Note: Abbreviations, Cond. – Condition; No Task – walking without texting; Typing – copying sentences; Texting – answering questions; SD – Standard Deviation; Low – Low Obstacle; Medium – Medium Obstacle; High – High Obstacle; TotalD.Supportsec –total double limb support time; CVStepTimesec – Coefficient of variation (%CV) of step time; CVStrideTimesec – Coefficient of variation (%CV) of stride time; CVStrideLengthcm – Coefficient of variation (%CV) of stride length; CVStepLengthcm – Coefficient of variation (%CV) of step length; CVTotalD.Supportsec – coefficient of variation (%CV) of total double limb support time; ML_A – Mean mediolateral acceleration; Rate – Rate of response (characters/sec); Accuracy – Response accuracy (number of correct responses/total number of answers provided); DTC_rate – Dual task cost of rate; DTC_accuracy – Dual task cost of accuracy.