

RESEARCH ARTICLE

Reliability of the thumb localizing test and its validity against quantitative measures with a robotic device in patients with hemiparetic stroke

Eri Otaka¹, Yohei Otaka^{1,2*}, Shoko Kasuga^{3,4}, Atsuko Nishimoto¹, Kotaro Yamazaki¹, Michiyuki Kawakami¹, Junichi Ushiba^{3,5}, Meigen Liu¹

1 Department of Rehabilitation Medicine, Keio University School of Medicine, Shinjuku-ku, Tokyo, Japan, **2** Department of Rehabilitation Medicine I, School of Medicine, Fujita Health University, Toyoake, Aichi, Japan, **3** Keio Institute of Pure and Applied Sciences (KiPAS), Keio University, Yokohama, Kanagawa, Japan, **4** Graduate School of Science and Technology, Keio University, Yokohama, Kanagawa, Japan, **5** Department of Biosciences and Informatics, Faculty of Science and Technology, Keio University, Yokohama, Kanagawa, Japan

* otaka119@mac.com



OPEN ACCESS

Citation: Otaka E, Otaka Y, Kasuga S, Nishimoto A, Yamazaki K, Kawakami M, et al. (2020) Reliability of the thumb localizing test and its validity against quantitative measures with a robotic device in patients with hemiparetic stroke. PLoS ONE 15(7): e0236437. <https://doi.org/10.1371/journal.pone.0236437>

Editor: François Tremblay, University of Ottawa, CANADA

Received: March 8, 2020

Accepted: July 6, 2020

Published: July 24, 2020

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0236437>

Copyright: © 2020 Otaka et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files.

Abstract

Objectives

To examine the inter-rater reliability of the thumb localizing test (TLT) and its validity against quantitative measures of proprioception.

Methods

The TLT was assessed by two raters in a standardized manner in 40 individuals with hemiparetic stroke. Inter-rater reliability was examined with weighted Kappa. For the quantitative measures, a bimanual matching task in a planar robotic device was performed. Without vision, each participant moved the unaffected hand to the perceived mirrored location of the affected hand, which was passively moved by the robot. Three measures were taken after 54 trials: *Variability*, trial-to-trial variability of the mirrored-matched locations; *Area*, the ratio of the area enclosed by the active hand relative to the passive hand; and *Shift*, systematic shifts between the passive and active hands. The correlation between the TLT and each robotic measure was examined with Spearman's rank correlation coefficient.

Results

The overall weighted kappa of the TLT was 0.84 ($P < 0.001$). The TLT correlated highly with *Area* ($r = -0.71$, $P < 0.001$) and moderately with *Variability* ($r = 0.40$, $P = 0.011$). No significant correlation was found between the TLT and *Shift*.

Conclusions

The TLT had a high inter-rater reliability, and was validated against quantitative measures of proprioception reflecting the perceived area of movement and variability of the limb location.

Funding: This study was supported in part by collaborative research funds from Toyota Tsusho Corporation (<https://www.toyota-tsusho.com>). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: This study was supported in part by collaborative research funds from Toyota Tsusho Corporation. M.L., M.K. and J.U. are the founding scientists of Connect Inc., a commercial company for the development of rehabilitation devices since May 2018. They have received a salary from Connect Inc., and have held shares in Connect Inc. They currently hold managerial positions at Connect Inc. These conditions were disclosed to the Universities, and were approved. Connect Inc. does not have any relationship with the present study. This does not alter our adherence to PLOS ONE policies on sharing data and materials.

Introduction

Proprioception is necessary for the control of limb posture, and the coordination of multijoint movements [1–6]. In stroke patients, proprioceptive deficits are quite common [7, 8] and known to be partly related to motor recovery [9, 10]. As its longitudinal process of recovery has been partly revealed recently [8, 11, 12] and novel rehabilitative approaches are sought more than ever in clinical practice, indispensability for robust assessment tools of proprioception is increasing to accurately evaluate its improvement with progress of time or by any intervention.

Conventionally, the thumb localizing test (TLT) [13–15] or similar techniques [9, 16] are widely used for the assessment of proprioception for patients with stroke, both in clinical and research settings. The TLT is a test for “limb localization” and is clinically useful in that it can examine invisible perceptual deficits using a visible motor task and be easily performed at the patient’s bedside, although it is a four-point ordinal scale and thus disadvantageous as a quantification tool in detecting subtle deficits or changes over time compared with continuous scales. In the development of the TLT, Hirayama et al. reported that the TLT showed greater frequency of abnormality than other physical examinations for proprioceptive sensation [15], and the TLT deficits had strong categorical relationship with joint position and movement test, and tactile cutaneous localization test [14]. However, reliability of the TLT has not been precisely shown. Furthermore, validity of the TLT has yet to be revealed because there has been no other conventional assessment of proprioception that is reliable and quantitative enough to be regarded as a reference.

Several groups have designed reliable and objective procedures to quantify proprioception [17–19]. Among these, the arm position matching task [20] provided by the KINARM Exoskeleton robotic device (Kinarm, Kingston, ON, Canada) [21] is a robotic assessment which describes kinematic characteristics of the individual’s mirror-matched arm movement, visualizing and quantifying proprioception using a motor task same as the TLT. The measures derived from this robotic task have been shown to be adequately reliable [20] and correlated with the measures of activities of daily living and dexterity in patients with subacute stroke [22], enough to be considered as a reference in a validation study for other assessments. Regarding the TLT, only categorical relationships have been shown with some types of quantitative measures driven by the robotic device [20, 22] in the mixed group of healthy adults and stroke patients. However, no attempt has been made to examine the correlation between the TLT and quantitative measures driven from the device.

To confirm the robustness of the TLT as a clinical measure of proprioception, the present study aimed (1) to examine the inter-rater reliability of the TLT, and (2) to validate the TLT by examining correlation with the quantitative measures derived from the exoskeleton robotic device.

Materials and methods

Participants

The recruited participants included 40 patients with chronic stroke who visited the Department of Rehabilitation Medicine at Keio University Hospital for the rehabilitation of upper extremity paresis from July 2013 to January 2015. Patients with the following criteria were included: more than 150 days since stroke onset; stroke in single cerebral hemisphere; no clinically obvious cognitive deficits; living independently; absence of pain in the paretic upper limb; no impairments in the non-paretic upper limb; ability to maintain sitting position without difficulty; and adequate range of motion in the upper extremity to perform tasks in the

KINARM. Patients with history of epilepsy, implants (pacemakers, shunts, or clipping), and those who could not perform the KINARM tasks owing to pain and/or contractures were excluded.

The study protocol was approved by the institutional ethics committee (Keio University School of Medicine Ethics Committee #20120070). All participants gave written informed consent prior to their participation, according to the Declaration of Helsinki.

Clinical assessments

To evaluate the participants' upper limb impairments, we assessed the affected arm using the Fugl-Meyer Assessment, the part A score of the upper limb section to quantify motor function (0 (worst) - 36 (best)) [23], and the Modified Ashworth Scale for elbow flexors to grade spasticity (0 (no spasticity), 1, 1+, 2, 3, 4 (rigid)) [24].

Thumb localizing test

In accordance with Hirayama et al. [13, 14], we adhered to the following procedure: an examiner positioned the affected upper limb of the participant at a certain position, and asked participants to pinch the tip of the thumb with the unaffected fingers. For the standardization, three limb positions were defined (Fig 1), and all participants were assessed in the same order. Before each limb position was fixed, the examiner moved the limb multi-directionally to prevent the subject from guessing the limb position based on the starting position. The participants performed the procedure twice, with vision occluded condition (with a blindfold) and with vision restored condition (without a blindfold). This was to analyze the effect of vision, as the importance of visual information in compensating proprioceptive deficit is clinically well known [4, 25].

To score the TLT, all performances were recorded on videotape and were independently rated afterwards by two physiatrists; EO, with 6 years of clinical experience and YO, with 18 years of clinical experience. Three physiatrists, EO, YO, KY, with 5 years of clinical experience, and an occupational therapist, AN, with 12 years of clinical experience, served as examiners. As the descriptions of the TLT ratings are slightly different in two papers written by the original authors [13, 15], we adopted the rating of the more recent paper [13]: 0, normal (the reaching limb can go straight and quickly toward the tip of the fixed thumb and pinch it easily); 1, slightly impaired (the fingers are brought several centimeters apart from the fixed thumb and then corrected immediately, or the reaching limb sways on the way); 2, moderately impaired (the reaching limb seeks around the fixed thumb for more than several centimeters, touches the thumb by chance, or reaches the tip of the thumb by tracing from the fixed hand or other fingers); 3, severely impaired (cannot reach the fixed limb at all, or the reaching limb touches the fixed arm by chance and reaches the tip of the thumb by tracing the skin).

Quantitative assessment of proprioception with the robotic device

Robotic device. The KINARM Exoskeleton was used for robotic assessment (Fig 2A). This is a planar robotic device, which has one degree-of-freedom for the shoulder joint (horizontal adduction/abduction) and one for the elbow joint (flexion/extension). It allows free horizontal arm movement, and can monitor shoulder and elbow motion [21]. Participants were initially seated in a modified wheelchair base with each arm set in a trough so that the arms were approximately in the same plane as the shoulder (80–90° abducted), the examiner then slid the whole wheelchair into a planar monitor. Visual feedback or occlusion can be given on the virtual reality display just above the plane of the arms.

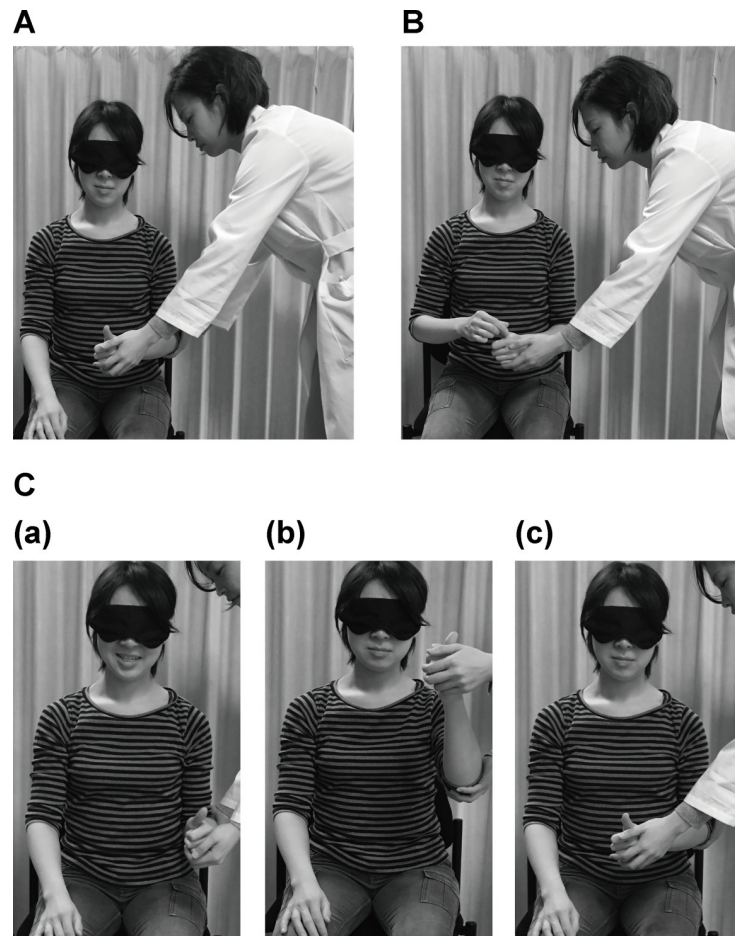


Fig 1. Procedure of the thumb localizing test (TLT). A. The reaching limb is on the ipsilateral knee at the start of the test. B. Participants are instructed to pinch the opposite thumb. C. The limb position of the fixed limb: (a) The forearm in the neutral position, with the elbow at 90° of flexion and the shoulder at 0° of flexion. (b) The forearm in the neutral position, with the elbow at 90° of flexion and the shoulder flexed so that the thumb is at the same level as the mouth. (c) The forearm in the neutral position, with the elbow at 90° of flexion and the shoulder internally rotated so that the thumb is over the midline of the trunk. The individuals pictured in this figure have provided written informed consent (as outlined in PLOS consent form) to publish their image alongside the manuscript.

<https://doi.org/10.1371/journal.pone.0236437.g001>

Assessment of proprioception. The arm position matching task is a bimanual matching task [20]. The central target, which is the starting position, was set such that the shoulder was at approximately 30° of horizontal abduction and the elbow at 90° of flexion. The subject was instructed to relax the affected arm (passive limb) and allow the robot to move the arm in the horizontal plane to adopt a certain position. The subject was asked to then move the unaffected arm (active limb) to the perceived mirrored location. Starting from the preceding position, each participant performed the matching tasks consecutively for nine spatial targets separated by 10 cm for each block, and six such blocks for a session (total of 54 trials in a session); the order of the target positions in each block was pseudo-randomized. Similar to the TLT, all the participants completed the session twice: first with a shield above the arms so that vision was occluded, and then repeated the task with vision restored (Fig 2B). It took approximately 10–15 minutes per session, and all sessions were usually performed consecutively on the same day. Examiners routinely advised participants to take enough time and rest between the testing sessions to fully concentrate on the task, and most participants needed no rest or only a few

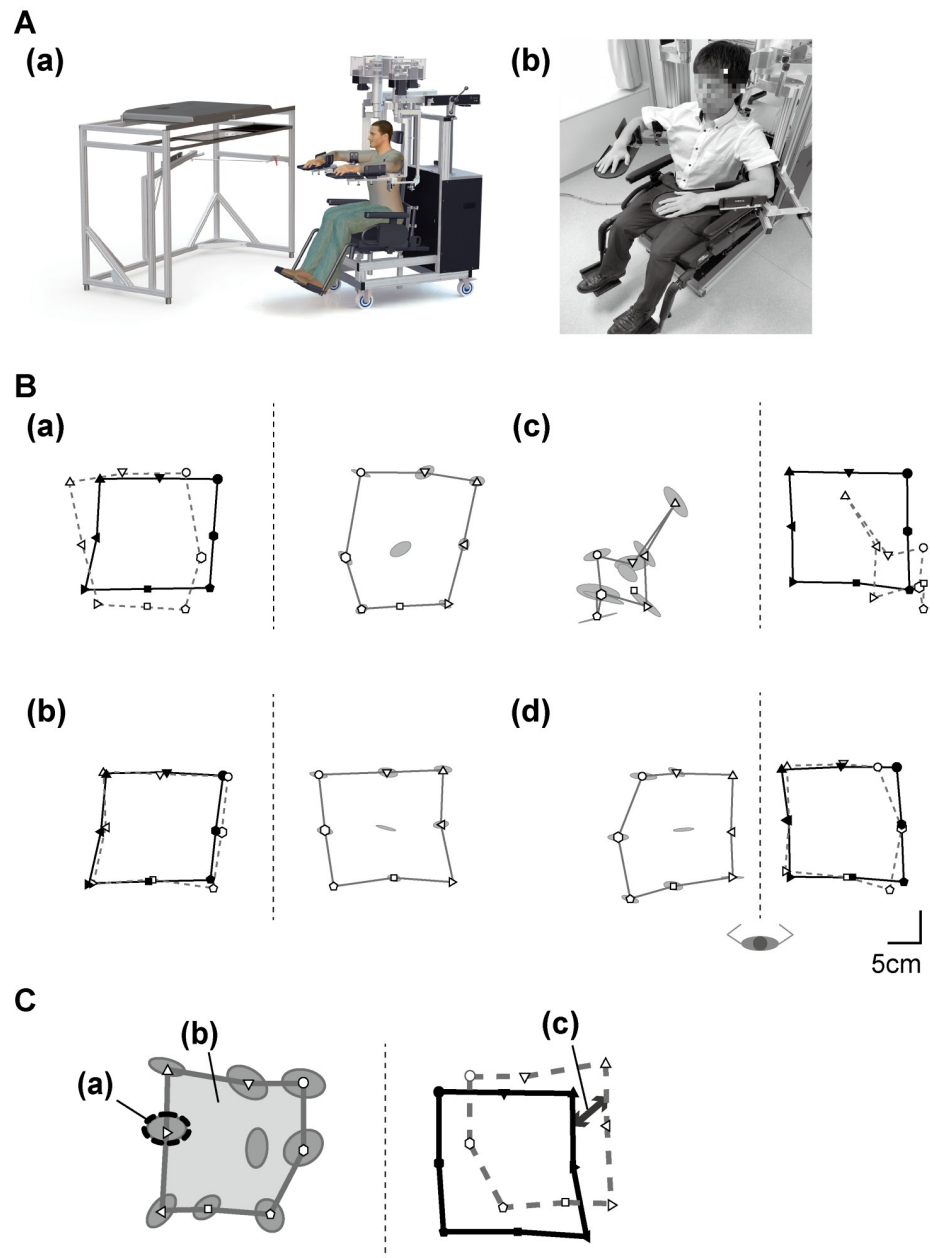


Fig 2. The KINARM Exoskeleton robotic device and measures of the arm position matching task. A. Overview of the robotic apparatus. (a) An overall cartoon depiction (courtesy of Kinarm Ltd., Kingston, Ontario). (b) Frontal view of the wheelchair base and exoskeleton arms. In performing the arm position matching task, the central target, which is the starting position, was set such that the shoulder was at approximately 30° of horizontal abduction and the elbow at 90° of flexion. B. Schema of the arm position matching task. Filled symbols (outlined with solid black lines) represent the positions where the robot moves the passive arm. Open symbols (outlined with solid gray lines) represent where the participants matched the perceived mirrored position with the active arm. The performance of the active arm is mirrored on the passive side using gray dashed lines. Ellipses around targets demonstrate variability of six matching attempts and represent one standard deviation. (a) A typical left-hemiparetic patient with the thumb localization test (TLT) score 0 displays little variability and is relatively accurate in matching the arm position with vision occluded. (b) In the same subject as (a), the matching is more accurate with vision restored. (c) A typical right-hemiparetic participant with the TLT score 3 shows increased variability, contraction of the workspace, and poor performance in overall matching with vision occluded. (d) In the same participant as (c), the accuracy of the matching is improved with vision restored. C. Illustration of the parameters. (a) Variability; trial-to-trial variability of the active hand, (b) Area; the overall spatial area matched by the active hand, (c) Shift; systematic shifts between the active and passive hands. The individual pictured in this figure has provided written informed consent (as outlined in PLOS consent form) to publish his image alongside the manuscript.

<https://doi.org/10.1371/journal.pone.0236437.g002>

minutes. The following robotic measures, calculated using Eqs (1) to (3), were used for the analyses (Fig 2C). These measures have been confirmed as having good inter-rater reliability in patients with stroke [20].

- *Variability*: This refers to the trial-to-trial variability of the perceived mirrored location by the active hand. This measure is calculated by calculating the standard deviation of the active hand's position for each target location, then calculating the mean of the standard deviations for all target locations in the x-coordinate (*Variability X*) and the y-coordinate (*Variability Y*). A smaller value for this measure indicates less variability among the repeated attempts.

$$\text{Variability} = \sqrt{(\text{Variability } X)^2 + (\text{Variability } Y)^2} \quad (1)$$

- *Area*: This refers to the ratio of the spatial area enclosed by the active hand (representing the perceived workspace of the passive hand) relative to the actual area enclosed by the passive hand. A value of < 1 indicates contraction of the perceived area compared to the actual movement area, while a value of > 1 indicates expansion of the perceived area. A value of 1 indicates a perfect match.

$$\text{Area} = \text{area}_{\text{active}} / \text{area}_{\text{passive}} \quad (2)$$

- *Shift*: This indicates the constant bias between the active and passive hands. This measure is calculated by finding the mean error between the active and passive hands for each target location, then calculating the mean of mean errors for all target locations in the x-coordinate (*Shift X*) and the y-coordinate (*Shift Y*). A smaller value indicates less bias in the horizontal plane.

$$\text{Shift} = \sqrt{(\text{Shift } X)^2 + (\text{Shift } Y)^2} \quad (3)$$

Analysis. The TLT with vision restored was not considered for the analysis as all participants obtained the normal score (0). The TLT score with vision occluded was adopted for further analysis. Quadratically weighted Kappa [26] was used to examine the inter-rater reliability of the TLT. The strength of agreement with kappa statistics were interpreted as follows: < 0.00 , poor; $0-0.20$, slight; $0.21-0.40$, fair; $0.41-0.60$, moderate; $0.61-0.80$, substantial; and $0.81-1.00$, almost perfect agreement [27]. Any disagreements in the scores were discussed by the two raters and rescored. The median value of the scores in the three positions was selected for the following analysis.

Prior to examining the correlation between the TLT and the robotic measures, each robotic measure with vision occluded and vision restored conditions was compared using the Wilcoxon matched-pairs signed-ranks test to ascertain whether or not the values significantly deteriorated without visual cues in the same manner as the TLT. Then, the correlation between each robot-driven measure with vision occluded and the TLT score was investigated using the Spearman's rank correlation coefficients. The 95% confidence interval of the Spearman's rank correlation coefficient was calculated based on Fisher's transformation. The strength of the correlation coefficients was interpreted as described by Guilford [28]: $0.0-0.2$, slight; $0.2-0.4$, low; $0.4-0.7$, moderate; $0.7-0.9$, high; and $0.9-1.0$, very high correlation. In addition, the values

Table 1. Participant characteristics.

Variables		(N = 40)
Age, mean \pm SD (range)		47.6 \pm 12.2 (18–78)
Sex, male/female		25/15
Type of stroke	Cerebral infarction	16
	Cerebral hemorrhage	23
	Subarachnoid hemorrhage	1
Days after the onset of hemiparesis, median (range)		486.5 (164–6456)
Part A score of the Fugl-Meyer Assessment of Upper Extremity on the affected side, median (range)		22 (5–35)
Modified Ashworth scale for elbow flexors in the affected arm, n in [0, 1, 1+, 2, 3, 4]		[8, 14, 17, 1, 0, 0]
Thumb localizing test, n in [0, 1, 2, 3]		[16, 11, 8, 5]

<https://doi.org/10.1371/journal.pone.0236437.t001>

of the robotic measure that correlated best with the TLT were compared between the TLT rating groups (0 to 3) using Kruskal-Wallis rank test with Dunn's post-hoc test. Statistical analyses were performed with STATA/SE 13.1 (StataCorp., Texas, USA). Any *P*-value less than 0.05 was considered statistically significant. Through an a priori power analysis for calculating correlation coefficients using G*Power 3.1 [29], we estimated at least a sample size of 29 in total to provide effect size of 0.5, power of 80%, and type 1 error of 0.05.

Results

The clinical characteristics of the participants are presented in Table 1. Participants with a wide range of motor impairments were included. Spasticity of the participants ranged from mild to moderate. Although all the participants obtained the normal score (0) in the vision restored condition, the TLT determined 24 out of 40 participants (60%) as having proprioceptive impairment with vision occluded condition. The overall weighted kappa of the TLT scores between the two raters was 0.84 ($P < 0.001$), indicating almost perfect agreement (Table 2).

On comparing each of the robotic measures between the two visual conditions, *Variability* and *Shift* were significantly larger with vision occluded than with vision restored (*Variability*: median value 5.76 cm vs 3.34 cm, $P < 0.001$; *Shift*: median value 6.35 cm vs 3.81 cm, $P < 0.001$). *Area* was significantly smaller with vision occluded than with vision restored (median value 0.56 vs 0.85, $P = 0.010$) (Figs 3 and 4).

As shown in Table 3, in the correlational analyses between the TLT and the robotic measures, *Area* correlated most strongly with the TLT with vision occluded condition. This correlation was high and negative. The scatterplot of *Area* with the two visual conditions (Fig 4) also implies the decrease of *Area* in the patients with severe impairment (TLT = 2,3). *Variability* showed moderate correlations with the TLT. The correlation between *Shift* and the TLT was not significant.

Table 2. Scores of the thumb localizing test and quadratically weighted kappa.

	Scores of rater A, n in [0, 1, 2, 3]	Scores of rater B, n in [0, 1, 2, 3]	Quadratically weighted kappa	<i>P</i> values
Position 1	[15,17,1,7]	[15,15,5,5]	0.86	< 0.001
Position 2	[15,12,5,8]	[17,7,11,5]	0.86	< 0.001
Position 3	[20,8,9,3]	[17,7,11,5]	0.81	< 0.001
Total	[50,37,15,18]	[49,29,27,15]	0.84	< 0.001

Position 1, the forearm in the neutral position, with the elbow at 90° of flexion, and the shoulder at 0° of flexion; Position 2, the forearm in the neutral position, with the elbow at 90° of flexion, and the shoulder flexed so that the thumb is at the same level as the mouth; Position 3, the forearm in the neutral position, with the elbow at 90° of flexion, and the shoulder internally rotated so that the thumb is over the midline of the trunk.

<https://doi.org/10.1371/journal.pone.0236437.t002>

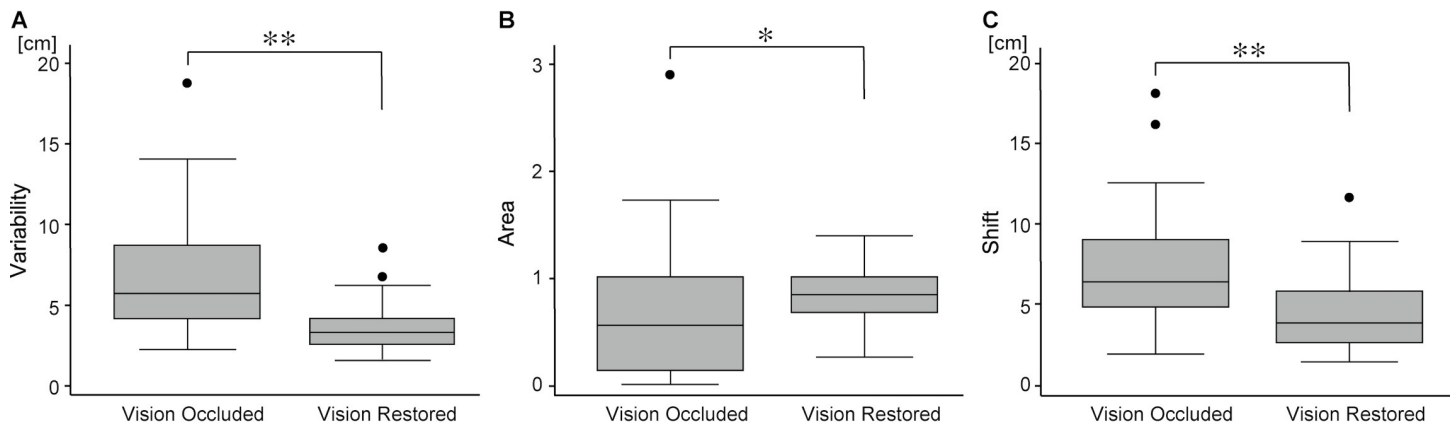


Fig 3. Comparisons between the measures of arm position matching task with vision occluded and vision restored. Boxplots of Variability, Area, and Shift are shown with vision occluded (left) and vision restored (right). Error bars indicate standard deviations. ** $P < 0.01$, * $P < 0.05$.

<https://doi.org/10.1371/journal.pone.0236437.g003>

On performing the Kruskal-Wallis rank test, the values of *Variability* ($P = 0.049$) and *Area* ($P < 0.001$) were found to be significantly different across the TLT rating groups, but the values of *Shift* were not ($P = 0.248$). Multiple comparisons using Dunn's post-hoc test revealed that the values for *Variability* were significantly different between the TLT scores of 0 and 2 ($P = 0.039$), 0 and 3 ($P < 0.001$), and 1 and 3 ($P = 0.017$). As for the values for *Area*, significant differences were seen between the TLT scores of 0 and 1 ($P = 0.027$), 0 and 2 ($P < 0.001$), 0 and 3 ($P < 0.001$), and 1 and 3 ($P = 0.015$) (Fig 5).

Discussion

This study examined, for the first time, the reliability of the TLT and its validation against quantitative measure of proprioception derived from the robotic device.

The TLT is the proprioceptive assessment frequently used in both clinical and research settings. Nevertheless, its psychometric properties have not been adequately shown, as was

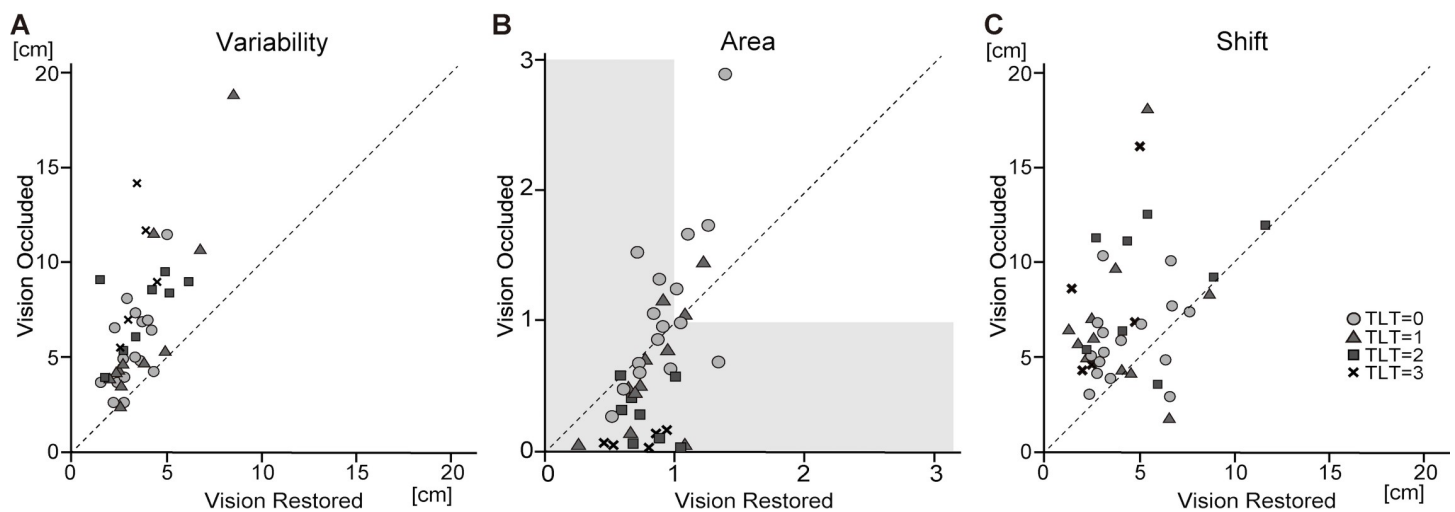


Fig 4. Scatterplots of robotic measures with vision occluded and with vision restored in each grade of the thumb localizing test (TLT). Scatterplots of (A) *Variability*, (B) *Area*, and (C) *Shift* are shown. The dots under the dotted line ($y = x$) indicate that the value is smaller with vision occluded than with vision restored, and the dots above the dotted line ($y = x$) indicate that the value is larger with vision occluded than with vision restored. The value of *Area* < 1 (gray shaded area) implies contraction and > 1 implies expansion of the workspace.

<https://doi.org/10.1371/journal.pone.0236437.g004>

Table 3. Values of the robotic measures with vision occluded and correlations with thumb localizing test.

Measures	Median value (range)	Spearman's rho (95% confidence intervals)	P values
Variability	5.76 (2.34–18.78)	0.40 (0.10–0.63)	0.011
Area	0.56 (0.01–2.90)	-0.71 (-0.84 - -0.51)	< 0.001
Shift	6.35 (1.81–18.11)	0.27 (-0.05–0.54)	0.093

<https://doi.org/10.1371/journal.pone.0236437.t003>

discussed in the Background. Contrary to the previous findings indicating the poor inter-rater reliability of some clinical sensory assessments [30], this study revealed that the TLT had almost perfect inter-rater reliability in the standardized way.

All participants in the present study could successfully reach the opposite thumb in the TLT under vision regardless of their TLT scores in the blindfold condition. Similarly, all three robotic measures were significantly worse in the vision occluded condition, compared to the vision restored condition. The improvement in motor performance when aided by visual information in patients with severely impaired proprioception is known [4, 25]. These findings indicate that clinical measures and robotic assessments have the same characteristics.

Although some studies report that vision does not always compensate for kinesthetic impairments in stroke [31, 32], the subjects in these studies were in the relatively early phases (average of less than 2 weeks post stroke). On the contrary, the present study was performed in the chronic phase, in which visuospatial deficits or fatigue were more likely to diminish compared to early or subacute phases post stroke, and thus visual compensation was considered to be more effective than the above-mentioned studies.

The TLT scores were significantly correlated with two of three quantitative measures of proprioception derived from the robotic device. Especially, a high correlation was found with *Area* in the vision occluded condition; the values of *Area* decreased as the TLT scores increased. Furthermore, the majority of values for *Area* in the vision occluded condition were < 1 and smaller than those with vision restored, especially in those with TLT ≥ 2 . The findings could be interpreted as showing that, without vision, the participants with worse TLT score might perceive the passive limb movement by the robot to be smaller than the actual movement. This result is quite reasonable considering the procedure of the TLT, in which one judges the thumb position after being moved from the preceding position similar to the

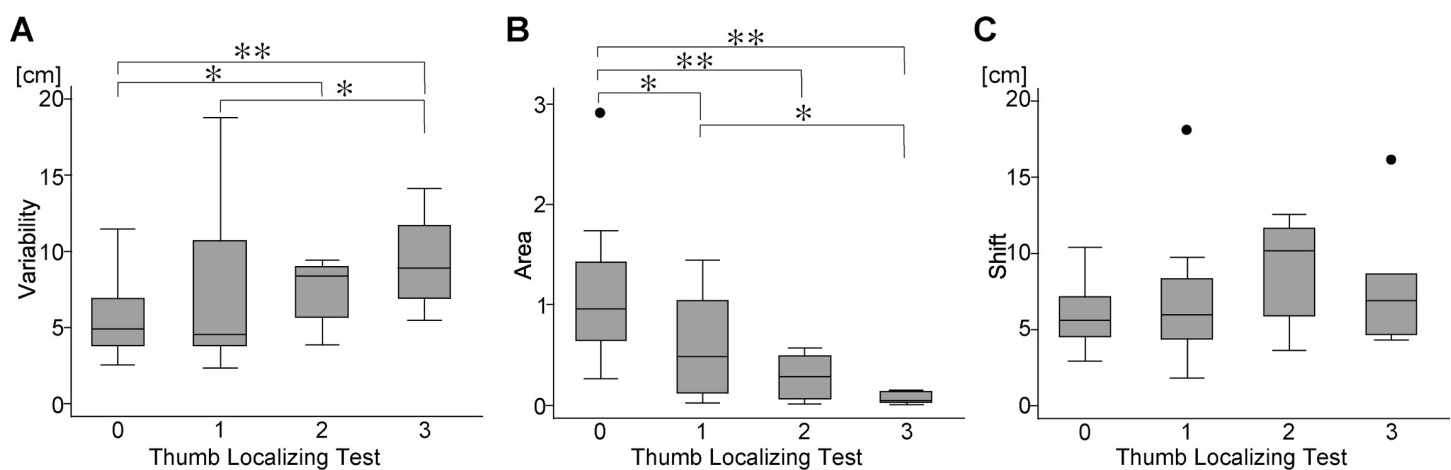


Fig 5. Robotic measures with vision occluded of the arm position matching task in each grade of the thumb localizing test (TLT). Boxplots of (A) Variability, (B) Area, and (C) Shift are shown. Error bars indicate standard deviations and the lines within the box indicate median values. ** $P < 0.01$, * $P < 0.05$ with the Dunn's test for multiple comparisons.

<https://doi.org/10.1371/journal.pone.0236437.g005>

robotic task, and it also indicates that the TLT may reflect at least a part of “kinesthetic sensation”, which is generated from muscle spindles in response to limb movement and changes in position [33]. The TLT scores also correlated moderately with *Variability*. The result seems to be reasonable because *Variability* is similar to the TLT rating in that it evaluates the gap or sway of the active limb. This measure was also previously reported as having a significant association with the TLT in its ability to discriminate between the normal and abnormal [22, 34], which was confirmed by our study. On the other hand, this study did not show significant correlation between the TLT and *Shift*. The measure reflects the bias of the expression of the perceived area, and could be influenced by rather other factors such as spatial cognition than proprioception. Therefore, no correlation with *Shift* may not hamper the validity of the TLT for proprioceptive sensation.

Although no study has examined the correlation between TLT and quantitative measures, i.e. Area, Variability, and Shift, some conflicts were found in a few previous studies [20, 22], where no categorical relationships were demonstrated between the TLT and Area in stroke patients, and significant categorical relationships were found between the TLT and Shift. This difference in the findings may be related to two reasons. First, in our study, an almost perfect inter-rater reliability in the TLT was achieved by standardizing the details of testing and the rating. This is the strong point of the present study, ensuring a robustness of the following analysis. Second, our study investigated chronic patients living independently, which suggests they have mild, if any, cognitive problems. In contrast, the previous studies [20, 22] investigated subacute patients whose functional abilities are various, and thus might have various confounding factors such as disturbance in spatial cognition due to cognitive impairments, including visuospatial neglect [31] or deficits in generalized attention. Our study might successfully focus on sensorimotor impairments with less confounding factors.

In the comparisons between the grades of TLT, *Variability* and *Area* were not statistically different and substantially overlapped between grade 1 and 2, and between 2 and 3 of the TLT score. This result suggests that the TLT rating may have some ambiguity in assessing the degree of proprioceptive deficits between these grades. Clinicians should bear in mind that the difference of 1 point in the TLT score between grade 1–3 should not be emphasized. In the context of robotic devices being more objective and quantitative than clinical assessment [35–37], the robotic assessment can be considered for further assessment of quantification of proprioception such as research purpose.

This study had a few limitations. First, another study design is needed to study test-retest reliability of the TLT. Second, the generalizability of our findings may be somewhat limited as the present study investigated chronic patients living independently, who were relatively young and high functioning compared to the general stroke population. In addition, further investigation will be needed to generalize our findings to individuals with other pathological conditions. Finally, another study would be required to reveal the relationship between the proprioceptive deficits represented by the TLT and brain lesions that may affect passive kinesthetic sensation or sensorimotor integration, as indicated in recent studies [8, 31].

Conclusions

The present study demonstrated that the TLT had a high inter-rater reliability when the testing procedure was standardized. The TLT correlated highly with the quantitative measures which reflects the perceived area of movement and variability of the limb location. These findings contribute to the robustness of the TLT as a clinical measure. Furthermore, the overlapping of values in quantitative measures between consecutive grades 1–3 of the TLT gave valuable information for interpretation of the TLT score in the clinical settings. In spite of the ordinal

nature of the scale, which tends to be a disadvantage when compared with continuous scales, the TLT has proven to be essential as an easily performed and robust tool to detect proprioceptive deficits in clinical scenes.

Supporting information

S1 Data. Data for assessing the inter-rater reliability of the TLT. Details of the variables are as follows: SubjectNo, an anonymous subject number for the present study only; PareticSide_R1L2, paretic side (1 = right, 2 = left); TL_sv_A1/2/3, the scores of TLT with vision occluded condition by one of the two raters; TL_sv_B1/2/3, the scores of TLT with vision occluded condition by another rater; each number describes three limb positions, as defined in the manuscript (1 = Fig 1C(a), 2 = Fig 1C(b), 3 = Fig 1C(c)). (XLSX)

S2 Data. Data of participant characteristics, clinical measures, and robotic measures. Details of the variables are as follows: TYPE_C1_H2_S3, the type of diseases (1 = cerebral infarction, 2 = cerebral hemorrhage, and 3 = subarachnoid hemorrhage); PareticSide_R1L2, paretic side (1 = right, 2 = left); MAS_elbow/wrist/finger, Modified Ashworth scale, originally scaled as 0, 1, 1+, 2, 3, 4, described in the dataset as 0, 1, 2, 3, 4, 5 respectively; Variability_sv, Area_sv, Shift_sv, robotic measures obtained with vision occluded condition; Variability_cv, Area_cv, Shift_cv, robotic measures with vision restored condition; TL_sv_1, TL_sv_2, TL_sv_3, the ultimate TLT scores discussed and rescored by the two raters; TL_sv_median, the median value of the TLT scores in the three limb positions. (XLSX)

Acknowledgments

We thank Ms. Sawako Ohtaki and Ms. Miho Hiramoto for their technical support.

Author Contributions

Conceptualization: Eri Otaka, Yohei Otaka.

Data curation: Eri Otaka, Yohei Otaka, Shoko Kasuga, Atsuko Nishimoto, Kotaro Yamazaki.

Formal analysis: Eri Otaka.

Methodology: Eri Otaka, Yohei Otaka.

Supervision: Michiyuki Kawakami, Junichi Ushiba, Meigen Liu.

Writing – original draft: Eri Otaka.

Writing – review & editing: Eri Otaka, Yohei Otaka, Shoko Kasuga, Atsuko Nishimoto, Kotaro Yamazaki, Michiyuki Kawakami, Junichi Ushiba, Meigen Liu.

References

1. Rothwell JC, Traub MM, Day BL, Obeso JA, Thomas PK, Marsden CD. Manual motor performance in a deafferented man. *Brain: a journal of neurology*. 1982; 105 (Pt 3):515–42. Epub 1982/09/01. <https://doi.org/10.1093/brain/105.3.515> PMID: 6286035.
2. Cordo P, Carlton L, Bevan L, Carlton M, Kerr GK. Proprioceptive coordination of movement sequences: role of velocity and position information. *Journal of neurophysiology*. 1994; 71(5):1848–61. Epub 1994/05/01. <https://doi.org/10.1152/jn.1994.71.5.1848> PMID: 8064352.
3. Gordon J, Ghilardi MF, Ghez C. Impairments of reaching movements in patients without proprioception. I. Spatial errors. *Journal of neurophysiology*. 1995; 73(1):347–60. Epub 1995/01/01. <https://doi.org/10.1152/jn.1995.73.1.347> PMID: 7714577.

4. Ghez C, Gordon J, Ghilardi MF. Impairments of reaching movements in patients without proprioception. II. Effects of visual information on accuracy. *Journal of neurophysiology*. 1995; 73(1):361–72. Epub 1995/01/01. <https://doi.org/10.1152/jn.1995.73.1.361> PMID: 7714578.
5. Kita K, Otaka Y, Takeda K, Sakata S, Ushiba J, Kondo K, et al. A pilot study of sensory feedback by transcutaneous electrical nerve stimulation to improve manipulation deficit caused by severe sensory loss after stroke. *Journal of neuroengineering and rehabilitation*. 2013; 10:55. Epub 2013/06/15. <https://doi.org/10.1186/1743-0003-10-55> PMID: 23764012; PubMed Central PMCID: PMC3701472.
6. Carey LM. Somatosensory Loss after Stroke. *Crit Rev Phys Rehabil Med*. 1995; 7(1):51–91. <https://doi.org/10.1615/CritRevPhysRehabilMed.v7.i1.40>
7. Carey LM, Matyas TA. Frequency of discriminative sensory loss in the hand after stroke in a rehabilitation setting. *Journal of rehabilitation medicine*. 2011; 43(3):257–63. Epub 2011/02/10. <https://doi.org/10.2340/16501977-0662> PMID: 21305243.
8. Kessner SS, Schlemm E, Cheng B, Bingel U, Fiehler J, Gerloff C, et al. Somatosensory Deficits After Ischemic Stroke. *Stroke*. 2019; 50(5):1116–23. Epub 2019/04/05. <https://doi.org/10.1161/STROKEAHA.118.023750> PMID: 30943883.
9. Kwakkel G, Kollen BJ, van der Grond J, Prevo AJ. Probability of regaining dexterity in the flaccid upper limb: impact of severity of paresis and time since onset in acute stroke. *Stroke*. 2003; 34(9):2181–6. Epub 2003/08/09. <https://doi.org/10.1161/01.STR.0000087172.16305.CD> PMID: 12907818.
10. Winters C, van Wegen EE, Daffertshofer A, Kwakkel G. Generalizability of the Proportional Recovery Model for the Upper Extremity After an Ischemic Stroke. *Neurorehabilitation and neural repair*. 2015; 29(7):614–22. Epub 2014/12/17. <https://doi.org/10.1177/1545968314562115> PMID: 25505223.
11. Semrau JA, Herter TM, Scott SH, Dukelow SP. Examining Differences in Patterns of Sensory and Motor Recovery After Stroke With Robotics. *Stroke*. 2015; 46(12):3459–69. Epub 2015/11/07. <https://doi.org/10.1161/STROKEAHA.115.010750> PMID: 26542695.
12. Zandvliet SB, Kwakkel G, Nijland RHM, van Wegen EEH, Meskers CGM. Is Recovery of Somatosensory Impairment Conditional for Upper-Limb Motor Recovery Early After Stroke? *Neurorehabilitation and neural repair*. 2020; 34(5):403–16. Epub 2020/05/12. <https://doi.org/10.1177/1545968320907075> PMID: 32391744; PubMed Central PMCID: PMC7222963.
13. Hirayama K. [Thumb/big-toe localizing test: examination for deficit of proprioception through the posterior column-medial lemniscal system]. *Brain and nerve = Shinkei kenkyu no shinpo*. 2011; 63(8):851–60. Epub 2011/08/06. PMID: 21817176.
14. Hirayama K, Fukutake T, Kawamura M. 'Thumb localizing test' for detecting a lesion in the posterior column-medial lemniscal system. *Journal of the neurological sciences*. 1999; 167(1):45–9. Epub 1999/09/29. [https://doi.org/10.1016/s0022-510x\(99\)00136-7](https://doi.org/10.1016/s0022-510x(99)00136-7) PMID: 10500261.
15. Hirayama K, Fukutake T, Kawamura M. [Thumb localizing test—examination for a disturbance of articular localization]. *Rinsho shinkeigaku = Clinical neurology*. 1986; 26(5):448–54. Epub 1986/05/01. PMID: 3742897.
16. Prescott RJ, Garraway WM, Akhtar AJ. Predicting functional outcome following acute stroke using a standard clinical examination. *Stroke*. 1982; 13(5):641–7. Epub 1982/09/01. <https://doi.org/10.1161/01.str.13.5.641> PMID: 7123597.
17. Carey LM, Oke LE, Matyas TA. Impaired limb position sense after stroke: a quantitative test for clinical use. *Archives of physical medicine and rehabilitation*. 1996; 77(12):1271–8. Epub 1996/12/01. [https://doi.org/10.1016/s0003-9993\(96\)90192-6](https://doi.org/10.1016/s0003-9993(96)90192-6) PMID: 8976311.
18. Simo L, Botzer L, Ghez C, Scheidt RA. A robotic test of proprioception within the hemiparetic arm post-stroke. *Journal of neuroengineering and rehabilitation*. 2014; 11:77. Epub 2014/06/03. <https://doi.org/10.1186/1743-0003-11-77> PMID: 24885197; PubMed Central PMCID: PMC4011167.
19. Mrotek LA, Bengtson M, Stoeckmann T, Botzer L, Ghez CP, McGuire J, et al. The Arm Movement Detection (AMD) test: a fast robotic test of proprioceptive acuity in the arm. *Journal of neuroengineering and rehabilitation*. 2017; 14(1):64. Epub 2017/07/01. <https://doi.org/10.1186/s12984-017-0269-3> PMID: 28659156; PubMed Central PMCID: PMC5490232.
20. Dukelow SP, Herter TM, Moore KD, Demers MJ, Glasgow JI, Bagg SD, et al. Quantitative assessment of limb position sense following stroke. *Neurorehabilitation and neural repair*. 2010; 24(2):178–87. Epub 2009/10/02. <https://doi.org/10.1177/1545968309345267> PMID: 19794134.
21. Scott SH. Apparatus for measuring and perturbing shoulder and elbow joint positions and torques during reaching. *Journal of neuroscience methods*. 1999; 89(2):119–27. Epub 1999/09/24. [https://doi.org/10.1016/s0165-0270\(99\)00053-9](https://doi.org/10.1016/s0165-0270(99)00053-9) PMID: 10491942.
22. Dukelow SP, Herter TM, Bagg SD, Scott SH. The independence of deficits in position sense and visually guided reaching following stroke. *Journal of neuroengineering and rehabilitation*. 2012; 9:72. Epub 2012/10/06. <https://doi.org/10.1186/1743-0003-9-72> PMID: 23035968; PubMed Central PMCID: PMC3543214.

23. Fugl-Meyer AR, Jaasko L, Leyman I, Olsson S, Steglind S. The post-stroke hemiplegic patient. 1. a method for evaluation of physical performance. *Scandinavian journal of rehabilitation medicine*. 1975; 7(1):13–31. Epub 1975/01/01. PMID: [1135616](#).
24. Bohannon RW, Smith MB. Interrater reliability of a modified Ashworth scale of muscle spasticity. *Physical therapy*. 1987; 67(2):206–7. Epub 1987/02/01. <https://doi.org/10.1093/ptj/67.2.206> PMID: [3809245](#).
25. Newport R, Hindle JV, Jackson SR. Links between vision and somatosensation. Vision can improve the felt position of the unseen hand. *Curr Biol*. 2001; 11(12):975–80. Epub 2001/07/13. [https://doi.org/10.1016/s0960-9822\(01\)00266-4](https://doi.org/10.1016/s0960-9822(01)00266-4) PMID: [11448775](#).
26. Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological bulletin*. 1968; 70(4):213–20. Epub 1968/10/01. <https://doi.org/10.1037/h0026256> PMID: [19673146](#).
27. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977; 33(1):159–74. Epub 1977/03/01. PMID: [843571](#).
28. Guilford JP, Lyons TC. On determining the reliability and significance of a tetrachoric coefficient of correlation. *Psychometrika*. 1942; 7(4):243–9. <https://doi.org/10.1007/BF02288627>.
29. Faul F, Erdfelder E, Buchner A, Lang AG. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behavior research methods*. 2009; 41(4):1149–60. Epub 2009/11/10. <https://doi.org/10.3758/BRM.41.4.1149> PMID: [19897823](#).
30. Lincoln N, Crow J, Jackson J, Waters G, Adams S, Hodgson P. The unreliability of sensory assessments. *Clinical Rehabilitation*. 1991; 5:273–82.
31. Semrau JA, Herter TM, Scott SH, Dukelow SP. Vision of the upper limb fails to compensate for kinesthetic impairments in subacute stroke. *Cortex; a journal devoted to the study of the nervous system and behavior*. 2018; 109:245–59. Epub 2018/11/06. <https://doi.org/10.1016/j.cortex.2018.09.022> PMID: [30391879](#).
32. Herter TM, Scott SH, Dukelow SP. Vision does not always help stroke survivors compensate for impaired limb position sense. *J Neuroeng Rehabil*. 2019; 16(1):129. Epub 2019/11/02. <https://doi.org/10.1186/s12984-019-0596-7> PMID: [31666135](#); PubMed Central PMCID: [PMC6822422](#).
33. Proske U, Gandevia SC. The proprioceptive senses: their roles in signaling body shape, body position and movement, and muscle force. *Physiological reviews*. 2012; 92(4):1651–97. Epub 2012/10/18. <https://doi.org/10.1152/physrev.00048.2011> PMID: [23073629](#).
34. Kuczynski AM, Dukelow SP, Semrau JA, Kirton A. Robotic Quantification of Position Sense in Children With Perinatal Stroke. *Neurorehabilitation and neural repair*. 2016; 30(8):762–72. Epub 2016/01/10. <https://doi.org/10.1177/1545968315624781> PMID: [26747126](#).
35. Balasubramanian S, Colombo R, Sterpi I, Sanguineti V, Burdet E. Robotic assessment of upper limb motor function after stroke. *American journal of physical medicine & rehabilitation*. 2012; 91(11 Suppl 3):S255–69. Epub 2012/10/25. <https://doi.org/10.1097/PHM.0b013e31826bc1> PMID: [23080041](#).
36. Scott SH, Dukelow SP. Potential of robots as next-generation technology for clinical assessment of neurological disorders and upper-limb therapy. *Journal of rehabilitation research and development*. 2011; 48(4):335–53. Epub 2011/06/16. <https://doi.org/10.1682/jrrd.2010.04.0057> PMID: [21674387](#).
37. Otaka E, Otaka Y, Kasuga S, Nishimoto A, Yamazaki K, Kawakami M, et al. Clinical usefulness and validity of robotic measures of reaching movement in hemiparetic stroke patients. *Journal of neuroengineering and rehabilitation*. 2015; 12:66. Epub 2015/08/13. <https://doi.org/10.1186/s12984-015-0059-8> PMID: [26265327](#); PubMed Central PMCID: [PMC4533801](#).