Design a Haptic-Combined Virtual Reality System to Improve Box and Block Test (BBT) for Upper Extremity Function Assessment

Ying Dong  
BeiHang University School of Biological Science and Medical Engineering

Xiaoyu Liu (✉️ x.y.liu@buaa.edu.cn)  
BeiHang University School of Biological Science and Medical Engineering  
https://orcid.org/0000-0001-7854-4866

Min Tang  
BeiHang University School of Biological Science and Medical Engineering

Hongqiang Huo  
BeiHang University School of Biological Science and Medical Engineering

Duo Chen  
BeiHang University School of Biological Science and Medical Engineering

Zhixin Wu  
BeiHang University School of Biological Science and Medical Engineering

Ran An  
BeiHang University School of Biological Science and Medical Engineering

Yubo Fan  
BeiHang University School of Biological Science and Medical Engineering

Research

Keywords: Stroke, Upper extremity assessment, Virtual Box and Blocks Test, Haptic device, Concurrent validity, Test-retest reliability

Posted Date: June 5th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-32776/v1

License: ☎️ This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Background: The Box and Block Test (BBT) has been widely used to assess gross upper extremity (UE) motor function. We designed a haptic-combined virtual reality (VR) system, named VBBT, to improve BBT with kinematic parameters for more specific assessments.

Methods: According to the block-transfer task in BBT, we designed a VBBT system with a haptic device to provide a sense of the grasping force and block’s gravity during task performance for UE function assessment. Besides the number of transferred blocks (N-TB), some kinematic parameters were also collected in VBBT for more specific assessment, including the number of zero-crossings of moving acceleration (N\textsubscript{ZC-ACC}), the number of zero-crossings of derivative of releasing force (N\textsubscript{ZC-DRF}), the ratio of path length and linear length (PLR) and the distance between barrier partition and drop position (DDP). 16 post-stroke patients and 113 healthy subjects were recruited to examine quantitative performances, concurrent validity, test-retest reliability and subjective preference between the BBT and VBBT.

Results: Based on the measurements of 95% healthy subjects, the normative ranges of quantitative performances were established for BBT and VBBT. Deficiencies in patients’ UE function could be identified when their measurements fell outside. A moderate correlation was found in the N-TBs between the VBBT and BBT ($r = 0.42$). The measurements in VBBT presented a stronger age-related correlation than that in BBT ($R^2 = 0.57$ and $R^2 = 0.16$). The N-TBs in both BBT and VBBT were strongly correlated to the Action Research Arm Test (ARAT) ($|r| = 0.84$ and $0.83$), and the N\textsubscript{ZC-ACC} and N\textsubscript{ZC-DRF} in VBBT also showed significant correlations. ($|r| = 0.76$ and $0.79$). The N-TB, N\textsubscript{ZC-ACC} and N\textsubscript{ZC-DRF} in VBBT showed a good test-retest reliability (ICC = 0.75, 0.78 and 0.80) while a moderate reliability was found in BBT (ICC = 0.62). For the patient preference, VBBT was given a higher score than the BBT ($p<0.05$) for its enjoyment and completion effort performing.

Conclusion: The VBBT improved a gross manual assessment in BBT, which can provide clinically validated, reliable and motiveive assessment with kinematic parameters for specific UE motor functions of post-stroke patients.

Introduction

Stroke is one of the most threatening diseases to human health worldwide due to its extremely high mortality and disability [1]. With the increasing and ageing world’s population, the number of stroke patients will inevitably continue to increase. It was reported that the lifetime risk of stroke among adults aged 25 years or older ranged from 22.8% in 1990 to 24.9% in 2016, showing a relative increase of 8.9% [2]. The number of stroke patients was estimated to be double the present level by 2050 [3]. Impairments in upper extremity (UE) function is a common consequence in post-stroke patients. Approximately 60–80% of patients remain deficient in UE motor function into the chronic phases [4]. Thus, the recovery of UE function requires long-term rehabilitative intervention. Most patients have to continue their self-rehabilitation at home or in the local community, since they cannot afford the high cost of rehabilitation training in special rehabilitation units [5, 6].

At present, some intelligent rehabilitation devices have been developed to make home-based training possible for post-stroke patients [7–9]. After a period of device-assisted training at home, the recovery of UE function should be accurately evaluated for further effective intervention [10]. Several assessments of UE function have been widely used in stroke rehabilitation, such as the Fugl-Meyer Assessment (FMA) [11–13], Action Research Arm Test (ARAT) [14], Wolf Motor Function Test (WMFT) [15], and Modified Ashworth Scale (MAS) [16, 17]. However, these assessment scales seem to be unsuitable for home-based rehabilitation due to several limitations [18, 19]. One limitation refers to the requirement of on-site supervision and subjective ratings by occupational therapists, making them impossible for self-use assessments at home. Additionally, most assessments are insensitive to fine motor skills, known as the ceiling effect, and fail to respond
to the recovery of UE function in patients in the later periods of rehabilitation [20]. These limitations of the assessment scales probably result in non-selective and targeted rehabilitative training [21, 22]. Therefore, an objective and sensitive assessment of UE motor function is necessary for rehabilitation at home. This would enable the transmission of quantitative data of a patient's UE motor function to a professional therapist for remote administration and guidance, speeding up the patient's recovery progress.

Recent advancements in virtual reality (VR) rehabilitation have allowed automated and remote assessments to be commercially available for post-stroke patients with low cost. VR rehabilitation assessments are characterized by various functional tasks [23]. The recovery of motor function can be evaluated based on the quality of task performance [19]. Some quantitative kinematic and biomechanical parameters have been considered critical references for off-site assessment [24]. In addition, VR environments with high immersion and interaction improve patient motivation to participate in unsupervised assessments [23, 25, 26]. Recently, haptic devices have been incorporated into VR rehabilitation. As an interactive tool, the haptic device provides force feedback such as tactile or grasping sensations to patients during task performance. For example, Gerber et al. combined haptic devices with a 3D virtual environment to design functional tasks, such as tool use, object moving and spelling in activities of daily life (ADL), to assess the cognitive and fine motor functions of chronic traumatic brain injury (TBI) patients [27]. Fluet et al. developed the haptic Virtual Peg Insertion Test (VPIT) to quantitatively measure performance while placing nine pegs into holes [28]. The VPIT test-retest study concluded that some kinematic parameters acquired by the haptic device remained consistent and reliable when performed by stroke patients [29]. In our previous study, we designed several virtual guiding tasks combined with haptic feedback to evaluate wrist motor function in patients with upper motor neuron lesions. Deficiencies in wrist motor function were identified when patient performance was outside normative ranges [30]. Related studies have indicated that rehabilitative assessments in VR combined with haptic rendering not only provide a more immersive interacting environment but also allow more reliable measurements of patients’ quantitative performances.

The Box and Block Test (BBT) has been widely used to assess UE motor function due to its merits, such as simple operation, short time consumption and high validity [31]. Kim et al. developed a virtual BBT program using a Microsoft Kinect system for unsupervised assessment. The test result indicated a strong correlation between the numbers of blocks moved in the BBT and virtual BBT for both nonhemiplegic and hemiplegic sides [32]. However, without haptic rendering, movements of grasping in virtual environments is significantly different from natural grasping in real physical environments [33]. In addition, the virtual BBT system returned the number of blocks to evaluate patient UE motor function but lacked the acquisition of kinematic parameters for specific analysis. In the current study, we designed a virtual BBT system combined with a haptic device (VBBT) that can provide a sense of the grasping force and the block's gravity during block-transfer task performance. In addition to the number of transferred blocks (N-TB), some kinematic parameters could also be collected in VBBT for more specific assessment of UE motor functions, such as motion smoothness, hand dexterity and motion efficiency. We recruited healthy subjects and post-stroke patients to examine the quantitative performance, concurrent validity, test-retest reliability and subjective preference between the BBT and VBBT. The VBBT system is expected to improve BBT for home-based rehabilitation.

**Methods**

**Subjects**

Healthy right-handed subjects aged over 18 years without any neurological disorder or musculoskeletal impairment were recruited from universities and local communities. Patients with chronic stroke were recruited from a rehabilitation unit in Beijing, China. The patient inclusion criteria were as follows: 1) received a stroke diagnosis at least three months earlier confirmed by brain CT or MRI findings, 2) was aged older than 18 years, 3) was right handed with an affected right hand, 4) could sit steadily on a chair without armrest support, 5) was able to move 3 blocks in the BBT within one minute, and
6) understood the whole experimental procedure. The exclusion criteria were as follows: 1) unstable fracture of the upper extremity on the hemiplegic side; 2) spatial or visual disorders, such as visual neglect; 3) epilepsy caused by visual stimuli (lights, television, etc.) in the previous six months; and 4) dizziness in the VR environment. All subjects were given written and verbal information on the current study. A signed informed consent statement was received from each subject. This study adhered to the tenets of the Declaration of Helsinki, and ethical approval was approved by Beihang University (BM20180017).

**Apparatus and program**

A haptic feedback device (Omega.7, Force Dimension Inc., Switzerland; Fig. 1a) was used to provide interactive forces, including grasping force and block activity. The haptic device allowed a translating force of 12.0 N and grasping force of 8.0 N, as well as an operating space of Φ160 × 110 mm for translation and 240°×140°×180° for rotation. A VR headset (Oculus Rift, Facebook Inc., US) was used to provide a 3D virtual environment that allowed spatial visualization and operation. An open source software library Chai3D combined with the OpenGL library was used to render visualization and haptic interaction in the VBBT program.

A virtual test box with a barrier partition in the middle was created in the VR environment (Fig. 1b). The block was created, one by one, in the compartment of the box on the side of the tested hand. In the case of the VBBT, when a subject had completed one trial in which a block was moved from one compartment to the other, another block was then automatically created (Supplementary Video). This was designed to provide movement consistency and avoid obstructions to the target block by other blocks during grasping. Each block was attributed physical properties, including tactile contact and gravity. A virtual grasping tool was attached to the handle of a haptic device. As a subject moved the handle in the real environment, the virtual tool synchronously performed the same motion in the virtual environment. When contact occurred between the tool and a block, haptic interactions were computed through force rendering algorithms. During the interaction, a haptic thread was created to compute the resulting forces between the virtual tool and block, providing a sense of haptic interaction for the subjects (Fig. 1c).

In the VBBT, some kinematic parameters can be collected by the haptic device to evaluate specific UE motor function in detail. Originally, the haptic device collected the position and velocity of virtual objects, as well as the grasping force of the virtual tool (Fig. 1c). Considering previously validated parameters for UE assessment in the literature [28, 34, 35], the kinematic parameters used in the VBBT were determined and included the following:

1) N\textsubscript{ZC-ACC}: The number of zero-crossings of the moving acceleration in a block transfer, which was used to assess the smoothness and coordination of UE movement.

2) N\textsubscript{ZC-DRF}: The number of zero-crossings of the derivative of the releasing force, which was used to assess hand dexterity.

3) PLR: The ratio of the path length and linear length in a block transfer trial, which was used to assess the efficiency of UE movement.

4) DDP: The distance between the barrier partition and the drop position of a block, which was used to present the efficiency of UE movement.

**Experimental procedure**
For the healthy subjects, they were asked to perform the BBT and VBBT, respectively. The BBT was performed according to previously published instructions [31]. In the VBBT, the subjects were seated on a standard height chair facing the haptic device that was placed at their middle line. We first introduced the operation of the haptic device to the subjects until they understood how to use it in the experiment. In the familiarization session of the VBBT, the subjects, wearing the VR headset, were instructed to move the blocks to the opposite compartment until they became fully familiarized with how to operate the haptic device to interact with virtual objects in a VR environment. In the formal session of the VBBT, the subjects were given one minute to move as many blocks as possible until the program automatically stopped. Four weeks after the first experimental session, the healthy subjects were asked to perform the BBT and VBBT again as a retest.

For the patients, they were also asked to perform several widely used assessment scales of UE motor function and cognitive function, including the BBT, ARAT, FMA-UE, Brunnstrom and Mini-Mental State Examination (MMSE). All assessments were performed according to the standard instructions reported in previous studies [11, 31, 36–38]. Then, the patients were asked to perform the VBBT with the same procedures as the healthy subjects. Adequate rest was provided for the patients when they felt tired during the performance. Specially, the test was immediately stopped once the patients felt uncomfortable. When the patients finished the VBBT, they were given a questionnaire to evaluate their subjective preferences for the BBT and VBBT, and an informal interview was conducted regarding their preference. There were 7 questions corresponding to 7 items in the questionnaire, including difference, understandability, enjoyment, attraction, relaxation, effort and tiredness. The patients gave scores (from 1 to 7) to show how true each question was for both the BBT and VBBT, in which 1 indicated “not at all true” and 7 indicated “very true”. The questions in the questionnaire were as follows:

Q1: I don't think there is a significant difference between the BBT and VBBT. (Difference)
Q2: I think the BBT/VBBT is quite easy to understand. (Understandability)
Q3: I enjoy to perform BBT/VBBT very much. (Enjoyment)
Q4: I think the BBT/VBBT can hold my attention very well. (Attraction)
Q5: I feel very relaxed in performing the BBT/VBBT. (Relaxation)
Q6: I put a lot of effort into the BBT/VBBT. (Effort)
Q7: I feel very tired after the BBT/VBBT. (Tiredness)

Data analysis

A demographic analysis was performed with both the healthy and patient subjects. We divided the subjects into three groups according to their ages: the young group, 18–44 years; the middle-aged group, 45–59 years; and the senior group, 60 years or older. A regression analysis was performed between the age and the quantitative performances (N-TB) in both BBT and VBBT. The kinematic parameters (N<sub>ZC</sub>-ACC, N<sub>ZC</sub>-DRF, PLR, and DDP) were compared between the healthy and patient subjects. We determined a normative range by 95% confidence interval (CI) for 2-sided measurements (2.5%-97.5%) of healthy-subject performances for each kinematic parameter. Specific deficiencies in a patient’s motor function were identified when the patient’s measurements fell outside of the normative ranges. A Correlation analysis was conducted between the parameters collected in the BBT (N-BT) as well as the VBBT (N-BT, N<sub>ZC</sub>-ACC, N<sub>ZC</sub>-DRF, PLR, and DDP) and the clinical scales, such as FMA-UE and ARAT for concurrent validity. Spearman rank correlation coefficient was used for the classification [39]: 0.20 or below, little if any correlation; 0.20–0.40, weak correlation; 0.40–0.70, moderate correlation; and 0.70-1.0, strong correlation. A intraclass correlation coefficient (ICC) was used to examine
the test-retest reliability of BBT and VBBT [40, 41]: 0.5 or below, poor reliability; 0.50–0.75, moderate reliability; 0.75–0.9, good reliability; and 0.9–1.0, excellent reliability. SPSS version 22.0 (SPSS Inc, Chicago, Illinois) was used to analyse all the data.

**Results**

**Demographics and clinical characteristics of subjects**

A total of 113 healthy subjects and 16 patients with chronic stroke were included in the current study. For the healthy subjects, 19 males and 26 females aged 28.9 ± 7.2 years were in the young group; 12 males and 19 females, aged 51.4 ± 4.0 years were in the middle-aged group; 12 males and 25 females aged 70.8 ± 9.2 years were in the senior group. For the patients, none were in the young group; 3 males and one female aged 52.3 ± 4.4 years were in the middle-aged group; 8 males and 4 females aged 73.1 ± 6.0 years were in the senior group. All patients with Brunnstrom stage (BS) ≥ Ⅲ for the arm and BS ≥ Ⅳ for the hand were subjected to stroke between 3 months and 36 months. The scores on the MMSE, FMA-UE, ARAT, and BBT were 26.2 ± 4.7, 53.7 ± 10.6, 45.3 ± 11.5 and 33.1 ± 11.9, respectively. More details on the demographic and clinical characteristics of patients are listed in Table 1.
Table 1
Demographic and clinical characteristics of patients (n = 16)

| Patient | Age | Age Group | Gender | Stroke Cause | Stroke Duration | BS-Arm | BS-Hand | MMSE | FMA-UE | ARAT | BBT |
|---------|-----|-----------|--------|--------------|----------------|--------|---------|------|--------|------|-----|
| P1      | 48  | Middle-aged | Male   | Ischemic     | 5 months       | −      | −       | 28   | 34     | 17   | 13  |
| P2      | 49  | Middle-aged | Female | Ischemic     | 4 months       | −      | −       | 30   | 58     | 57   | 40  |
| P3      | 55  | Middle-aged | Male   | Ischemic     | 15 months      | −      | −       | 29   | 63     | 53   | 53  |
| P4      | 57  | Middle-aged | Male   | Ischemic     | 4 months       | −      | −       | 27   | 60     | 51   | 27  |
| P5      | 64  | Senior     | Male   | Ischemic     | 15 months      | −      | −       | 25   | 54     | 45   | 27  |
| P6      | 65  | Senior     | Female | Ischemic     | 3 months       | −      | −       | 27   | 61     | 57   | 44  |
| P7      | 66  | Senior     | Male   | Hemorrhage   | 6 months       | −      | −       | 25   | 63     | 57   | 44  |
| P8      | 70  | Senior     | Male   | Ischemic     | 5 months       | −      | −       | 28   | 59     | 45   | 42  |
| P9      | 72  | Senior     | Male   | Ischemic     | 5 months       | −      | −       | 27   | 63     | 52   | 30  |
| P10     | 72  | Senior     | Male   | Ischemic     | 5 months       | −      | −       | 25   | 38     | 32   | 20  |
| P11     | 73  | Senior     | Male   | Ischemic     | 5 months       | −      | −       | 30   | 41     | 39   | 32  |
| P12     | 75  | Senior     | Female | Ischemic     | 7 months       | −      | −       | 15   | 46     | 34   | 16  |
| P13     | 76  | Senior     | Female | Hemorrhage   | 6 months       | −      | −       | 28   | 63     | 39   | 37  |
| P14     | 80  | Senior     | Female | Ischemic     | 4 months       | −      | −       | 16   | 38     | 37   | 21  |
| P15     | 81  | Senior     | Male   | Ischemic     | 4 months       | −      | −       | 30   | 55     | 57   | 49  |
| P16     | 83  | Senior     | Male   | Ischemic     | 36 months      | −      | −       | 30   | 63     | 53   | 35  |
| Mean    | 67.88 | −         | −      | −           | −              | −      | −       | 26.2 | 53.7   | 45.3 | 33.1|
| SD      | 10.93 | −         | −      | −           | −              | −      | −       | 4.7  | 10.6   | 11.5 | 11.9|

SD Standard Deviation, BS Brunnstrom stage on the hemiplegic side, MMSE Mini Mental Status Examination, FMA-UE Fugl-Meyer Assessment for the Upper Extremity, ARAT Action Research Arm Test, N-BBT Number of Box and Block Test, N-VBBT Number of Virtual Box and Block Test.

Block-transfer performance in BBT and VBBT
The healthy subjects moved more blocks in the BBT (95% range = 53–89, median = 77 in the young group; 95% range = 61–91, median = 74 in the middle-aged group; 95% range = 48–87, median = 68 in the senior group; Fig. 2a) than in the VBBT (95% range = 28–45, median = 37 in the young group; 95% range = 20–45, median = 28 in the middle-aged group; 95% range = 11–37, median = 24 in the senior group; Fig. 2b), and there was a medium correlation between the two assessments ($r$ = 0.42, $p < 0.001$; Fig. 3a). In addition, the age of healthy subjects accounted for more of the variance in VBBT performance ($R^2$ = 0.57, $p < 0.001$; Fig. 3b) compared to that in BBT performance ($R^2$ = 0.16, $p < 0.001$; Fig. 3c), suggesting that age differences should be taken into consideration for VBBT performance. None of the middle-aged patients could move more blocks than 2.5% of the middle-aged healthy subjects (lower limit of the normative range) in either the BBT or the VBBT. In the senior group, only one patient moved more VBBT blocks than the lowest 2.5% of the senior healthy subjects; this individual also moved more BBT blocks than the lower limit of the normative range in the BBT.

**Kinematic measurements in VBBT**

The normative ranges of the kinematic parameters in the VBBT were determined by measurements of healthy subjects (n = 113; 95% CI for 2-sided metrics, Fig. 4). For the $N_{ZC}$-ACC (Fig. 4a), all the middle-aged patients fell outside the normative range (1.2–4.2), and 3 out of 12 senior patients fell outside the normative range (1.3–7.3). For the $N_{ZC}$-DRF (Fig. 4b), all the middle-aged patients fell outside the normative range (3.2–7.7), and 3 out of 12 senior patients fell outside the normative range (4.3–12.3). For the PLR (Fig. 4c), only one middle-aged patient and 2 senior patients fell outside the normative ranges (1.1–1.8 and 1.1–1.7, respectively). For the DDP (Fig. 4d), only one middle-aged patient and 5 senior patients fell outside the normative ranges (10.0–91.2 and 19.7–76.8, respectively). Specific characteristics for a patient’s UE motor function, such as the dexterity, smoothness and coordination of limb movement, were evaluated according to his/her quantitative performance compared to the normative range.

**Concurrent validity for BBT and VBBT**

The results of correlation analyses for clinical validity in the patients (n = 16) are presented in Table 2. The N-TB in BBT was strongly correlated with ARAT ($r$ = 0.83, $p < 0.001$) and moderately correlated with FMA-UE ($r$ = 0.66, $p < 0.01$). A strong correlation was found between N-TB in VBBT and ARAT ($r$ = 0.84, $p < 0.001$), while a moderate correlation was found between N-TB in VBBT and FMA-UE ($r$ = 0.61, $p < 0.01$). The kinematic measurement $N_{ZC}$-ACC, $N_{ZC}$-DRF and DDP in VBBT were strongly correlated to ARAT ($|r| \geq 0.76; p < 0.001$), except for PLR ($|r| = 0.56; p < 0.05$). All the kinematic measurements in VBBT were weakly correlated to FMA-UE scores ($|r| \leq 0.48$). This result indicated that the VBBT was validated for task-oriented assessments of UE motor function in chronic stroke patients.
Table 2
Spearman’s rank correlation coefficients between clinical scale scores and BBT/VBBT parameters

| Clinical scale | N-TB (BBT) | N-TB (VBBT) | N\textsubscript{ZC}-ACC (VBBT) | N\textsubscript{ZC}-DRF (VBBT) | PLR (VBBT) | DDP (VBBT) |
|----------------|------------|-------------|------------------------------|---------------------|-----------|-----------|
| ARAT           | 0.83***    | 0.84***     | -0.76***                    | -0.79***            | -0.56*    | -0.78***  |
| FMA-UE         | 0.66**     | 0.61**      | -0.48*                      | -0.45*              | -0.23     | -0.47*    |

N-TB, number of transferred blocks; N\textsubscript{ZC}-ACC, number of zero-crossings of the moving acceleration; N\textsubscript{ZC}-DRF, number of zero-crossings of the derivative of releasing force; PLR, ratio of the path length and linear length in a block transfer; DDP, distance between the barrier partition and the drop position of a block; ARAT, Action Research Arm Test; FMA-UE, Fugl-Meyer Assessment for the Upper Extremity. A single asterisk (*), double asterisks (**) and triple asterisks (***), indicate significant differences between the two age groups with $p<0.05$, $p<0.01$ and $p<0.001$, respectively.

Test–retest reliability for BBT and VBBT

The results of the test-retest analysis on the BBT and VBBT for healthy subjects (n = 95) are shown in Table 3. The N-TB in VBBT (ICC = 0.75) exhibited better consistency than that in the BBT (ICC = 0.62). The test-retest correlations were good for the kinematic measurement N\textsubscript{ZC}-ACC and N\textsubscript{ZC}-DRF (ICC = 0.78 and 0.80, respectively) and were moderate for PLR and DDP (ICC = 0.68 and 0.58, respectively).

Table 3
Test-retest reliability for BBT and VBBT parameters in healthy subjects (n = 95)

|                    | Test mean (SD) | Retest mean (SD) | ICC (95%CI) | SEM |
|--------------------|----------------|------------------|-------------|-----|
| N-TB (BBT)        | 72.20 (9.37)   | 73.74 (9.58)     | 0.62 (0.48–0.73) | 5.85 |
| N-TB (VBBT)       | 30.49 (7.92)   | 34.95 (10.57)    | 0.75 (0.65–0.83) | 4.79 |
| N\textsubscript{ZC}-ACC (VBBT) | 2.88 (1.24)    | 2.53 (1.09)      | 0.78 (0.68–0.84) | 0.56 |
| N\textsubscript{ZC}-DRF (VBBT) | 5.21 (1.84)    | 4.65 (1.79)      | 0.80 (0.71–0.86) | 0.82 |
| PLR (VBBT)        | 1.21 (0.16)    | 1.16 (0.15)      | 0.68 (0.56–0.78) | 0.087|
| DDP (VBBT)        | 44.43 (17.00)  | 39.15 (15.83)    | 0.58 (0.43–0.70) | 0.34 |

N-TB, number of transferred blocks; N\textsubscript{ZC}-ACC, number of zero-crossings of the moving acceleration; N\textsubscript{ZC}-DRF, number of zero-crossings of the derivative of releasing force; PLR, ratio of the path length and linear length in a block transfer; DDP, distance between the barrier partition and the drop position of a block; SD, standard deviation; ICC, Intraclass correlation coefficient; CI Confidence interval; SEM, Standard error of mean.

Subjective preference for BBT and VBBT

As shown in Fig. 5, the patients scored 3.7 ± 1.6 for the difference of the subjective preference in performing the BBT and VBBT. Specifically, they gave higher scores for enjoyment in the VBBT than in the BBT ($6.0 ± 1.5$ and $4.1 ± 0.9$, respectively, $p<0.05$) and gave higher scores for effort in the VBBT than in the BBT ($3.7 ± 1.8$ and $1.7 ± 1.0$, $p<0.05$). There were no significant differences in scores for understandability ($6.2 ± 1.5$ and $6.9 ± 0.3$, $p>0.05$), attraction ($6.6 ± 1.3$ and $5.6 ± 1.1$, $p>0.05$), relaxation ($4.6 ± 2.5$ and $5.2 ± 2.5$, $p>0.05$) and tiredness ($1.6 ± 1.1$ and $1.3 ± 0.7$, $p>0.05$) between the VBBT and BBT.
Discussion

The BBT has been considered to be an effective tool to evaluate gross manual dexterity in post-stroke patients. However, due to the simplicity of the measurement, the results from the BBT cannot represent specific impairments in motor function. Our study aimed to virtualize the BBT using a VR display combined with a haptic device to improve an assessment of UE function in more details. Specifically, four kinematic parameters, including \( N_{ZC} \)-ACC, \( N_{ZC} \)-DRF, PLR and DDP, were used to reflect motion smoothness, hand dexterity and motion efficiency during task performance.

Compare to the BBT performance, the number of transferred blocks (N-TB) in VBBT presented a stronger age-related correlation (Fig. 2 and Fig. 3). As the age increased, N-TB completed by the subjects significantly decreased. This might be due to a better learning ability to use a new device for young people than for senior people. Therefore, we divided the subjects into three groups (young, middle-aged and senior) to characterize the subjects’ performances. Compared to the BBT, the VBBT can provide several kinematic parameters, in addition to the number of transferred blocks, to reflect a patient’s specific impairment in UE function in more details. The kinematic parameters refer to \( N_{ZC} \)-ACC, \( N_{ZC} \)-DRF, PLR, and DDP. Normative ranges of the kinematic parameters were determined by 95% measurements of healthy-subject performances. Impairments in UE function in a stroke patient can be identified when his/her measurements fall outside of the normative ranges. The number of zero-crossings collected by the haptic device has been regarded as the number of submovements involved in a block-transfer movement, which has been commonly accepted to reflect movement smoothness \[42, 43\]. In our experiment, the \( N_{ZC} \)-ACC is an indicator of the smoothness of the trunk, shoulder, elbow and wrist of a patient while transferring a block; the \( N_{ZC} \)-DRF refers to a measurement of the smoothness of the hand and fingers during grasping and releasing a block \[34\]. Since the PLR in the VBBT was invalid to some well-accepted clinical assessment scales, we used the DDP to reflect the efficiency of UE movement. According to the kinematic measurements of the patients (Supplementary Table), the therapists could individually design rehabilitative training strategies for them. For example, patients P2-P5, P8 and P11 with \( N_{ZC} \)-ACC and \( N_{ZC} \)-DRF measures falling outside the normative ranges but DDP in the normative ranges could be suggested to focus their treatment on enhancing their UE movement smoothness. For patients P1 and P13, all their measurements fell outside the normative ranges, and the therapist could decide to improve both the smoothness and efficiency of their UE function.

The VBBT is a task-oriented assessment that is also concerned with patients’ motivation \[49–52\]. Based on the Intrinsic Motivation Inventory, we designed several questions to examine the subjective preference of the patients. The results showed that most patients preferred to perform the assessment using the VBBT rather than the BBT, due to the enjoyment of the immersive environment and effort needed for task performance. Interestingly, some patients thought that the use of the device in the VBBT required more effort than that in BBT, which could have motivated their task completion.

We compared the measurements on the BBT and VBBT with the scores of clinical assessment scales to validate our current design. To be expected, the measurements in BBT were strongly correlated with ARAT and FMA-UE scores. There is also a strong correlation between the parametric measurements in VBBT and ARAT. However, FMA-UE scores showed weak and moderate correlations with the VBBT measurements (Table 2). The ARAT, related to the International Classification of Functioning, Disability and Health (ICF) Activity level \[18\], is usually used for concurrent validity of UE function assessment with novel technology \[35, 39\]. The FMA-UE has also been commonly used for concurrent validity. However, the FMA is related to the ICF Body Function/Body Structure level \[18\], which was considered a complement for the designed objective assessment \[44, 45\]. In our experiment, two patients (P2 and P15) exhibited apparent differences between the VBBT and FMA-UE measures. For example, P2 was unable to perform scapula adduction completely on her affected side and had limited shoulder flexion up to 140°. P15 had limited range of motion (ROM) of wrist extension for 50° and limited ROM of wrist supination for 45°. These physical impairments or compensatory movements were the result of low FMA-UE scores for the two patients but made little difference in BBT and ARAT performance. We also found
that the PLR was weakly or moderately correlated with the clinical scale scores ($0.23 \leq |r| \leq 0.56$; Table 2). This is possibly due to individual habits in transferring movements. Therefore, the PLR was not recommended as an indicator in a patient's UE functional assessment. Furthermore, we conducted a test-retest analysis to evaluate the reliability of the VBBT assessments (Table 3). A higher ICC was found in the N-TB in VBBT than that in the BBT. In addition, the ICCs tended to be between moderate and good for the four kinematic parameters in the VBBT, including N$_{ZC}^{-\text{ACC}}$, N$_{ZC}^{-\text{DRF}}$, PLR and DDP. Additionally, the subjects' VBBT performances exhibited significant improvement between test and retest sessions, though all the subjects were given a familiarization session up to 15 minutes to minimize the learning effect [46–48]. In fact, most subjects finished the familiarization session within 3 minutes, since they thought they were fully familiarized with how to operate the haptic device. However, the test-retest results indicated that this was not the case. Therefore, more complex virtual tasks should be included in the familiarization session to reduce the learning effect in the VBBT assessment. It should be noted that the test-retest analysis was available only for healthy subjects, not the patients. All patients in this study were recruited from rehabilitation units of the hospital. Their UE motor function probably improved in the 4-week interval between test and retest performance since they were receiving physical therapy or occupational therapy during that time. Above all, the VBBT was concurrently validated for task-oriented assessment of UE motor function and had good test-retest reliability.

There are still some limitations in the current study. One limitation refers to the "floor effect" of the VBBT. The patients could operate the device and perform the virtual task only if their UE function reached Brunnstrom stage IV or later. In fact, most patients tend to conduct rehabilitation at home when they have recovered major functions of the UE. For some patients who are unable to independently operate the device, we suggest providing some support for their arm during VBBT performance. Another limitation is the higher price of the haptic device for VBBT system than the BBT instrument, which makes it impossible to be commercially available for some patients. The application of the VBBT is oriented to home-based rehabilitation. However, this does not mean that every patient must buy it. Patients can rent the device from the rehabilitation unit. When they regain motor function and can maintain recovery, they will send back the device to the rehabilitation unit. Therefore, the patients do not need to afford too much for the use of the VBBT.

**Conclusion**

We designed the VBBT with a haptic device to improve BBT for assessing UE motor function in patients with chronic stroke. The performances of healthy subjects were collected to establish normative ranges in different age groups. Deficiencies in a patient's UE motor function could be determined when his/her measurements were not within the normative ranges. The concurrent validity and test-retest evaluations were examined to indicate that the VBBT is an effective and reliable task-oriented assessment. Besides the number of transferred blocks, the VBBT can provide clinically validated kinematic parameters, including N$_{ZC}^{-\text{ACC}}$, N$_{ZC}^{-\text{DRF}}$ and DDP, to reflect patients' specific impairments in UE function, including the motion smoothness, hand dexterity and motion efficiency. Additionally, the patients showed a significant preference for the VBBT due to the enjoyment and completion effort, which would improve their motivation to participate in the unsupervised assessment for home-based rehabilitation.

**Abbreviations**

UE: upper extremity, FMA: Fugl-Meyer Assessment, ARAT: Action Research Arm Test, BS: Brunnstrom stage on the hemiplegic side, VR: virtual reality, BBT: Box and Block Test, VBBT: Virtual Box and Block Test, N-TB: Number of transferred blocks, N$_{ZC}^{-\text{ACC}}$: Number of zero-crossings of the moving acceleration, N$_{ZC}^{-\text{DRF}}$: Number of zero-crossings of the derivative of releasing force, PLR: Ratio of the path length and linear length in a block transfer, DDP: Distance between the barrier partition and the drop position of a block, ICC: Intraclass Correlation Coefficient, CI: Confidence Interval, SEM: Standard Error of Mean.
Declarations

Ethics approval and consent to participate

The current study adhered to the tenets of the Declaration of Helsinki, and ethical approval was approved by Beihang University (BM20180017). Each participant was provided written informed consent prior to their participation in the study. A signed informed consent statement was received from each participant.

Consent for publication

Not applicable.

Availability of data and materials

All the data and materials in the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the National Key Research and Development Plan of China under Grant 2020YFC2004200, 2016YFC1102203 and 2016YFC1101100, the National Natural Science Foundation of China under Grant 11421202, Beijing Municipal Natural Science Foundation: Haidian Original Innovation Joint Fund under Grant L172016 and the Open Project funding from the State Key Laboratory of Virtual Reality Technology and Systems, Beihang University under Grant VRLAB2018T01.

Authors' contributions

YD contributed to designing and conducting the experiment, analyzing the experimental data, and drafting this manuscript. XL contributed to leading this work, securing the funding, guiding the experiment and drafting and editing the manuscript. MT contributing to programing the virtual task, conducting the experiment and interpreting the data. HH contributed to analyzing and interpreting the data. DC contributed to conducting the experiment. YF contributed to leading this work, securing the funding, editing and approving the final manuscript.

Acknowledgements

The authors would like to thank Ms. Jige Dong and Ms. Zhihua Xu for their contribution to recruitment of patients and healthy subjects in this study.

References

1. Ovbiagele B, Goldstein LB, Higashida RT, Howard VJ, Johnston SC, Khavjou OA, Lackland DT, Lichtman JH, Mohl S, Sacco RL et al: Forecasting the future of stroke in the United States: a policy statement from the American Heart
1. Bohannon RW, Smith MB: Interrater reliability of a modified Ashworth scale of muscle spasticity. Phys Ther 1987, 67(2):206-207.

2. Ansari NN, Naghdi S, Arab TK, Jalaie S: The interrater and intrarater reliability of the Modified Ashworth Scale in the assessment of muscle spasticity: limb and muscle group effect. NeuroRehabilitation 2008, 23(3):231-237.

3. Santisteban L, Térémetz M, Bleton JP, Baron JC, Maier MA, Lindberg PGJPO: Upper Limb Outcome Measures Used in Stroke Rehabilitation Studies: A Systematic Literature Review. 2016, 11(5):e0154792.

4. Thompson-Butel AG, Lin G, Shiner CT, McNulty PAJN, repair n: Comparison of three tools to measure improvements in upper-limb function with poststroke therapy. 2015, 29(4):341-348.

5. Krabben T, Molier BI, Houwink A, Rietman JS, Buurke JH, Prange GBJJOn, rehabilitation: Circle drawing as evaluative movement task in stroke rehabilitation: an explorative study. 2011, 8(1):15.

6. Cunningham DA, Potter-Baker KA, Knutson JS, Sankarasubramanian V, Machado AG, Plow EB: Tailoring brain stimulation to the nature of rehabilitative therapies in stroke: a conceptual framework based on their effectiveness. Clinics R, 2017, 72(4):1853-1862.
22. Walker RW, Wakefield K, Gray WK, Jusabani A, Swai M, Mugusi F: Case-fatality and disability in the Tanzanian Stroke Incidence Project cohort. Acta Neurol Scand 2016, 133(1):49-54.

23. Saposnik G, Cohen LG, Mamdani M, Pooyania S, Ploughman M, Cheung D, Shaw J, Hall J, Nord P, Dukelow S et al: Efficacy and safety of non-immersive virtual reality exercising in stroke rehabilitation (EVREST): a randomised, multicentre, single-blind, controlled trial. Lancet Neurol 2016, 15(10):1019-1027.

24. Kwakkel G, Lannin NA, Borschmann K, English C, Ali M, Churilov L, Saposnik G, Weinstein C, van Wegen EEH, Wolf SL et al: Standardized Measurement of Sensorimotor Recovery in Stroke Trials: Consensus-Based Core Recommendations from the Stroke Recovery and Rehabilitation Roundtable. Neurorehabil Neural Repair 2017, 31(9):784-792.

25. Perez-Marcos D, Chevalley O, Schmidlin T, Garipelli G, Serino A, Vuadens P, Tadi T, Blanke O, Millán JdRJon, rehabilitation: Increasing upper limb training intensity in chronic stroke using embodied virtual reality: a pilot study. 2017, 14(1):119.

26. Vourvopoulos A, i Badia SBJJon, rehabilitation: Motor priming in virtual reality can augment motor-imagery training efficacy in restorative brain-computer interaction: a within-subject analysis. 2016, 13(1):69.

27. Gerber LH, Narber CG, Vishnoi N, Johnson SL, Chan L, Duric Z: The feasibility of using haptic devices to engage people with chronic traumatic brain injury in virtual 3D functional tasks. Journal of Neuroengineering and Rehabilitation 2014, 11.

28. Fluet M-C, Lambercy O, Gassert R: Upper limb assessment using a virtual peg insertion test. In: 2011 ieee international conference on rehabilitation robotics: 2011: IEEE; 2011: 1-6.

29. Tobler-Ammann BC, de Bruin ED, Fluet MC, Lambercy O, de Bie RA, Knols RH: 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 2016, 13:8.

30. Liu X, Zhu Y, Huo H, Wei P, Wang L, Sun A, Hu C, Yin X, Lv Z, Fan Y: Design of Virtual Guiding Tasks With Haptic Feedback for Assessing the Wrist Motor Function of Patients With Upper Motor Neuron Lesions. IEEE Trans Neural Syst Rehabil Eng 2019, 27(5):984-994.

31. Mathiowetz V, Volland G, Kashman N, Weber K: Adult norms for the Box and Block Test of manual dexterity. Am J Occup Ther 1985, 39(6):386-391.

32. Cho S, Kim W-S, Paik N-J, Bang HJlccg, applications: Upper-limb function assessment using VBBTs for stroke patients. 2015, 36(1):70-78.

33. Furmanek MP, Schettino LF, Yarossi M, Kirkman S, Adamovich SV, Tunik EJJJon, rehabilitation: Coordination of reach-to-grasp in physical and haptic-free virtual environments. 2019, 16(1):78.

34. Schwarz A, Kanzler CM, Lambercy O, Luft AR, Veerbeek JMJS: Systematic review on kinematic assessments of upper limb movements after stroke. 2019, 50(3):718-727.

35. Hussain N, Sunnerhagen KS, Murphy MAJJJon, rehabilitation: End-point kinematics using virtual reality explaining upper limb impairment and activity capacity in stroke. 2019, 16(1):82.

36. Yozbatiran N, Der-Yeghiaian L, Cramer SCJN, repair n: A standardized approach to performing the action research arm test. 2008, 22(1):78-90.

37. Müller G: [Movement therapy in hemiplegia]; 1970.

38. Marshal, F., Folstein, and, Susan, E., Folstein, and, Paul, R.: "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician.

39. Germanotta M, Cruciani A, Pecchioli C, Loreti S, Spedicato A, Meotti M, Mosca R, Speranza G, Cecchi F, Giannarelli G et al: Reliability, validity and discriminant ability of the instrumental indices provided by a novel planar robotic device
for upper limb rehabilitation. J Neuroeng Rehabil 2018, 15(1):39.
40. Koo TK, Li MY. Jocm: A guideline of selecting and reporting intraclass correlation coefficients for reliability research. 2016, 15(2):155-163.
41. Weir JP. JoS, Research C: Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. 2005, 19(2):231-240.
42. Rohrer B, Fasoli S, Krebs HI, Hughes R, Volpe B, Frontera WR, Stein J, Hogan N: Movement smoothness changes during stroke recovery. J Neurosci 2002, 22(18):8297-8304.
43. Bardorfer A, Munih M, Zupan A, Primozic AJ. To: Upper limb motion analysis using haptic interface. 6(3):0-260.
44. Colombo R, Pisano F, Micera S, Mazzone A, Delconte C, Carrozza MC, Dario P, Minuco G: Robotic techniques for upper limb evaluation and rehabilitation of stroke patients. IEEE Trans Neural Syst Rehabil Eng 2005, 13(3):311-324.
45. Lin KC, Chuang LL, Wu CY, Hsieh YW, Chang WY: Responsiveness and validity of three dexterous function measures in stroke rehabilitation. J Rehabil Res Dev 2010, 47(6):563-571.
46. Davis GA, Purcell LK: The evaluation and management of acute concussion differs in young children. Br J Sports Med 2014, 48(2):98-101.
47. Little CE, Emery C, Black A, Scott SH, Meeuwisse W, Nettel-Aguirre A, Benson B, Dukelow S: Test-retest reliability of KINARM robot sensorimotor and cognitive assessment: in pediatric ice hockey players. J Neuroeng Rehabil 2015, 12:78.
48. Elaine Little C, Woollacott M: Effect of attentional interference on balance recovery in older adults. Exp Brain Res 2014, 232(7):2049-2060.
49. Mihelj M, Novak D, Milavec M, Zihrl J, Munih MJ. PT, Environments V: Virtual Rehabilitation Environment Using Principles of Intrinsic Motivation and Game Design. 2012, 21(1):1-15.
50. Colombo R, Pisano F, Mazzone A, Delconte C, Micera S, Carrozza MC, Dario P, Minuco G: Design strategies to improve patient motivation during robot-aided rehabilitation. J Neuroeng Rehabil 2007, 4:3.
51. Novak D, Nagle A, Keller U, Riener R: Increasing motivation in robot-aided arm rehabilitation with competitive and cooperative gameplay. J Neuroeng Rehabil 2014, 11:64.
52. Gorsic M, Cikajlo I, Novak D: Competitive and cooperative arm rehabilitation games played by a patient and unimpaired person: effects on motivation and exercise intensity. J Neuroeng Rehabil 2017, 14(1):23.

Figures
Figure 1

The haptic device, the VBBT task environment and the measurements in VBBT performance in the current experiment. a) The seven-degree-of-freedom (7-DoF) haptic feedback device (Omega.7, Force Dimension Inc., Switzerland) used to provide human-VBBT interaction. b) VBBT task that requires users to move a virtual block from one compartment of a virtual box to the other in one trial. c) A schematic diagram of the measurements in VBBT performance, including a block trajectory and the grasping force from a start position (SP) to the end position (EP) in one block transfer. [Double column]

Figure 2

The numbers of transferred blocks (N-TB) in BBT and VBBT performed by the healthy subjects and patients in different age groups. a) N-TB in BBT performed by the healthy subjects and the patients in the young, middle-aged and senior groups. b) N-TB in VBBT performed by the healthy subjects and the patients in the young, middle-aged and senior groups. In each boxplot, the 2.5th and 97.5th percentiles are shown at the end of the lines, and black diamonds indicate the
outliers. A single asterisk (*), double asterisks (**) and triple asterisks (***)) indicate significant differences between the two age groups with \( p < 0.05 \), \( p < 0.01 \) and \( p < 0.001 \), respectively. [Double column]

**Figure 3**

Correlation and linear regression analysis on the numbers of transferred blocks (N-TB) in BBT and VBBT performed by the healthy subjects with the age distributions. a) A Spearman correlation on N-TBs between the BBT and VBBT in the healthy subjects \( (r = 0.42, p < 0.001) \). b) A linear regression between N-TB in VBBT in the healthy subjects and their ages \( (R^2 = 0.57, p < 0.001) \). c) A linear regression between N-TB in BBT in the healthy subjects and their ages \( (R^2 = 0.16, p < 0.001) \). The green, blue and red diamonds indicate the subjects in the young, middle-aged and senior groups, respectively. [Single column]
Figure 4

The kinematic parameters in the VBBT performed by the healthy subjects and patients. a1) Two typical curves of block-moving acceleration vs. normalized time in one trial performed by a healthy subject and a patient. The blue dots and the red asterisks indicate the zero-crossings of the block-moving acceleration for the healthy subject and patient, respectively. a2) The numbers of zero-crossings of the block-moving acceleration performed by the healthy subjects and patients in the young, middle-aged and senior groups. b1) Two typical curves of releasing-force derivative vs. normalized time in one trial performed by a healthy subject and a patient. The blue dots and the red asterisks indicate the zero-crossings of the releasing-force derivative for the healthy subject and patient, respectively. b2) The numbers of zero-crossings of the releasing-force derivative performed by the healthy subjects and patients in the young, middle-aged and senior groups. c1) Two typical curves of the ratio of the path distance and linear distance (PLR) vs. normalized Y position (PY) in the VBBT performed by a healthy subject and a patient. The belts indicate 2-sided measurements (2.5%–97.5%) of the PLR performance. c2) The PLRs performed by the healthy subjects and patients in the young, middle-aged and senior groups.
d1) Two typical scatter plots of the distance between the barrier partition and the drop position of a virtual block (DDP) performed by a healthy subject and a patient in the VBBT. d2) The DDPs performed by the healthy subjects and patients in the young, middle-aged and senior groups. A single asterisk (*), double asterisks (**) and triple asterisks (***) indicate significant differences between the two age groups with p < 0.05, p < 0.01 and p < 0.001, respectively. [Double column]

![Figure 5](image)

**Figure 5**

The scores of the patients’ subjective preferences for the BBT and VBBT. The scores for understandability, enjoyment, attraction, relaxation, effort and tiredness for the BBT and VBBT are 6.2 ± 1.5 and 6.9 ± 0.3, 6.0 ± 1.5 and 4.1 ± 0.9, 6.6 ± 1.3 and 5.6 ± 1.1, 4.6 ± 2.5 and 5.2 ± 2.5, 3.7 ± 1.8 and 1.7 ± 1.0, 1.6 ± 1.1 and 1.3 ± 0.7, respectively. A single asterisk (*) indicates a significant difference (p < 0.05) between the subjective preferences for the BBT and VBBT. [One and a half column]

**Supplementary Files**

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementarytable.docx
- Supplementaryvideo.mp4