scientific reports

OPEN

Evaluating eye-hand coordination with digital technologies

Francesca Di Rocco[1,2], Marianna De Maio[1,2], Emanuel Festino[1,2], Olga Papale[1,2], Philip X. Fuchs[3], Rubens A. Da Silva[4], Alexis Barbeau[1,2], Cristina Cortis[1,2]✉ & Andrea Fusco[5]

Digital tools, such as computerized wobble boards (WB), offer a novel approach to assess dynamic balance in lower limbs. Since their potential for evaluating eye-hand coordination remains unexplored, this study assessed the inter-test reliability and concurrent validity of WB measurements for upper limb fine motor skills, proposing WB as an alternative to the Grooved Pegboard test (GPT). Fifty-three healthy participants completed WB and GPT tests, and a WB retest after 48 h. The custom WB software displayed real-time performance via a motion marker and target zone. Participants moved the marker within the target zone following predefined patterns (clockwise, counterclockwise, anteroposterior, mediolateral) across four 15-s trials per hand. Performance was quantified as the duration (s) the marker remained in the target zone under each condition. According to Intraclass correlation coefficients (ICC) WB demonstrated good to excellent reliability (ICC: 0.62–0.80), acceptable Standard Error of Measurement (*SEM*: 0.96–2.14 s), and minimal detectable change (MDC_{95}: 1.90–4.25 s). Moderate to strong correlations ($r = -0.30$ to -0.54) between WB and GPT outcomes suggested WB captured related aspects of fine motor coordination. These findings confirm WB's reliability and validity as a tool for assessing eye-hand coordination. Further validation is needed in training or rehabilitation contexts.

Keywords Neuromuscular performance, Motor skills, Validity, Reliability, Upper limb coordination

Motor skills are fundamental to daily activities, such as walking, grasping, cooking, work-related tasks, and driving[1–3]. These skills are categorized into gross and fine motor skills[4], based on movement patterns[3] and task complexity[5]. Fine motor skills, unlike gross motor skills, involve precise and coordinated movements of smaller muscle groups, primarily in the hands and fingers[6]. The literature[7,8] further distinguishes related concepts such as fine motor precision, visual-motor integration, and manual dexterity. Fine motor precision refers to the accurate control of hand and finger movements while visual-motor integration involves responding to visual stimuli with appropriate motor action, and manual dexterity relates to object manipulation and coordination[6,9].

Eye-hand coordination refers to a sensorimotor ability in which the brain integrates visual and tactile information to guide deliberate movements. This skill can be categorized into two types: probation, or closed motor skills, where the movement is self-initiated and controlled by the individual; and reaction, or open motor skills, which are triggered in response to external stimuli[10]. This coordination ability involves the visual system's capacity to process input from the eyes and translate it into precise motor commands that guide hand actions toward a specific goal. Enhancing eye-hand coordination typically involves tasks that challenge the integration of visual perception and motor control. In particular, interacting with unstable surfaces, or manipulating small objects in dynamic contexts like video gaming, can effectively support the development of fine motor control[11]. Moreover, the effectiveness of these tasks depends on the ability to align visual attention with accurate hand execution, a process during which manual dexterity plays a crucial role[12].

To assess fine motor skills, various tests have been developed[13–16], with the Grooved Pegboard test (GPT) widely recognized as the gold standard for evaluating motor coordination and manual dexterity[17,18]. The GPT is valued for its precision, ease of administration across diverse populations, including children[19], adults[18], the elderly[20], and individuals with neurological conditions[21], and its high test–retest reliability[22]. However, GPT primarily assesses static motor skills and may not fully capture the dynamic coordination required in everyday tasks[23]. This limitation is particularly relevant in conditions such as Parkinson's disease and older adults, where

[1]Department of Human Sciences, Society and Health, University of Cassino and Lazio Meridionale, Via S. Angelo–Località Folcara, 03043 Cassino, FR, Italy. [2]European University of Technology EUt+, Viale Dell'Università - Rettorato, 03043 Cassino, FR, Italy. [3]Department of Physical Education and Sport Sciences, National Taiwan Normal University, 162, Section 1, Heping E. Rd., Taipei City 106, Taiwan. [4]Department of Health Sciences, Physical Therapy Program and Research Laboratory BioNR, Université du Québec à Chicoutimi, 555 Bd de l'Université, Chicoutimi, QC G7H 2B1, Canada. [5]Department of Medicine and Aging Sciences, University of "G. D'Annunzio" Chieti-Pescara, 66100 Chieti, Italy. ✉email: c.cortis@unicas.it

GPT has been shown to predict cognitive impairments rather than purely motor deficits[21]. This limitation arises because GPT performance is influenced not only by fine motor execution, but also cognitive domains such as attention, visuospatial processing, executive function, and processing speed[24,25]. In populations with neurological disorders or in aging, deficits in these cognitive domains can significantly impact GPT scores, making it a sensitive tool for detecting cognitive impairments rather than being a pure measure of motor function[21]. As a result, relying solely on the GPT may obscure motor deficits that are independent of cognitive status, highlighting the need for alternative or complementary tools that specifically target dynamic aspects of motor coordination with reduced cognitive confounding effects. These findings further highlight the need for alternative tools able to assess fine motor coordination under dynamic conditions, better reflecting real-world functional movements.

Computerized proprioceptive platforms such as wobble boards (WB) offer a novel approach to motor skill assessment. Traditionally used for postural control and balance assessment,[26,27] WB can be adapted for upper-limb coordination evaluations by integrating hand-based control tasks. Unlike GPT, which operates in a stable environment, WB protocols introduce dynamic control elements representing an additional aspect of real-world motor challenges, particularly those requiring continuous adaptation to unpredictable conditions (e.g., across diverse daily tasks and sports scenarios). The emphasis on dynamic tasks highlights the importance of assessing not only static postural control, but also the ability to adapt and react to unpredictable conditions, crucial for functional motor performance[28]. Linear movements, such as anteroposterior or mediolateral, are typically repetitive and rely heavily on feedback control, where real-time sensory information (visual or proprioceptive) is used to adjust ongoing movement. In contrast, circular tasks involve more complex trajectories that require anticipatory motor planning, making greater use of feedforward control mechanisms. These movements demand predictive coordination based on internal models rather than continuous sensory correction. Therefore, circular WB tasks may impose higher neuromuscular and cognitive demands than linear movements, offering a more sensitive assessment of fine motor control and adaptability[29,30]. Furthermore, while both GPT and WB tests use task completion time as a primary outcome, WB incorporate advanced sensors (e.g., accelerometers) that enable precise measurement of movement parameters such as acceleration, angles, and movement smoothness, by providing a more comprehensive assessment of motor performance.

While previous studies have focused on the application of WBs for dynamic balance and postural control,[26,27] their use in upper-limb sensorimotor evaluation remains largely unexplored. Only preliminary information is available on computerized WB system to assess upper-limb control via visual biofeedback,[31] and to the best of our knowledge, no empirical studies have yet evaluated the reliability or validity of similar approach in healthy or clinical populations. Therefore, this study aimed to (i) determine the inter-test reliability of WB outcomes in evaluating upper-limb fine motor skill and (ii) test the concurrent validity of WB performance compared to the GPT. We hypothesized that the WB protocol would yield reliable fine motor skill measurements and offer additional insight into execution speed and hand–eye coordination, complementing the limitations of GPT in assessing dynamic motor tasks.

Methods

Study design

This study assessed the validity and reliability of WB as a measure of manual dexterity through two test sessions. In the first session, participants completed a WB test followed by a GPT with 2-min rest interval in between. During the second session, conducted five days later to avoid learning effects[32], participants repeated the WB test. This interval was based on previous motor control studies using similar upper-limb tasks[32], where 3–7 days intervals were reported to be adequate to reduce performance transfer in repeated motor assessments. A trained evaluator administered both tests to ensure standardization. Intra-individual, intra-session comparison between WB and GPT performances assessed concurrent validity, while an intra-individual, inter-session comparison between repeated WB tests determined test–retest reliability.

Participants

The required sample size for test–retest reliability analysis was determined via G*power software (version 3.1 for Mac OS, Heinrich Heine Universität Düsseldorf, Kent) for Intraclass Correlation Coefficient (ICC) analysis. The minimum acceptable reliability (ICC ρ0) was set at 0.5, with an expected reliability (ICC ρ1) of 0.75[33]. The analysis indicated a minimum sample size of 45 participants for a two-tailed test with a significance level (α) of 0.05 and a desired power of 80%[34]. Inclusion criteria were the absence of the following conditions: upper limb fractures and surgeries, visual disorders, neuromuscular diseases, and cognitive impairment. Participants unavailable for the WB retest or with missing data were excluded. Of the 65 participants initially assessed, five were excluded due to unavailability for the WB retest, and seven due to missing data, resulting in a final sample of 53 students (aged 21–34 years) from the University of Cassino and Lazio Meridionale. The STROBE cross sectional reporting guidelines were used[35]. Anthropometric data were collected before testing. Body mass and height were measured using a scale with an integrated stadiometer (Seca, model 709, Vogel & Halke, Hamburg, Germany; accuracy: ± 0.1 kg and ± 0.1 cm). Body mass index (BMI) was subsequently calculated, and preferred hand was determined based on the participant's writing hand,[36] following LICI recommendations[37].

Wobble board test

The WB (WSP, Well Sport Project, G.S.J. Services S.r.l., Rome, Italy) featured a triaxial accelerometer (Phidget Spatial 0/0/3 Basic 1041, Phidgets Inc. 2016, Calgary, AB, Canada) and consisted of a circular standing surface (diameter: 40 cm, height: 2 cm) placed on a semi-spherical base (diameter: 12 cm, height: 6 cm). A USB connection (200 Hz) linked the WB to a laptop, where customized software displayed real-time performance via a motion marker on a monitor (screen resolution: 1920 × 1080) (Supplementary S1). Participants controlled the WB tilt angle and, consequently, the motion marker position using hand movement.

The WB test required participants to keep the motion marker (yellow circle; diameter: 6 mm) within a target zone (red circle; diameter: 6.5 cm) for as long as possible. The software recorded full seconds within the target zone as the performance measure. The target zone followed different movement patterns:

- *Clockwise*: The target zone moved from the middle of the screen to the left in 6 ms before completing a full clockwise rotation.
- *Counterclockwise*: The target zone moved from the middle of the screen to the left in 6 ms before completing a full counterclockwise rotation.
- *Anteroposterior*: The target zone moved from the middle of the screen to the top in 8 ms before continuing along a vertical trajectory.
- *Mediolateral*: The target zone moved from the middle of the screen to the top in 8 ms before continuing along a horizontal trajectory.

For anteroposterior and mediolateral patterns, the target zone's movement speed was adjusted based on the participant's performance. In particular, reaching the target zone in the middle in the beginning of a trial and keeping the marker in the zone during the trial increased the target zone speed. In clockwise and counterclockwise patterns, the target zone followed a predetermined trajectory independent of participant movement. Specifically, in the anteroposterior and mediolateral conditions, the speed of the target zone was not fixed but dynamically adjusted based on the participant's ability to keep the motion marker within the target zone. If the participant maintained the motion marker inside the target zone for a continuous duration of 2 ms, the software increased the target zone's movement speed by 10% in real time. Conversely, if the motion marker moved outside the target zone for more than 1 ms, the speed returned to baseline. This adaptive algorithm, embedded in the proprietary and closed-source software, was designed to automatically adjust task difficulty based on individual performance and skill level. As it cannot be modified by the user, the system dynamically regulates target zone speed in real time, helping to prevent ceiling effects and improving the WB's sensitivity in detecting subtle variations in motor control. After a 1-min familiarization with all movement patterns, participants completed one 15-s trial per pattern for each hand (8 trials total), with 30-s recovery between trials. The sequence of patterns and hands was randomized. The WB was placed on a table, with a monitor positioned in front of the participants, which, standing upright, placed the tested hand on the WB, maintaining a 90° elbow angle and a relaxed shoulder, while the contralateral arm rested at their side (Fig. 1). Trials were repeated if participants (a) did not maintain the contralateral arm position, (b) use the contralateral arm for support, or (c) spoke during the test. Out of 424 total WB trials (53 participants × 4 conditions × 2 limbs), 26 trials (6.1%) were considered invalid and had to be repeated. Most repetitions were due to participants speaking during the test (n = 15), followed by incorrect contralateral arm positioning (n = 7) and use of the contralateral arm for support (n = 4).

Grooved Pegboard test

The GPT (model 32025, Lafayette Instrument Company, Inc., Lafayette, Indiana, USA) consisted of a 10 × 10 cm square pegboard with 25 holes with random keyhole orientation arranged in a 5 × 5 grid. Steel pegs (diameter: 0.4 cm; model 32104) with side keys were in a spherical tray above the keyholes. After a short demonstration and verbal instructions, participants familiarized themselves with the task by inserting pegs into the top row[18]. Participants were then instructed to insert pegs sequentially, as quickly as possible, completing rows from left to right for the right hand and from right to left for the left hand, moving from top to bottom. A 1-min recovery period was allowed between trials. Participants began the test whenever they were ready. Recording time started when the first peg was picked up and ended when the last peg was inserted. A digital stopwatch (Garmin International, Kansas City, MO, USA) recorded the completion time[38]. If a peg fell, participants had to continue without retrieving it. The GPT was placed on a table, and participants were seated with the non-testing hand resting on the table. Each evaluation lasted approximately 30 min, including familiarization and both tests (WB and GPT).

Statistical analysis

Dependent variables included GPT completion time (s) and WB motion marker time (s) within the target zone. Data normality was verified via Shapiro–Wilk test and means with standard deviations (SD) were calculated.

A 2-way mixed model evaluated intersession reliability, treating participants as random effect and measurement tools as a fixed effect[39]. $ICC_{2,1}$ was used to assess absolute agreement between measurements and reported in Table 1. ICC values, together with 95% confidence intervals, were interpreted as poor (0.00–0.39), fair (0.40–0.59), good (0.60–0.74), and excellent (0.75–1.00)[39]. Standard Error of Measurement (*SEM*) and the Minimal Detectable Change (MDC_{95}) were also computed. *SEM* estimates measurement error and within-subject variability ($SEM = SD * \sqrt{1 - ICC}$), providing a range of values within which a true score might fall based on an observed score[39]. MDC_{95} indicates the smallest detectable change required to exhibit an actual change in performance ($MDC_{95} = SEM * \sqrt{2} * 1.96$), suggesting that changes equal to or exceeding MDC_{95} indicate reliable and significant change at the 95% confidence level. Bland–Altman plots and 95% Limits of Agreement (95% LoA = mean ± 2 * SD) assessed intersession variability between WB test-retests[40]. Bland–Altman analysis was not conducted between WB test and GPT test because they are based on different measurement properties and scoring directions (i.e., higher WB scores indicate better performance, whereas lower GPT scores indicate better performance). In our study, WB and GPT were not considered as interchangeable tools; rather, they have been considered complementary assessments that capture overlapping but distinct aspects of upper-limb motor control.

Fig. 1. Starting position of the participant during the wobble board test.

Task		Limb	Test (s)	Retest (s)	ICC (95% CI)	*SEM* (s)	MDC_{95} (s)
WB tests	Anteroposterior	Preferred	12.4 ± 1.8	13.2 ± 1.7	0.52 (0.18–0.72)	1.24	2.45
		Non-preferred	12.2 ± 1.8	12.8 ± 2.0	0.74 (0.55–0.85)	0.96	1.90
		Average	12.3 ± 1.4	13.0 ± 1.7	0.74 (0.56–0.85)	0.75	1.50
	Mediolateral	Preferred	11.0 ± 1.8	12.6 ± 1.9	0.62 (0.35–0.78)	1.15	2.28
		Non-preferred	11.3 ± 2.1	12.0 ± 1.7	0.62 (0.34–0.78)	1.30	2.57
		Average	11.1 ± 1.6	12.3 ± 1.6	0.72 (0.52–0.84)	0.88	1.74
	Clockwise	Preferred	9.9 ± 3.8	10.0 ± 3.4	0.80 (0.67–0.89)	1.77	3.52
		Non-preferred	10.4 ± 3.4	11.8 ± 3.0	0.78 (0.62–0.87)	1.63	3.24
		Average	10.1 ± 3.4	11.3 ± 2.8	0.83 (0.71–0.90)	1.28	1.54
	Counterclockwise	Preferred	10.3 ± 3.4	11.6 ± 2.9	0.80 (0.65–0.88)	1.53	3.03
		Non-preferred	9.0 ± 4.0	11.0 ± 3.3	0.72 (0.52–0.84)	2.14	4.25
		Average	9.7 ± 3.3	11.3 ± 2.8	0.82 (0.69–0.89)	1.37	2.71
	Mean	Preferred	43.7 ± 8.4	48.4 ± 7.5	0.84 (0.72–0.90)	0.84	1.67
		Non-preferred	43.0 ± 9.3	47.8 ± 7.8	0.83 (0.72–0.90)	0.96	1.91
		Average	43.3 ± 8.4	48.1 ± 7.3	0.86 (0.76–0.92)	0.78	1.56

Table 1. Mean values ± standard deviations (SD) of wobble board (WB) test, and WB retest performances in seconds (s) for all conditions and limbs. WB test–retest intraclass correlation coefficient (ICC) with 95% confidence interval (95% CI), standard error of measurement (*SEM*) in seconds (s), and minimal detectable change (MDC_{95}) in seconds (s).

Pearson correlation coefficients evaluated concurrent validity between WB and GPT outcomes. Correlations were classified as small (< 0.25), fair (0.25–0.49), moderate (0.5–0.74), and strong (≥ 0.75)[34]. Only the first WB session was used for this comparison to ensure consistency with GPT data. Statistical analyses were conducted using STATA software (version 18, StataCorp LP, College Station, TX, USA) with significance set at $p \leq 0.05$.

Ethics

This study was conducted following the guidelines of the Declaration of Helsinki. The Institutional Review Board of the Department of Human Sciences, Society, and Health of the University of Cassino and Lazio Meridionale approved this study (approval No.: 11748; date: 09 June 2021). Participants provided written informed consent after receiving a full explanation of study procedures.

Results

A total of 23 female and 30 male students met the inclusion criteria and participated in the study (age: 25.0 ± 2.8 years, body height: 1.69 ± 0.09 m, body mass: 67.7 ± 13.3 kg, BMI: 23.2 ± 3.1 $kg \cdot m^{-2}$). Participants completed GPT with the preferred hand in 63.7 ± 9.3 s, with the non-preferred hand in 69.0 ± 12.7 s, and in 66.4 ± 9.9 s as the average values of both hands.

The WB test demonstrated intersession reliability coefficients ranging from fair to excellent, with the highest values observed in clockwise and counterclockwise tasks and when averaging across tasks and limbs. Specifically, ICC values exceeded 0.80 for preferred and non-preferred circular patterns, suggesting high reliability. Conversely, linear patterns (anteroposterior and mediolateral) showed slightly lower reliability although within acceptable ranges. *SEM* and MDC_{95} were the lowest when performance data were averaged across conditions, further supporting the robustness of WB outcomes in tracking repeatable motor performance. Table 1 shows mean values ± SD of WB test, and WB retest outcomes, along with WB intersession ICC, 95% CI, *SEM*, and MDC_{95}. Intersession reliability ranged from 0.52 to 0.86 (ICC), 0.75 to 2.14 s (*SEM*), and 1.50 to 4.25 s (MDC_{95}). Test–retest differences in WB performance plotted against the mean, including the 95% LoA for each condition (anteroposterior, mediolateral, clockwise, counterclockwise, and the average) and limb (preferred and non-preferred) are shown in Fig. 2, Fig. 3, Fig. 4, and Fig. 5.

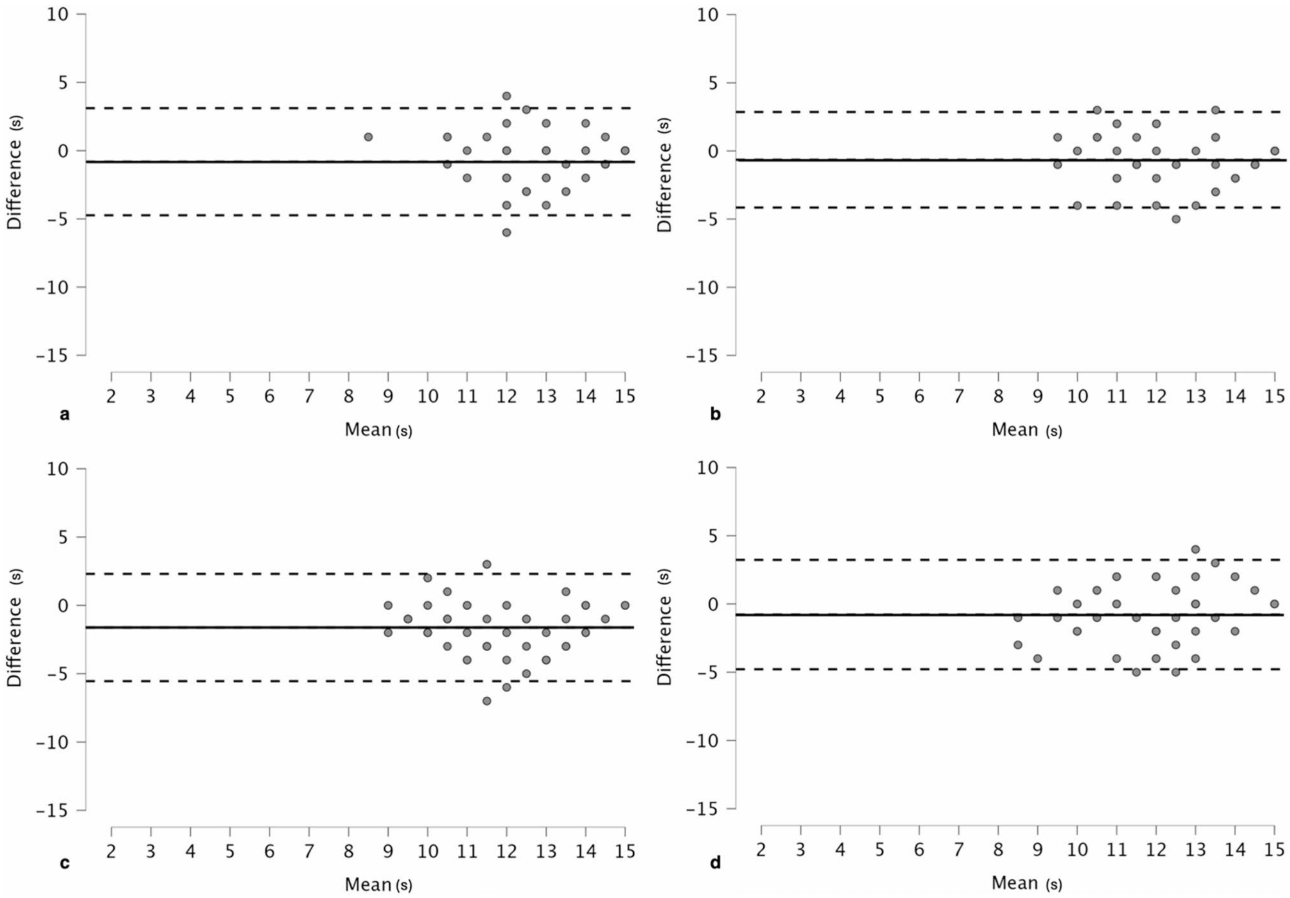

Fig. 2. Bland–Altman plot for: (**a**) preferred limb, (**b**) non-preferred limb in wobble board anteroposterior condition; (**c**) preferred limb, (**d**) non-preferred limb, in wobble board mediolateral condition. Single bold lines indicate the average of the differences, whereas two dotted lines indicate the 95% limit of agreement.

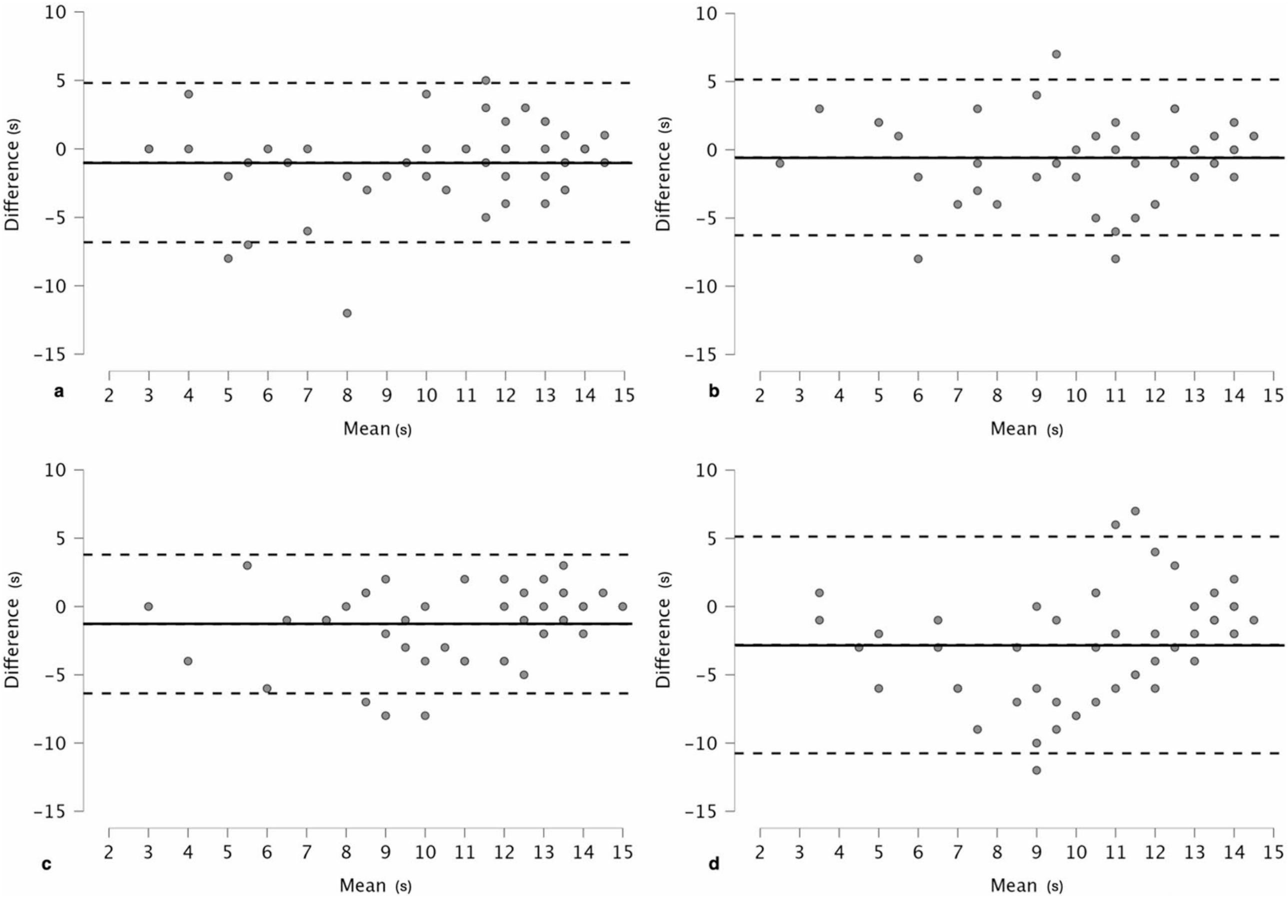

Fig. 3. Bland–Altman plot for: (**a**) preferred limb, (**b**) non-preferred limb in wobble board clockwise condition; (**c**) preferred limb, (**d**) non-preferred limb in wobble board counterclockwise condition. Single bold lines indicate the average of the differences, whereas two dotted lines indicate the 95% limit of agreement.

Regarding validity, Pearson correlation coefficients between WB and GPT scores ranged from − 0.26 to − 0.54, with stronger inverse correlations observed in non-preferred limb and circular movement tasks. This suggests that WB performance is meaningfully aligned with GPT outcomes, particularly in tasks that demand greater neuromuscular control. The strongest correlation was recorded in the non-preferred limb during the clockwise condition ($r = -0.52, p < 0.0001$), highlighting the relevance of task complexity and limb usage. Overall, aggregated WB performances (i.e., task and limb averages) showed the most consistent and interpretable relationships with GPT results. Table 2 shows the Pearson correlation coefficients between WB and GPT performances.

Discussion

This study aimed to determine the inter-test reliability of WB values during upper limb fine motor tasks and assess WB's concurrent validity with GPT by examining the relationship between the two measures within the same study design. The key findings were: the computerized WB demonstrated fair to excellent reliability (ICC: 0.52–0.86), with the highest values in the averaged WB tasks (ICC: 0.83–0.86), particularly when averaging both limbs (ICC: 0.86); significant correlations, ranging from − 0.26 to − 0.54 (fair to moderate), emerged between WB and GPT scores, with stronger correlations under specific conditions. These findings support WB's fair to excellent test–retest reliability in assessing upper limb fine motor skills, considering the diverse movement patterns involved. Specifically, clockwise and counterclockwise patterns exhibited good to excellent reliability (ICC: 0.72–0.83), reflecting acceptable consistency and absolute agreement across sessions. This aligns with previous studies suggesting that task complexity and variability influence reliability. In this study, linear movements (anteroposterior and mediolateral) were likely more repetitive and less complex than circular movements, potentially explaining the variability in reliability across conditions. The highest reliability was observed in the WB mean variable, consistent with findings suggesting that combining multiple measures or movement types reduces variability and enhances reliability[18,41].

Low *SEM* values further confirmed WB's precision in tracking motor coordination across sessions. The best *SEM* and MDC_{95} values (i.e., lowest average) were found when averaging across WB tasks, reinforcing WB's effectiveness in detecting motor skill improvements or deficits with high confidence. Compared to studies on fine motor coordination and visuomotor integration, often reporting high ICC values (0.80–0.95) in controlled settings[40], the WB achieved comparable reliability (ICC: 0.52–0.86). WB's ability to integrate linear and circular

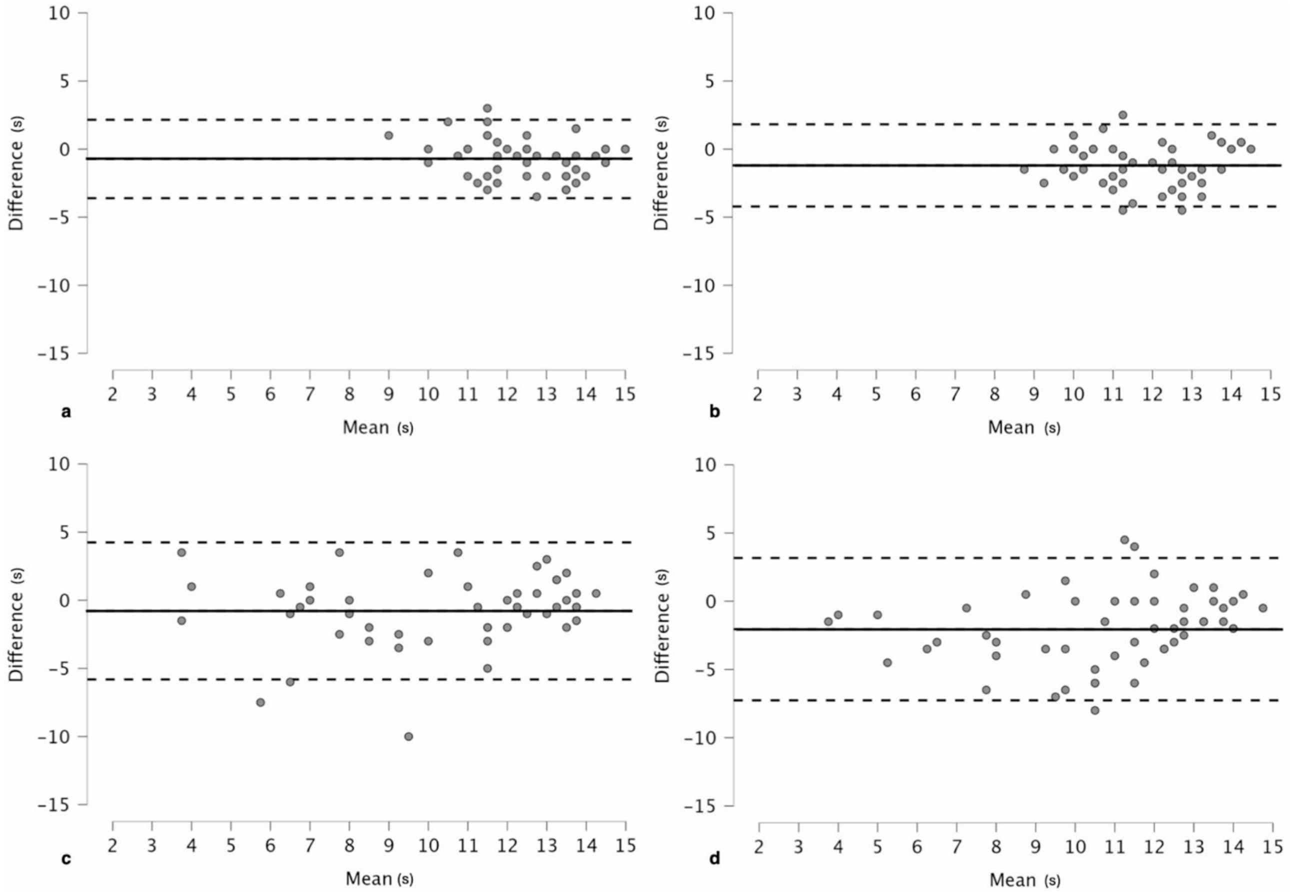

Fig. 4. Bland–Altman plot for: (**a**) average mean in wobble board anteroposterior condition; (**b**) average mean in wobble board mediolateral condition; (**c**) average mean in wobble board clockwise condition; (**d**) average mean in wobble board counterclockwise condition. Single bold lines indicate the average of the differences, whereas two dotted lines indicate the 95% limit of agreement.

movements further enhanced the reliability by minimizing trial-to-trial variability. These characteristics make WB a valuable tool for longitudinal monitoring of motor skills.

Bland–Altman plots (Figs. 2, 3, 4, 5) supported these findings, with negligible mean bias, indicating no systematic error between test–retest sessions. The relatively narrow LoA further confirmed WB's precision and consistency in capturing repeatable motor performance. However, LoA was wider for circular movements (clockwise and counterclockwise) than for linear movements (anteroposterior and mediolateral), likely reflecting the increased motor control demands of circular tasks. While high ICC values in circular tasks suggested good relative consistency across participants, broader LoA ranges revealed greater absolute variability in individual performance. This apparent contradiction arose because ICC measures relative consistency between sessions. In other words, ICC can remain high if the ranking of individuals is stable over time, even if the absolute differences between individuals become larger across sessions. The increased *SEM* and MDC_{95} in circular tasks compared to linear tasks (e.g., *SEM*: Clockwise = 1.28 vs. Anteroposterior = 0.75; MDC_{95}: Counterclockwise = 2.71 vs. Mediolateral = 1.74) further highlight the greater variability in complex movement patterns, highlighting the need for additional familiarization or practice to reduce variability and improve absolute agreement. These results highlight the WB's particular sensitivity in detecting subtle changes in motor performance, particularly in tasks characterized by higher inherent variability. This increased sensitivity likely stems from the dynamic nature of WB assessments, which require continuous motor adjustments and reactive control. Consequently, WB protocol may reveal fluctuations in motor coordination that remain undetected by more static or traditional evaluation methods, providing valuable insight into an individual's functional motor capacity and adaptability.

Compared to the traditional GPT, the WB demonstrated comparable or superior precision and reliability, despite the lack of repeated GPT measurements. This choice reflects the GPT's established role as a single-administration test[25]. To better contextualize our findings, under more demanding conditions, such as dual-task execution, GPT has shown *SEM* values ranging from 1.2 to 2.5 s and MDC_{95} values between 3.3 and 6.9 s[18]. These values are comparable to or higher than those obtained with WB in our study, further supporting its potential to detect subtle motor changes with high sensitivity. Similarly, studies on the Purdue Pegboard Test reported *SEM* values of 1.32–2.66 and MDC_{95} values of 3.66–7.37[42], whereas WB achieved *SEM* values as low as 0.75 and MDC_{95} values as low as 1.50 in linear tasks. This suggests WB's capability to detect subtle motor

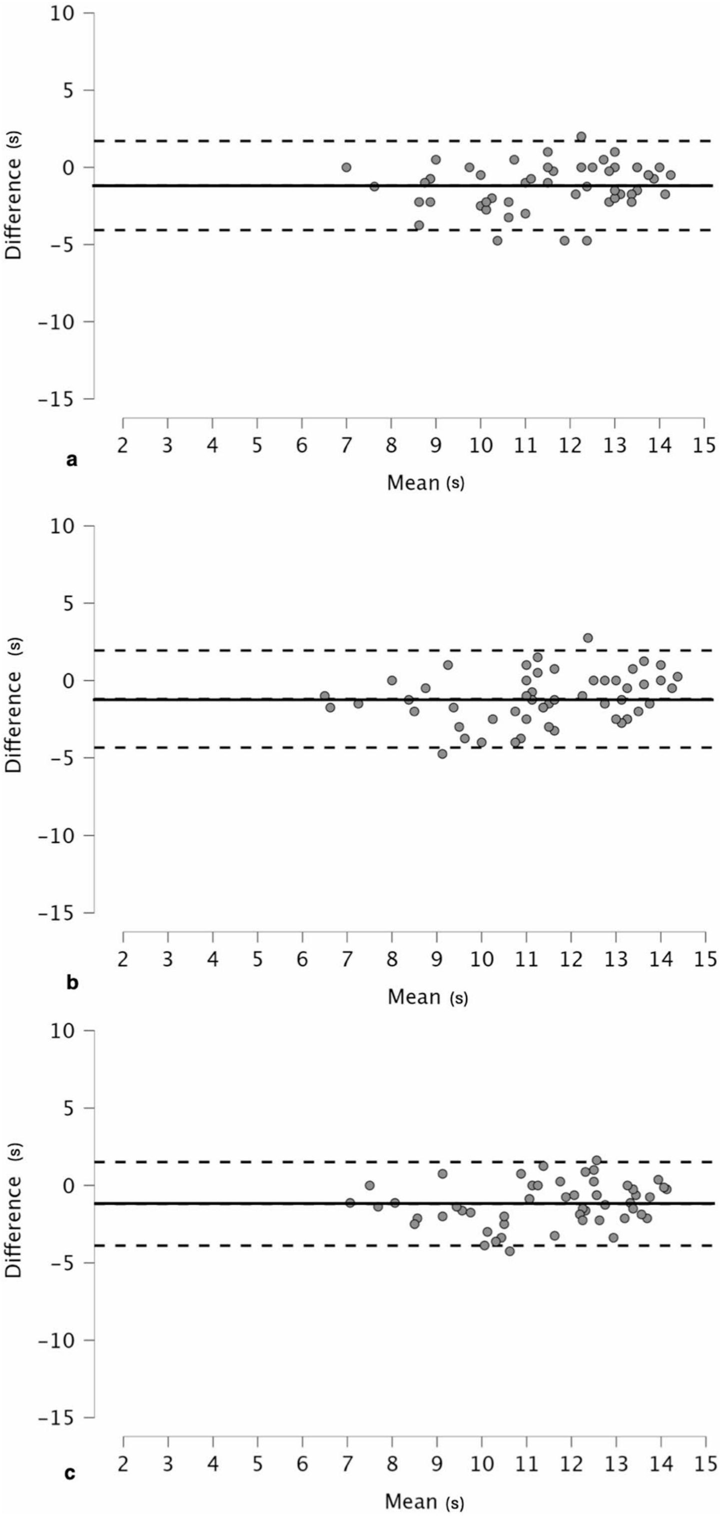

Fig. 5. Bland–Altman plot for: (**a**) mean of all conditions for preferred limb, (**b**) mean of all conditions for non-preferred limb and (**c**) average mean of all conditions. Single bold lines indicate the average of the differences, whereas two dotted lines indicate the 95% limit of agreement.

performance changes with high precision, making it suitable for longitudinal assessments. Moreover, while GPT evaluates simple grasp-and-release actions, WB incorporates both linear and circular movement patterns, providing a more comprehensive assessment of motor coordination. This versatility, combined with its high ICC values and sensitivity, supports WB's use as an effective tool for continual evaluation of motor skill evolution or decline in various contexts.

			GPT					
			Preferred		**Non-preferred**		**Limb average**	
		Condition	**r**	**p**	**r**	**p**	**r**	**p**
WB test	Preferred	Anteroposterior	**−0.26**	0.05	**−0.28**	0.03	**−0.30**	0.02
		Mediolateral	−0.22	0.10	**−0.36**	0.006	**−0.34**	0.01
		Clockwise	**−0.28**	0.03	**−0.52**	0.0001	**−0.47**	0.0004
		Counterclockwise	−0.18	0.18	**−0.37**	0.006	**−0.32**	0.01
	Non-preferred	Anteroposterior	−0.22	0.11	**−0.27**	0.04	**−0.27**	0.04
		Mediolateral	−0.06	0.65	**−0.29**	0.03	−0.21	0.11
		Clockwise	−0.21	0.12	**−0.43**	0.0001	**−0.37**	0.005
		Counterclockwise	−0.08	0.55	**−0.36**	0.007	**−0.27**	0.04
	Limb average	Anteroposterior	**−0.29**	0.02	**−0.34**	0.01	**−0.36**	0.007
		Mediolateral	−0.16	0.23	**−0.39**	0.003	**−0.33**	0.01
		Clockwise	**−0.27**	0.04	**−0.52**	0.0001	**−0.46**	0.0004
		Counterclockwise	−0.14	0.29	**−0.41**	0.02	**−0.33**	0.01
	Task average	Preferred	**−0.31**	0.02	**−0.54**	0.0001	**−0.49**	0.0002
		Non-preferred	−0.17	0.20	**−0.44**	0.0009	**−0.36**	0.0007
		Mean	−0.25	0.06	**−0.51**	0.0001	**−0.44**	0.007

Table 2. Pearson product moment correlation between wobble board (WB) and Grooved Pegboard test (GPT) performances for the preferred limb, non-preferred limb, limb averages, and the average across WB tasks, including correlation coefficients (r) and significance (p). Significant ($p \leq 0.05$) correlation coefficients were marked in bold.

Interestingly, circular tasks (clockwise and counterclockwise) demonstrated higher ICC values compared to linear conditions, while also exhibiting wider LoA, suggesting that although these tasks show strong relative reliability, effectively distinguishing between individuals, they are more variable at the intra-individual level. One possible explanation lies in the increased neuromuscular complexity of circular tracking, which requires feedforward planning, multijoint coordination, and continuous spatial recalibration. These demands may introduce more trial-to-trial fluctuations, especially under test–retest conditions, widening LoA despite strong overall consistency. However, this greater within-subject variability does not diminish the circular WB tasks' capacity to sensitively detect meaningful individual differences in motor control, highlighting their utility in assessing fine motor function.

Pearson correlation analysis further revealed negative relationships between WB and GPT outcomes, supporting WB validity. Correlations were generally stronger for the non-preferred hand and the hand averages than for the preferred hand. For example, non-preferred hand averages showed more negative correlations, suggesting that better WB performance (higher scores) corresponded with faster GPT times (lower scores), indicating a meaningful inverse relationship. No clear trends emerged when comparing linear and circular WB tasks, but task averages consistently yielded stronger correlations than individual tasks, emphasizing the value of aggregated measures in capturing overall motor performance relationships. This finding can be explained by the multifaceted nature of eye-hand coordination in real-life activities, which typically require the integration of multi-directional and multi-joint movements rather than isolated control in a single plane. By averaging performance across tasks and limbs, the aggregated WB score provides a more comprehensive and ecologically valid measure of fine motor ability, reflecting the overall integration and adaptability of sensorimotor control mechanisms[36,43]. This approach also reduces the influence of task-specific variability and measurement error, leading to more robust and representative associations with standard assessments such as the GPT.

The variation in correlation strength between preferred and non-preferred hands may reflect differences in neuromuscular proficiency and task difficulty[44,45]. The weaker correlations for preferred hand tasks may stem from their habitual use in daily activities, leading to more automatic and efficient motor performance with reduced variability. Similarly, athletes often rely on their preferred limb for sport-specific tasks like dribbling or shooting, further refining neuromuscular coordination and reducing variability in test outcomes. In contrast, the non-preferred limb requires greater conscious effort, potentially increasing variability and strengthening the inverse relationship between WB and GPT scores.

The negative correlation observed between WB and GPT scores is consistent with the nature of their performance. In WB, higher values indicate better performance (i.e., greater duration in maintaining the motion marker within the target zone), while in GPT lower completion times reflect higher dexterity. Therefore, an inverse relationship (negative correlation) is expected based on the nature of the tasks: participants who perform better on WB tend to complete the GPT faster. The inverse correlation may also be interpreted in light of the motor control theories[46]. In particular, WB emphasizes sustained precision during continuous dynamic tracking, which heavily relies on feedback mechanisms and postural control and eye-hand coordination, while GPT prioritizes speed and discrete object manipulation, tapping into more ballistic motor strategies. This difference may reflect a speed-accuracy trade-off, where proficiency in one domain does not linearly translate to the other. Therefore, such opposite motor demands can produce meaningful negative correlations without undermining concurrent validity.

The strongest correlation was observed in non-preferred hand clockwise movements ($r = -0.52$, $p < 0.0001$), while moderate correlations were found in preferred hand anteroposterior ($r = -0.30$, $p = 0.02$) and mediolateral ($r = -0.34$, $p = 0.01$) tasks. Interestingly, WB tasks showed a different pattern than GPT tasks with: tasks performed with the preferred hand having more significant correlations (10/12) compared to non-preferred hand (7/12). This discrepancy may be explained by the greater complexity of WB tasks, which require more coordination and control, particularly for the preferred limb. This increased difficulty likely results in greater variability in performance, enabling a clearer distinction between individuals with varying skill levels. These findings highlight WB's potential to assess fine motor skills and differentiate proficiency levels. While GPT may show stronger correlations in simpler tasks for the non-preferred hand due to increased variability, WB emphasizes the superior neuromuscular control of the preferred hand in complex tasks. This observation underscores WB's value in capturing fine motor performance. However, the variability in correlation outcomes across WB tasks and hands necessitates careful test selection for fine motor skill assessment. The strongest correlations were observed between WB task averages for the preferred hand and GPT non-preferred hand ($r = -0.54$), followed by counterclockwise average ($r = -0.52$) and clockwise average ($r = -0.43$) for the non-preferred hand. These findings suggested that tasks combining movements across hands or specific movement patterns (e.g., circular tasks) provide better insight into motor performance due to their higher correlation with GPT outcomes. This variability highlights the importance of careful test selection. Instead of matching specific target skills, selected tasks and hands should provide valid and reliable measures of fine motor performance. Some WB tasks or hand conditions showed weaker correlations, indicating they may be less useful for assessing fine motor skills in relation to GPT outcomes. A standardized approach prioritizing tasks and conditions with stronger correlations could enhance the accuracy and validity of WB-based motor assessments.

Compared to other digital tools for motor assessment, the WB system offers a practical balance between task specificity, portability, and low operational requirements. Kinect-based platforms allow for full-body motion capture and are suitable for assessing gross motor functions, particularly in gait and posture studies[47]. Force plates provide high-resolution data on balance and postural sway, although their use is generally restricted to lower-limb and stance evaluations[48]. While wearable sensors and inertial measurement units (IMUs) can capture detailed joint kinematics, they often require multi-point calibration and are more sensitive to placement error and noise[49]. In contrast, the WB allows for focused assessment of upper-limb coordination through standardized, visually guided tasks that require continuous motor control and feedback integration.

Conclusions

To the best of our knowledge, this is the first study to assess the validity and reliability of the WB as a tool for evaluating eye-hand coordination. The findings strongly support the WB as a valid and reliable instrument for assessing fine motor skills in the upper limbs. Its ability to integrate diverse movement patterns and detect subtle performance changes makes it particularly valuable for tracking motor skill development over time. Additionally, the WB shows promise for integration into training programs, complementing traditional assessment tools like the GPT. Future research should validate the WB in specific populations such as individuals with neurological conditions (e.g., Parkinson's disease or hand tremors), where it has already been applied in training interventions[10].Tailored protocols and specific validation procedures are needed to account for clinical implications and population-specific variability. Furthermore, investigating the effects of repeated WB testing is crucial to understanding potential motor learning effects and their influence on reliability and validity metrics. Addressing these aspects will enhance the WB's applicability in both clinical and research settings, maximizing its potential as a comprehensive assessment tool.

Practical and future implications

The findings provide clear guidance on using the WB as an assessment tool for upper-limb fine motor skills. Given the variability in correlation strengths across tasks and limbs, we recommend prioritizing circular WB tasks, which demonstrate consistently stronger correlations with GPT outcomes compared to linear tasks. Circular movements appear to better reflect the complexity and neuromuscular demands of motor skills measured by the GPT, making them a more valid assessment choice. This highlights the importance of selecting tasks and movement patterns that align with the target skill to ensure accurate and reliable assessments. Additionally, the results suggest that non-preferred limb assessments may be more sensitive in detecting motor performance variability, particularly in tasks requiring conscious effort and precision. This is especially relevant when assessing skills that are less automatic and more effort-dependent. In contrast, preferred limb tasks may be valuable for capturing automatic and highly coordinated motor performance, providing a complementary perspective. These insights support the use of WB-based assessments across various settings, including sports and clinical contexts, where fine motor skill evaluation is essential.

Future research should focus on validating the WB in specific populations such as individuals with neurological conditions. While previous studies have demonstrated the WB's utility in training interventions for such groups, tailored validation protocols are necessary to address condition-specific variability. Additionally, further investigation into motor learning effects during repeated WB testing is essential to determine how task familiarity could influence reliability and validity over time. Building on these findings, the WB can be further established as a complementary tool to traditional assessments, offering unique insights into fine motor skill performance by capturing complex movement patterns and subtle variations in motor coordination.

Strengths and limitations

This study represents a novel approach in evaluating the validity and reliability of the WB for fine motor skill assessment in healthy young adults. While this focus is a strength in terms of novelty, it also limits the generalizability of the findings. The homogeneity of the sample, restricted to young, neurologically healthy participants, limits the applicability of the findings to other populations, such as older adults or individuals with motor impairments. Moreover, although a 5-day rest period was observed between sessions, practice effects cannot be entirely excluded and may have contributed to improved performance at retest. An analysis of GPT performance in our sample, including SD and range, suggested that our results aligned with those reported in the literature. Specifically, mean times for the preferred (63.7 ± 9.3 s) and non-preferred hand (69.0 ± 12.7 s), as well as the average across both hands (66.4 ± 9.9 s), are comparable to those reported in previous studies (e.g., 71.0 ± 9.0 s for the dominant and 75.0 ± 9.0 s for the non-dominant hand[50] and 67.0 ± 8.0 s and 72.0 ± 10.0 s, respectively)[16]. These results indicate that the GPT performance of our sample was consistent with normative data from healthy populations, suggesting adequate representativeness within this demographic. However, including more diverse participants in terms of age, clinical status, and motor proficiency will be essential to confirm the WB's broader applicability and improve its integration with existing fine motor assessment tools.

Data availability

Anonymous data is available upon request (Dr. Francesca Di Rocco; francesca.dirocco1@unicas.it).

Received: 15 March 2025; Accepted: 30 September 2025

References

1. Wang, Y. C. et al. Assessing dexterity function: A comparison of two alternatives for the NIH toolbox. *J. Hand Ther.* **24**, 313–321 (2011).
2. Ostwald, S. K., Snowdon, D. A., Rysavy, S. D. M., Keenan, N. L. & Kane, R. L. Manual dexterity as a correlate of dependency in the elderly. *J. Am. Geriatr. Soc.* **37**, 963–969 (1989).
3. Voelcker-Rehage, C. Motor-skill learning in older adults-a review of studies on age-related differences. *Eur. Rev. Aging Phys. Act.* **5**, 5–16 (2008).
4. Ec, P. On prediction in skilled movements. *Psychol. Bull.* **54**, 467–478 (1957).
5. Voelcker-Rehage, C. & Willimczik, K. Motor plasticity in a juggling task in older adults—A developmental study. *Age Ageing* **35**, 422–427 (2006).
6. Luo, Z., Jose, P. E., Huntsinger, C. S. & Pigott, T. D. Fine motor skills and mathematics achievement in East Asian American and European American kindergartners and first graders. *Br. J. Dev. Psychol.* **25**, 595–614 (2007).
7. Cameron, C. E. et al. Fine motor skills and executive function both contribute to kindergarten achievement. *Child Dev.* **83**, 1229–1244 (2012).
8. Grissmer, D., Grimm, K. J., Aiyer, S. M., Murrah, W. M. & Steele, J. S. Fine motor skills and early comprehension of the world: Two new school readiness indicators. *Dev. Psychol.* **46**, 1008–1017 (2010).
9. Dankert, H. L., Davies, P. L. & Gavin, W. J. Occupational therapy effects on visual-motor skills in preschool children. *Am. J. Occup. Ther.* **57**, 542–549 (2003).
10. Papale, O. et al. Do hand exercises influence physiological hand tremor? An observational cohort study on healthy young adults. *Appl. Sci. Switz.* **14**, 4467 (2024).
11. Borecki, L., Tolstych, K. & Pokorski, M. *Computer Games and Fine Motor Skills* (Springer, 2013).
12. Nayak, A. K. Effect of hand-eye coordination on motor coordinative ability of tribal adolescents. *Int. J. Phys. Educ. Sports Health* **2**, 328–330 (2015).
13. Mathiowetz, V., Volland, G., Kashman, N. & Weber, K. Adult norms for the box and block test of manual dexterity. *Am. J. Occup. Ther. Off. Publ. Am. Occup. Ther. Assoc.* **39**, 386–391 (1985).
14. Lawson, I. Purdue pegboard test. *Occup. Med.* **69**, 376–377 (2019).
15. D'Hondt, E., Deforche, B., Bourdeaudhuij, I. D. & Lenoir, M. Childhood obesity affects fine motor skill performance under different postural constraints. *Neurosci. Lett.* **440**, 72–75 (2008).
16. Smith, Y. A., Hong, E. & Presson, C. Normative and validation studies of the nine-hole peg test with children. *Percept. Mot. Skills* **90**, 823–843 (2000).
17. Gershon, R. C. et al. Assessment of neurological and behavioural function: The NIH Toolbox. *Lancet Neurol.* **9**, 138–139 (2010).
18. Petrigna, L. *et al.* The execution of the Grooved Pegboard test in a dual-task situation: A pilot study. *Heliyon* **6**, (2020).
19. Rourke, B. P., Yanni, D. W., MacDonald, G. W. & Young, G. C. Neuropsychological significance of lateralized deficits on the Grooved Pegboard test for older children with learning disabilities. *J. Consult. Clin. Psychol.* **41**, 128–134 (1973).
20. Hamilton, L. D., Mazzo, M. R., Petrigna, L., Ahmed, A. A. & Enoka, R. M. Poor estimates of motor variability are associated with longer grooved pegboard times for middle-aged and older adults. *J. Neurophysiol.* **121**, 588–601 (2019).
21. Bezdicek, O. et al. Grooved pegboard predicates more of cognitive than motor involvement in Parkinson's disease. *Assessment* **21**, 723–730 (2014).
22. Wang, L.-I. The kinetics and stiffness characteristics of the lower extremity in older adults during vertical jumping. *J. Sports Sci. Med.* **7**, 379–386 (2008).
23. Tolle, K. A., Rahman-Filipiak, A. M., Hale, A. C., Andren, K. A. K. & Spencer, R. J. Grooved Pegboard Test as a measure of executive functioning. *Appl. Neuropsychol. Adult* **27**, 414–420 (2020).
24. Fellows, S. J., Noth, J. & Schwarz, M. Precision grip and Parkinson's disease. *Brain* **121**, 1771–1784 (1998).
25. Bryden, P. & Roy, E. A new method of administering the Grooved Pegboard Test: Performance as a function of handedness and sex. *Brain Cogn.* **58**, 258–268 (2005).
26. Fusco, A., Fuchs, P. X., De Maio, M., Wagner, H. & Cortis, C. A novel approach to measuring wobble board performance in individuals with chronic ankle instability. *Heliyon* **6**, e04937 (2020).
27. De Maio, M., Cortis, C., Iannaccone, A., Da Silva, R. A. & Fusco, A. Association between anthropometric variables, sex, and visual biofeedback in dynamic postural control assessed on a computerized wobble board. *Appl. Sci.* **11**, 8370 (2021).
28. Riemann, B. L. & Lephart, S. M. The Sensorimotor System, Part II: The Role of Proprioception in Motor Control and Functional Joint Stability.
29. Keele, S. W. & Posner, M. I. Processing of visual feedback in rapid movements. *J. Exp. Psychol.* **77**, 155–158 (1968).
30. Shadmehr, R. & Krakauer, J. W. A computational neuroanatomy for motor control. *Exp. Brain Res.* **185**, 359–381 (2008).
31. De Maio, M., Castellani, L., Cugusi, L., Cortis, C. & Fusco, A. The effect of a combined exercise program on postural control and fine motor skills in Parkinson's disease: Study design. *Int. J. Environ. Res. Public. Health* **19**, 15216 (2022).

32. Tobler-Ammann, B. C. et al. Concurrent validity and test-retest reliability of the virtual peg insertion test to quantify upper limb function in patients with chronic stroke. *J. NeuroEng. Rehabil.* **13**, 8 (2016).
33. Weir, J. P. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J. Strength Cond. Res.* **19**, 231–240 (2005).
34. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences Second Edition*.
35. Von Elm, E. et al. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: Guidelines for reporting observational studies. *J. Clin. Epidemiol.* **61**, 344–349 (2008).
36. Corey, D., Hurley, M. & Foundas, A. Right and left handedness defined: A multivariate approach using hand preference and hand performance measures. *Neuropsychiatry. Neuropsychol. Behav. Neurol.* **14**, 144–152 (2001).
37. Vingerhoets, G. et al. Laterality indices consensus initiative (LICI): A Delphi expert survey report on recommendations to record, assess, and report asymmetry in human behavioural and brain research. *Laterality* **28**, 122–191 (2023).
38. Amato, A. et al. Effects of a resistance training protocol on physical performance, body composition, bone metabolism, and systemic homeostasis in patients diagnosed with parkinson's disease: A pilot study. *Int. J. Environ. Res. Public. Health* **19**, 13022 (2022).
39. Bland, J. M. & Altman, D. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **327**, 307–310 (1986).
40. Flores, P., Coelho, E., Mourão-Carvalhal, M. I. & Forte, P. M. Preliminary adaptation of motor tests to evaluate fine motor skills associated with mathematical skills in preschoolers. *Eur. J. Investig. Health Psychol. Educ.* **13**, 1330–1361 (2023).
41. Fusco, A. et al. Dynamic balance evaluation: Reliability and validity of a computerized wobble board. *J. Strength Cond. Res.* **34**, 1709–1715 (2020).
42. Lee, P. et al. The test-retest reliability and the minimal detectable change of the Purdue pegboard test in schizophrenia. *J. Formos. Med. Assoc.* **112**, 332–337 (2013).
43. Latash, M. L. The bliss (not the problem) of motor abundance (not redundancy). *Exp. Brain Res.* **217**, 1–5 (2012).
44. Torell, F., Franklin, S., Franklin, D. W. & Dimitriou, M. Goal-directed modulation of stretch reflex gains is reduced in the non-dominant upper limb. *Eur. J. Neurosci.* **58**, 3981–4001 (2023).
45. Noguchi, T., Demura, S., Takahashi, K., Demura, G. & Mori, Y. Differences in muscle power between the dominant and nondominant upper limbs of baseball players. *J. Strength Cond. Res.* **28**, 82–86 (2014).
46. Latash, M., Levin, M., Scholz, J. & Schöner, G. Motor control theories and their applications. *Med. Mex.* **46**, 382 (2010).
47. Clark, R. A. et al. Validity of the microsoft kinect for assessment of postural control. *Gait Posture* **36**, 372–377 (2012).
48. Paillard, T. & Noé, F. Techniques and methods for testing the postural function in healthy and pathological subjects. *BioMed. Res. Int.* **2015**, 1–15 (2015).
49. Sabatini, A. M. Estimating three-dimensional orientation of human body parts by inertial/magnetic sensing. *Sensors* **11**, 1489–1525 (2011).
50. Ruff, R. M. & Parker, S. B. Gender- and age-specific changes in motor speed and eye-hand coordination in adults: Normative values for the Finger Tapping and Grooved Pegboard Tests. *Percept. Mot. Skills* **76**, 1219–1230 (1993).

Acknowledgements

This work was financially supported by the National Taiwan Normal University (NTNU) within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan.

Author contributions

All the authors contributed to the completion of this article. Their individual contributions were as follows: conceptualization, F.D.R., M.D.M, C.C., and A.F.; methodology, F.D.R., M.D.M., O.P., E.F., A.B., C.C., and A.F.; formal analysis, F.D.R and M.D.M.; investigation, F.D.R; data curation, F.D.R; writing—original draft preparation, F.D.R, M.D.M, and P.X.F.; writing—review and editing, F.D.R, M.D.M., O.P, E.F., P.X.F., R.D.S. , A.B., C.C. and A.F.; supervision, C.C. and A.F.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-22782-w.

Correspondence and requests for materials should be addressed to C.C.