The Foot 48 (2021) 101851

Contents lists available at ScienceDirect

The Foot

journal homepage: www.elsevier.com/locate/foot

Original Article

An evaluation of the reliability of the foot-tapping test in a healthy sample

Brian A. Pribble *, Christopher D. Black, Daniel J. Larson, Rebecca D. Larson

The University of Oklahoma, Department of Health and Exercise Science, 1401 Asp Ave, Room 104, Norman, OK 73019, United States

ARTICLE INFO

Keywords:
Foot-Tapping test
FTT
Repetitive tapping
Test reliability
Upper motor neuron assessment

ABSTRACT

The foot-tapping test (FTT) can be used to assess upper motor neuron dysfunction in clinical populations. However, relatively little is known regarding the reliability or normative values of the FTT in either healthy or clinical populations. Although several different FTT methods have been used, no study to date has demonstrated the reliability or validity of FTT by comparing it across several different counting methods in healthy persons. This unfortunately limits its usefulness in medicine and research.

Objective: This study sought to examine the reliability and validity of the FTT in healthy individuals to determine its usefulness and to make recommendations for its implementation in clinical populations. Furthermore, the concurrent validity and reliability of using a force plate as an objective measure of foot-taps was considered.

Design: Thirty-eight healthy individuals had their foot-tapping assessed using Live, Force Plate, and Video Counting methods over four separate visits.

Methods: Participants were seated as per previous FTT recommendations and asked to tap their foot in 10-second intervals while the number of taps was counted via Live, Video, and Force Plate counters. This was done with both legs, with shoes ON and OFF, and repeated over four separate visits.

Results: Despite significant differences between repeat trials for Force Plate and Video Counts (~2 foot-taps, $p < 0.01$), test-retest reliability was high for all three methods (Pearson's $R > 0.90$). Dominant foot trials were higher (~2 foot-taps, $p < 0.05$) than Non-dominant for all three counts. When performed with shoes ON, counts were higher (~2 foot-taps, $p < 0.05$) than OFF for the Live and Force Plate counts. Reliability between visits was high ($ICC > 0.80$) and only the Video count was significantly lower for Visit 1 ($p < 0.01$).

Conclusions: Given findings, the authors suggest using a Force Plate counting method and have compiled a list of suggestions for future implementation of the FTT.

1. Introduction

Proper dorsi- and plantarflexion of the foot is crucial in performing activities of daily living (ADL). Impairments in either dorsi or plantarflexion during ADL can result in compensatory movements [2,3,5,13,23,24] thereby negatively affecting gait [22], movement efficiency [5,22], fatigue, and increase the likelihood of falls or injury [8,14,18]. Though inevitable with aging, such changes tend to be more pronounced in clinical populations with upper motor neuron (UMN) deficits such as those with cervical myelopathy, multiple sclerosis, amyotrophic lateral sclerosis, and Parkinson's disease. Impairments in UMN are common in these individuals and may result in a reduced ability to produce rapid and repetitive movements, particularly those of the hands and feet [1,4,21,6,7,10,11,16,17,19,20]. To that end, a simple evaluation known foot-tapping test (FTT) has been utilized to assess UMN function in these clinical populations [4,6,11,16,17,19–21]. Previous studies utilizing the FTT found a tendency for foot-tapping rates to be lower in older adults and those with known UMN weakness compared to controls [1,4,7,10,11,17,19,20].

Despite clinically significant findings with the FTT, very little normative data or reliability statistics exist for either healthy or clinical populations. The few studies that have evaluated the use of the FTT found it to have a high test-retest [7,9,12,19], inter-rater reliability and agreement with known UMN weakness [1,15]. It has been previously demonstrated that the FTT can be used to track changes in UMN function over longer periods of time (months to years) [11,21] or shorter term following clinical interventions [6,19]. However, no study to date has successfully demonstrated the short-term stability (days to weeks) of the FTT in healthy individuals. It is remains unclear whether or by how much the average healthy person's foot-tapping ability fluctuates over repeated trials and visits or what is even considered a clinically significant change in foot-tapping ability.

* Corresponding author.
E-mail addresses: Brian.a.pribble-1@ou.edu (B.A. Pribble), cblack@ou.edu (C.D. Black), larsondj@ou.edu (D.J. Larson), rdlarson@ou.edu (R.D. Larson).

https://doi.org/10.1016/j.foot.2021.101851
Received 28 September 2020; Received in revised form 3 May 2021; Accepted 8 July 2021
Available online 12 July 2021
0958-2592/

Previous FTT studies have utilized either visual counting [1,10,11,16,17,19] or automated counting devices [4,6,7,21,25] to count the number of foot-taps, however none have sought to compare foot-tapping rates between different counting methods. Visual counting tends to be the most utilized method as it does not require special equipment and can easily performed. This comes at the cost however of potentially introducing human error. Previous studies have utilized automated foot-tap counting [4,6,7,21,25] such as foot pedals, foot mounted camera systems, or force plate like measurement devices. It becomes difficult to say however whether these automated methods of foot-tap counting are either valid or reliable without first comparing them to some other "gold standard". However, a "gold standard" of foot-tap counting has yet to be identified, nor has what constitutes a "foot-tap" been defined. We must ask, does a foot-tap need to meet a certain threshold to be considered a foot-tap? Must the foot strike the ground in a particular manner, make complete contact with the ground, use a certain range of motion, or exert a certain amount of force? These questions have yet to be answered and would likely implications for how foot-tapping is conducted and interpreted.

The FTT offers a simple means of assessing UMN function yet remains underutilized outside of clinical research. This is perhaps due in part to a lack of standardized FTT protocol which has led to a lack of normative data in both healthy and clinical populations. Although significant findings have been reported in clinical populations, authors often neglect to describe their FTT procedures or what constitutes a true "foot-tap" leaving a lot of room for ambiguity. In order for the FTT to be recognized as a clinically relevant measure of UMN dysfunction, it must be first shown to be a reliable and objective measure of foot-tapping ability in healthy individuals. It also remains unclear whether variables such as testing conditions, counting method, or subject characteristics significantly contribute to foot-tapping rates. To that end, this study sought to examine the FTT in a healthy population under various testing and counting conditions in order to evaluate its reliability and make recommendations for the standardization of the FTT procedures using a force plate measurement.

2. Methods

Thirty-eight healthy individuals free of any known disorders or injuries affecting the legs participated in this study. The University of Oklahoma Institutional Review Board approved this study and all subjects provided written consent prior to participation.

Subject characteristics were assessed using a series of health and physical activity questionnaires and a full body DXA scan. Subjects performed a total of thirty-two FTT trials over four separate visits. Each visit consisted of eight FTT trials in which each leg (dominant and non-dominant limb in randomized order) was assessed twice with the shoes ON and again with the shoes OFF. [4 visits * (4 trials w/shoes ON + 4 w/shoes OFF) = 32 trials]. Time of day and shoe type were kept consistent in order to eliminate them as confounding variables.

Trials were performed with subjects seated in a chair, with hips and knees at approximately 90 degrees, and feet positioned over two 10″x10″ force plates (NEULOG) (Fig. 1a & b). The feet were placed so that when the subject plantar flexed and their foot struck the ground, only the ball of the foot made contact with the plate, allowing the number of foot-taps to be counted. Foot position was kept consistent between trials using a grid system on the plates. A testing interval of 10-s was used, as it is most commonly reported [1,7,10,11,17,20,21]. Subjects were instructed to tap the ball of their foot on the force plate as quickly as possible while keeping heel firmly planted. Subjects were instructed to use a comfortable and consistent range of motion while tapping, so long as it allowed for complete clearance of the foot between taps. Subjects were given a countdown (3, 2, 1 & GO) to start tapping and 1–2 min rest between trials. The first four trials were performed

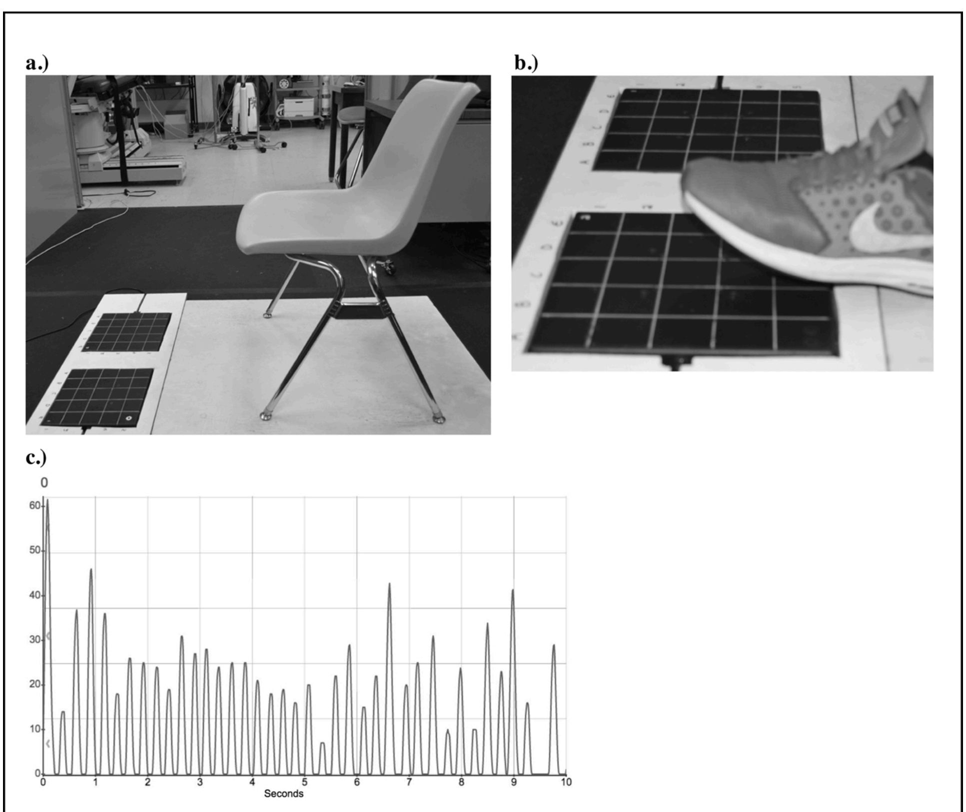

Fig. 1. Foot-tapping test experiment setup. a: FTT setup with chair and two force plates, b: example of foot proper foot placement with heel off of plate, and c: example of FTT trial count with force plate (n = 38).

with the shoes ON (two trials for each foot) and then another four trials with the shoes OFF. This was done so as to determine if shoes significantly altered the mechanics or count of foot-tapping.

Each trial was counted using three different methods. *Live Count.* This count was done using live visual inspection by the main investigator during testing. Using their best judgment of what constituted a tap, the investigator visually counted and recorded the number of times the subject tapped their foot in 10-s. Any foot-taps that did not appear to fully strike and/or clear the force plate were disregarded. *Force Plate.* Force tracings from each FTT trial were used to count the number of foot-taps performed. Each time the ball of the foot would strike the plate, a corresponding peak in force could be seen on the tracing. The number of peaks in the tracings were then counted by the researcher and recorded as the number of foot-taps (Fig. 1c). Again, best judgement was used and any taps that appeared incomplete were disregarded. *Video Count.* Each trial was videotaped at ground level, held stable by a miniature tripod and controlled via Bluetooth remote. All videotaped trials of the FTT were slowed to ½ speed, visually counted, and recorded. The slowed videotaped trials had the added benefit of giving off a distinct 'ringing' sound each time the foot struck the metal surface of the plate. Researchers were furthermore blinded from other counts so as to minimize bias during the various counts. A subset of the subjects (n = 23) had their Video and Force Plate Counts analyzed by a secondary researcher who was also blinded from previous counts.

Data are presented as means ± (SD). Two-way repeated measures ANOVAs were run to compare the number of foot-taps counted using the three counting methods (Live, Video, and Force Plate) under various conditions (i.e. Shoes ON vs. OFF, Dominant vs. Non-Dominant Foot, and over different trials and visits). The simple main effects were determined using either one-way repeated measure ANOVA or Paired samples T-test. Cohen's D and Partial Eta Squared were used to calculate effect sizes. Reliability of the FTT was examined using Pearson's R correlations, Coefficient of variation (%CV), Intraclass correlation Coefficient (ICC), and Bland Altmans plots.

3. Results

The subjects mean age, activity level, shoe size, weight, height, bone mineral density, and lower limb lean tissue mass were assessed (Supplementary Table 1). Foot-tapping ability did not significantly correlate with any of the subject characteristics for any of the counting methods (Supplementary Table 2).

No significant interaction for Counting_Method*Trial_Number ($p = 0.26$, $\eta^2 = 0.002$) was observed. Significant main effects were found for Counting_Method ($p < 0.01$, $\eta^2 = 0.418$) and Trial_Number ($p < 0.01$, $\eta^2 = 0.025$). Each counting method was found to be different from each other ($p < 0.01$) and Trial 1 tended to be higher than Trial 2 for the Force Plate and Video counts only ($p < 0.01$) (Table 1a). In subsequent analyses, trials 1 and 2 have been averaged together and are presented as their mean.

No significant interaction for Counting_Method*Foot_Dominance ($p = 0.19$, $\eta^2 = 0.006$) was observed. Significant main effects were observed for Counting_Method ($p < 0.01$, $\eta^2 = 0.460$) and Foot_Dominance ($p < 0.01$, $\eta^2 = 0.197$). The number of foot-taps counted for each counting method significantly differed from each other ($p < 0.01$) and the Dominant Foot trials tended to be higher than Non-Dominant ($p < 0.01$) (Table 1b). Due to significant differences found between feet, only the Dominant foot trials were considered in subsequent analyses.

A significant interaction effect for Counting_Method*Shoes ON/OFF ($p < 0.01$, $\eta^2 = 0.126$) and main effects for the Counting Method ($p < 0.01$, $\eta^2 = 0.464$) and the Shoes ON/OFF Trials ($p < 0.01$, $\eta^2 = 0.086$) were observed. The number of foot-taps counted for each method was found to be significantly different from each other ($p < 0.01$) and trials with Shoes ON tended to be significantly higher than shoes OFF for the Live and Force Plate counts ($p < 0.01$) (Table 1c). In subsequent analyses only the trials with Shoes ON were considered.

A significant interaction effect for Counting_Method * Visit_Number ($p < 0.01$, $\eta^2 = 0.153$) and main effect for Counting Method ($p < 0.01$, $\eta^2 = 0.633$) but not for Visit Number ($p = 0.20$, $\eta^2 = 0.041$) was observed. The number of foot-taps counted for the Live and Video counts did not significantly differ from each other ($p = 0.73$) but did significantly differ from the Force Plate count ($p < 0.01$). Foot-taps for Visits 1, 2, 3, and 4 did not significantly differ from each other for either the Live or Force Plate counts ($p > 0.05$) but was significantly less for the Video count on visit one ($p < 0.05$). Between visits ICC was high for all count methods (ICC > 0.80) (Table 2). In the subsequent analysis the FTT trials of all 4 visits were averaged together.

After collapsing across trials and variables, a significant difference in foot-taps was observed between the counting methods ($p < 0.01$, $\eta^2 = 0.633$). Pairwise comparisons revealed that Live and Video counts did not significantly differ from each other ($p = 0.73$); the Force Plate count however was significantly lower than the other two methods ($p < 0.01$) (Table 3). Bland-Altman plots show the agreement between each of the three counting methods (Supplementary Fig. 1).

It was found that foot-taps counted significantly differed between counters for both Video ($p = 0.01$) and Force Plate counts ($p < 0.01$) and had high inter-rater reliability ($R = 0.83$ and 0.99 respectively (Table 4).

4. Discussion

The FTT has previously been used in clinical populations to assess UMN function. Despite this, relatively little is known regarding the FTT's reliability or normative foot-tapping rates in either healthy or clinical populations. It is furthermore unclear how differences in FTT methodologies or subject characteristics may affect foot-tapping rates. Our study was not the first to attempt assessing the FTT with automated counters [4,6,7,21,25], yet no other study to date has assessed the reliability or concurrent validity of the FTT with a force plate by comparing it to other visual counting methods.

The current study evaluated the difference in foot-taps counted between two subsequent trials (Trial 1 and Trial 2) using three counting methods (Live, Force Plate, and Video Counting). There was a tendency for Trial 1 to be higher than Trial 2 for the Force Plate and Video ($p < 0.05$) but not Live count ($p = 0.09$). Despite statistical significance, the difference between trials 1 and 2 was <1 foot-tap for each counting method and showed a high level of test-retest reliability (Pearson's R > 0.90). This is in line with similar previous studies which found a Pearson's R correlation of 0.81–0.93 and ICC = 0.793 between trials in healthy subjects [7,9,19]. The non-significance between Trials 1 and 2 and lower %CV for the Live Count method might seem to imply it has a higher level of test-retest reliability. However, the opposite could also be argued for. Perhaps the live count lacks the necessary precision to detect a difference between repeated trials and is therefore not a valid or reliable method of foot-tap counting.

In the current study it was found that the dominant limb had a significantly higher number of foot-taps than the non-dominant limb

Table 1a
Mean difference in foot-tapping counts between Trial 1 and Trial 2.

Counting Method	Trial 1	Trial 2	Mean Dif.	Trials CV	Sig.	Pearson's R	Cohen's D
Live	48.3 (10.3)	48.0 (10.0)	0.3 (4.5)	150%	0.09	0.90	0.03
Force Plate	43.3 (8.8)	42.9 (8.5)	0.5 (3.6)	720%	<0.01*	0.92	0.05
Video	50.8 (10.7)	50.2 (10.3)	0.6 (4.2)	700%	<0.01*	0.93	0.06

Table 1b
Mean difference in foot-tapping counts between the dominant and non-dominant foot.

Counting Method	Dom. Foot	Non-Dom. Foot	Mean Difference	Trials CV	Sig.	Pearson's R	Cohen's D Effect size
Live	49.2 (9.7)	47.1 (10.0)	2.1 (4.8)	228.6%	<0.01*	0.88	0.21
Force Plate	44.2 (8.3)	42.0 (8.5)	2.1 (4.7)	223.8%	<0.01*	0.85	0.25
Video	51.3 (10.0)	49.7 (10.5)	1.7 (5.2)	305.9%	<0.01*	0.87	0.16

Table 1c
Mean difference in foot-tapping counts between shoes ON and shoes OFF trials.

Counting Method	Shoes ON	Shoes OFF	Mean Difference	Trials CV	Sig.	Pearson's R	Cohen's D Effect size
Live	50.5 (9.8)	47.9 (9.4)	2.6 (4.8)	184.6%	<0.01*	0.88	0.27
Force plate	44.6 (8.8)	43.7 (7.8)	0.90 (4.3)	477.8%	0.01*	0.87	0.11
Video	51.3 (10.1)	51.4 (9.8)	−0.08 (4.8)	−6000%	0.85	0.88	0.01

Values represent mean number of foot-taps ± (SD). Sig. mean differences ($p < 0.05$) denoted with *. CV = % coefficient of variation.

Table 2
Day-to-day differences in foot-tapping counts.

Counting method	Visit 1 Foot-Taps	Visit 2 Foot-Taps	Visit 3 Foot-Taps	Visit 4 Foot-Taps	Sig.	Between-visit ICC
Live	50.8 (10.1) 19.9	51.8 (10.5) 20.3	49.6 (9.4) 19.0	49.6 (9.5) 19.2	0.29	0.88
Force plate	43.4 (8.6) 19.8	44.7 (9.0) 20.1	45.1 (9.6) 21.3	45.2 (8.1) 17.9	0.11	0.94
Video	48.6 (9.5) 19.5	51.3 (9.9) 19.3	52.0 (10.7) 20.6	53.3 (10.2) 19.1	<0.01*	0.93

Values represent mean number of foot-taps ± (SD). Sig. mean differences ($p \leq 0.05$) denoted with *.

Table 3
Differences in foot-tapping count by counting method.

Counting Method	Mean	95% CI (Lower)	95% CI (Upper)	CV
Live	50.5 (8.5)	47.7	53.2	16.8%
Force Plate	44.6 (8.2)*	41.9	47.3	18.4%
Video	51.3 (9.3)	48.2	54.3	18.1%

Values represent mean number of foot-taps ± (SD).Sig. mean differences ($p < 0.05$) denoted with *. CV = % coefficient of variation.

Table 4
Differences in foot-taps counted (Counter 1–Counter 2).

Counting method	Counters Mean Diff.	Mean Dif. Sig.	CV	Pearson's R	Cohen's D effect size
Video	−4.3 (7.5)	0.01*	−174.4%	0.83	0.36
Force Plate	1.3 (1.2)	<0.01*	92.3%	0.99	0.14

Values represent mean number of foot-taps ± (SD).Sig. mean differences ($p < 0.05$) denoted with *. CV= % Coefficient of variation.

(~2 foot-taps) for each counting method. Furthermore, the three counting methods all had very high Pearson's R correlations between the dominant and non-dominant limbs (Pearson's R = 0.85−0.88) and similar % CV (223.8–305.9%). The between limb differences in the current study were similar to the studies by Hinman and Numasawa et al. who found a similar ~3 foot-tap difference between limbs [7,19]. In another study by Kalaycıoğlu et al. [9], foot preference highly correlated with foot-tapping ability and was more highly correlated with skilled than unskilled foot tasks. Despite significant differences in the current study, a 2 foot-tap difference is unlikely to be clinically significant. It is therefore our belief that the FTT may be done on the dominant limb for general screening purposes, but should be done bilaterally in cases of suspected limb assymetries.

To our knowledge, the current study is the first to examine the differences in foot-tapping rate with the shoes ON vs. shoes OFF. Previous studies of the FTT tended not to specify whether shoes were even worn or standardized across trials; a detail which would likely be of clinical interest. It was found that more foot-taps were counted for shoes ON than shoes OFF trials for Live and Force Plate counts ($p < 0.01$ and $p = 0.01$ respectively) but not video count ($p = 0.85$). All three methods showed high Pearson's R correlations between shoes ON and OFF trials (Pearson's R = 0.87−0.88). Though found to be significant, a difference of ~1 foot-tap would not likely be clinically significant and may actually be explained by the added benefits of wearing shoes during the test. Anecdotally speaking, the main investigator found it easier to count foot-taps with the shoes ON as it allowed them to better standardize what constitutes a tap (the midsole visually makes contact with ground) and gave off a more distinctive audible sound upon foot strike. And although the force plate is very sensitive (requiring at least 1 Newton of force) at picking up foot-taps, perhaps the added weight of the shoes allowed for better distinction of what was or wasn't a foot-tap from the force tracings. In populations with functional limitations, it may therefore actually be more beneficial to conduct the FTT with shoes ON for the safety and ease of the patient. What's more, it reasons that a patient's performance of the FTT with the shoes ON may more closely approximate performance of ADL in which shoes are commonly worn.

The time of day, shoe type, and foot positioning was kept consistent to minimize any day-to-day variability. It was found that foot-taps did not significantly differ on any of the days for either the live or force plate counts ($p > 0.05$). The video count however was significantly lower on Visit 1 compared to subsequent visits ($p < 0.05$). The between visits ICC was found to be high for the live, force plate, and video count (ICC = 0.88, 0.94, and 0.93 respectively). In the current study, the small but significantly lower foot-taps counted on the Visit 1 Video count may reflect a learning effect within either the subject or researcher themself. Similar % CV were found for each method for each of the four visits (Between 17.9–21.3%).

This study utilized three different counting methods (Live, Force Plate, and Video) during each trial to determine whether or not the number of foot-taps counted was influence by the counting method. Previous studies have utilized either automated counting device or visual inspection to count foot-tapping, yet have failed to clarify how they verified the reliability or validity of their counts. As to why we videotaped the trials, we hypothesized that by allowing the researcher to slow and re-watch foot-tapping trials that a larger number of foot-taps would be counted. Experience in our lab has shown that anything greater than

~50 taps in 10 s becomes difficult to count visually. Interestingly the current study found that the live count (50.5 ± 8.5 taps) did not significantly differ from the slowed video counts (51.3 ± 9.3 taps) ($p = 0.73$). It was found that the force plate (44.6 ± 8.2 taps) counted significantly less foot-taps than the other two methods ($p < 0.01$) but had similar (if not better) levels of reliability throughout all trials and methods of testing Furthermore, the coefficient of variation for the Live (16.8%), Force Plate (18.4%) and Video (18.1%) counting methods were all very similar to one another. Given the high precision of the force plate, capable of sensing even the lightest of taps, it is unlikely that the force plate "missed" any foot-taps. Instead it is much more likely that what the raters thought was a "tap" was not actually a tap at all because it did not make contact with the surface of the plate. As anyone who has spent a significant time counting foot-taps can tell you, it can become very difficult to count and judge "what is or isn't a foot-tap" in the moment, especially after your 500th foot-tap counted for the day.

Despite non-significant differences between the Live and Video counts, it is worth noting that the same researcher rated all of these trials. It is unlikely in a clinical setting that the same individual would be rating ALL trials themself. We were therefore interested to know how the FTT may differ when viewed by two different raters. A subset of the Video and Force Plate trials ($n = 26$) were reevaluated by a second rater and although raters tended to have significantly different counts ($p < 0.05$), the Pearson's R correlation between raters was found to be high for Video ($R = 0.83$) and Force Plate ($R = 0.99$) counts alike. According to the coefficient of variation, the Force Plate (92.3%) showed a relatively lower level of variability compared to that of the Video count (174.4%). Although the study did not have multiple raters for the live count, two previous studies found a higher inter-rater reliability for foot-tapping (kappa = 0.73−0.77) than the Babinski sign (kappa = 0.30−0.45) between multiple raters [1,15]. Furthermore, foot-tapping rate (83.8–85%) had a higher level of agreement with known UMN weakness than did the Babinski sign (56–63.7%) and had similar results in physicians and non-physicians alike [1,15].

In the current study, it was found that foot-tapping rate did not significantly correlate with any of the measured subject characteristics (Age, Height, Weight, Leg mass, Shoe size, or Activity level). (Pearson's R $p > 0.05$). This was expected given that the FTT is meant to be a measure of upper motor neuron function and is unlikely to be affected by these subject characteristics in otherwise healthy individuals. Contrary to other previous FTT studies, the authors did not find a significant negative correlation between age and foot-tapping ability [10,17,19]. This may however be explained by our comparatively younger and smaller sample size than studies that previously found a strong negative correlation between age and foot-taps performed.

Despite having measured foot-tapping rate with a force plate, the current study did not evaluate any of the additional kinematic data collected. The use of a force plate also allows for the measurement of individual foot-tap force, contact and clearance time, time to completion and changes in acceleration. In a study by Djurić-Jovičić et al., a similar foot-tapping protocol was evaluated using an in-house constructed force plate and inertial sensors [4]. By combining force plate and inertial sensor data, they were able quantify foot-taps as well as individual tap angle, speed, and force. This then allowed them to assess the more qualitative aspects of tapping such rhythmicity, regularity, smoothness, freezing, tremor, and motor patterns to discern differences between Parkinson's patients and controls. Although it was not done in the current study, the force tracings could also be used to qualitatively assess foot-tapping ability by examining the shape, size and trends of the foot-taps; a practice that would have clear clinical value in such populations.

The main findings of the current study were that the FTT tended to have a high level of reliability in healthy individuals. The FTT showed a high level of test-retest, day-to-day, and inter-rater reliability for all three counting methods (Pearson $R > 0.80$). Testing variables such as limb dominance and shoes tended to result in small (~1−2 foot-taps) but significant differences in foot-taps for all three counting methods. Despite significance however, these small differences are unlikely to amount to clinical significance. The force plate consistently showed a statistically significant lower number of foot-taps (~6 taps) compared to the live and video counting methods, but also showed a lower level of variability (SD) and had comparable levels of reliability. The Force plate (18.4%) also showed a similar % CV to both the Live (16.8%) and Video (18.1%) counts. When comparing the interrater reliability between two raters, the %CV between rater for the Force Plate (92.3%) was found to be better than that of the Video Count (174.4%). Contrary to some studies, a significant correlation between foot-tapping and subject age or any other subject characteristics was not found. From a clinical perspective this supports the hypothesis that foot-tapping rate reflects UMN functioning rather than confounding subject characteristics.

In the current study a "foot-tap" was defined as one complete cycle of dorsi flexion, requiring the ball of the foot to make contact with the force plate, and plantarflexion requiring complete clearance from the force plate. When visual counting was employed, if the ball of the foot did not appear to make contact, or for lack of better terms "hovered" over the plate, it was disregarded. Likewise, when utilizing the force plate, best judgement was used and any peaks that did not appear like a complete tap were disregarded. These disregarded force plate taps were most often ones who's force tracing did not return back to baseline during the dorsiflexion movement. These disregarded taps likely explain why the force plate counts tended to be lower than the two visual counting methods. When comparing two different raters, the reliability of the force plate count was found to be higher than that of the video counts. One could easily argue for either visual (video) counting or force plate as being the "Gold standard" of foot-tapping, however each has their own caveats. Visual counting has the benefit of simplicity at the cost of increased chances for human error and poorer agreement between raters. The force plate meanwhile tended to show a significantly lower count and requires special equipment but benefits from a better agreement between raters and comparable reliability to visual counting methods between trials. Given the increased potential for human error with visual counting methods, it's understandable why an automated counting method would be attractive to clinicians. Methods such as the force plate offers the additional, and relatively unexplored, foot kinematics data such as tapping force, time to peak tension, contact time and etc. The current study has taken a big step towards demonstrating the reliability and validating of using a force plate to conduct the FTT in healthy individuals. The logical next step is now for us to begin implementing the FTT in clinical populations.

Given the findings of the current study,a list of best practices when implementing the FTT in either health or clinical populationshas been compiled:

1. Use 10-second testing interval, procedures and subject placement as previously described in methods section. These tended to be the most utilized methods in previous studies and was found to be highly reliable in the current study.
2. In healthy individuals, test the dominant limb and in suspected bilateral deficits, test both the dominant and non-dominant limbs.
3. Conduct a familiarization to mitigate any learning effect and take the average of a minimum of 2–3 trials.
4. Conduct the test with shoes ON and standardize shoe type and foot placement between trials/visits. Tests with shoes ON may better reflect ADL.
5. Use an objective counting method such as a Force Plate to measure taps. If not possible, video record trials or have multiple counters present.
6. Regardless of counting methodology, have a clear criterion for what constitutes a "foot-tap". A true foot-tap requires the ball of the foot to strike and then clear the floor during the movement.

5. Conclusions

Previous studies have shown both quantitative and qualitative disparities in foot-tapping rates in clinical populations making it a useful index of UMN function. Despite its usefulness as a clinical measure, the FTT is underutilized or otherwise unknown by many. This may be in part due to a lack of standardized foot-tapping protocol but also the ambiguity around the best way to define and measure a foot-tap. Though our sample may only consist of a small sample of younger healthy individuals, we have taken the first big step by demonstrating a high reliability of the FTT across different trials, days, counters, and testing conditions. The authors have also demonstrated that it is possible to objectively test the FTT with a cheaper and readily available device (Neulog Force Plate Logger), thereby reducing counter bias and error. Besides offering an objective count of foot-taps, the force plate would also allow researchers to assess qualitative changes in foot-tapping trends. A list of best practices for conducting the FTT that should be followed by those who wish to implement the FTT in their future research has been compiled too.

Brief summary

What is already known:

- The FTT can be used as a means to assess upper motor neuron function and shows significant differences between clinical and healthy populations.
- Despite significance between groups, the reliability of the FTT has not been well demonstrated in either healthy or clinical populations.
- There tends to be methodological differences between studies, making results less generalizable and necessitates a more objective means of conducting the FTT.

What this study adds:

- The FTT showed a high level of reliability across several conditions, trials, and visits in healthy individuals. This suggests that a change in foot-tapping rate may represent a true effect, rather than being the result of error. Foot-tapping rates did not appear to be significantly correlated with any measured subject characteristics.
- All three counting methods (Live, Video, and Force Plate) were found to have high reliability. Despite the tendency for the force plate counts to be lower, it was found to be highly sensitive and displayed a higher reliability than the other two methods.
- A force plate may be useful for discerning questionable foot-taps. That is to say, it will tell us whether or not the foot actually made contact with the ground or just came close. Based on our observations of the FTT in the current study we recommend the use of a force plate and have compiled a list of best practices.

Conflict of interest

None declared.

Ethical approval

Study was approved by the University of Oklahoma Heath Sciences Center Institutional Review Board and all subjects were required to give written consent prior to participation.

Funding

None declared.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.foot.2021.101851.

References

[1] Appasamy PT, Dan TA, Bandyopadhyay V, Mathew V, Jeyaseelan V, Babu S, et al. Accuracy and reliability of Babinski sign versus finger and foot tapping in the diagnosis of corticospinal tract lesions. Neurol India 2018;66(5):1377–80. https://doi.org/10.4103/0028-3886.241370.

[2] Błażkiewicz M, Wit A. Compensatory strategy for ankle dorsiflexion muscle weakness during gait in patients with drop-foot. Gait Posture 2019;68:88–94. https://doi.org/10.1016/J.GAITPOST.2018.11.011.

[3] Cornwall MW, McPoil TG. Effect of ankle dorsiflexion range of motion on rearfoot motion during walking. J Am Podiatr Med Assoc 1999;89(6):272–7. https://doi.org/10.7547/87507315-89-6-272.

[4] Djurić-Jovičić M, Jovičić N, Radovanović S, Jecmenica-Lukic M, Belic M, Popović M, et al. Finger and foot tapping sensor system for objective motor assessment Senzorski sistem za objektivnu motornu procenu na osnovu tapping-a prstima i stopalom. Vojn Pregl 2018;75(1):68–77. https://doi.org/10.2298/VSP150502323D.

[5] Don R, Serrao M, Vinci P, Ranavolo A, Cacchio A, Ioppolo F, et al. Foot drop and plantar flexion failure determine different gait strategies in Charcot-Marie-Tooth patients. Clin Biomech 2007;22(8):905–16. https://doi.org/10.1016/j.clinbiomech.2007.06.002.

[6] Gunzler SA, Pavel M, Koudelka C, Carlson NE, Nutt JG. Foot-tapping rate as an objective outcome measure for Parkinson disease clinical trials. Clin Neuropharmacol 2009;32(2):97–102. https://doi.org/10.1097/WNF.0B013E3181684C22.

[7] Hinman M. Validity and reliability of a 10-second foot-tap test in older adults MOJ Gerontology &. Gerontology & Geriatrics 2021:4120194246. https://doi.org/10.15406/mojgg.2019.04.00175.

[8] Hyndman D, Ashburn A, Stack E. Fall events among people with stroke living in the community: circumstances of falls and characteristics of fallers. Arch Phys Med Rehabil 2002;83(2):165–70. https://doi.org/10.1053/APMR.2002.28030.

[9] Kalaycıoğlu C, Kara C, Atbaşoğlu C, Nalçacı E. Aspects of foot preference: differential relationships of skilled and unskilled foot movements with motor asymmetry. Laterality 2008;13(2):124–42. https://doi.org/10.1080/13576500701701704.

[10] Kent-Braun JA, Ng AV. Specific strength and voluntary muscle activation in young and elderly women and men. J Appl Physiol 1999;87(1):22–9. https://doi.org/10.1152/jappl.1999.87.1.22.

[11] Kent-Braun JA, Walker CH, Weiner MW, Miller RG. Functional significance of upper and lower motor neuron impairment in amyotrophic lateral sclerosis. Muscle Nerve 1998;21(6):762–8. https://doi.org/10.1002/(SICI)1097-4598(199806)21:6<762::AID-MUS8>3.0.CO;2-5.

[12] Knights RM, Moule AD. Normative and reliability data on finger and foot tapping in children. Percept Mot Skills 1967;25(3):717–20. https://doi.org/10.2466/pms.1967.25.3.717.

[13] Lamontagne A, Malouin F, Richards C, Dumas F. Mechanisms of disturbed motor control in ankle weakness during gait after stroke. Gait Posture 2002;15(3):244–55. https://doi.org/10.1016/S0966-6362(01)00190-4.

[14] Menz HB, Morris ME, Lord SR. Foot and ankle risk factors for falls in older people: a prospective study. J Gerontol A Biol Sci Med Sci 2006;61(8):866–70. https://doi.org/10.1093/gerona/61.8.866.

[15] Miller TM, Johnston SC. Should the Babinski sign be part of the routine neurologic examination? Neurology 2005;65(8):1165–8. https://doi.org/10.1212/01.wnl.0000180608.76190.10.

[16] Mitsumoto H, Ulug AM, Pullman SL, Gooch CL, Chan S, Tang M-X, et al. Quantitative objective markers for upper and lower motor neuron dysfunction in ALS [Accessed 6 August 2018]. 2007. www.neurology.org.

[17] Ng AV, Miller RG, Gelinas D, Kent-Braun JA. Functional relationships of central and peripheral muscle alterations in multiple sclerosis. Muscle Nerve 2004;29(6):843–52. https://doi.org/10.1002/mus.20038.

[18] Nitz J, Choy NL. The relationship between ankle dorsiflexion range, falls and activity level in women aged 40 to 80 years. NZ J Physiother 2004;32(3):121–5 [Accessed 3 October 2019], https://www.researchgate.net/profile/Jennifer_Nitz/publication/43767790_The_relationship_between_ankle_dorsiflexion_range_falls_and_activity_level_in_women_aged_40_to_80_years/links/5761ec3808aeeada5bc5034c/The-relationship-between-ankle-dorsiflexion-rang.

[19] Numasawa T, Ono A, Wada K, Yamasaki Y, Yokoyama T, Aburakawa S, et al. Simple foot tapping test as a quantitative objective assessment of cervical myelopathy. Spine (Phila Pa 1976) 2012;37(2):108–13. https://doi.org/10.1097/BRS.0b013e31821041f8.

[20] Sharma KR, Kent-Braun J, Mynhier MA, Weiner MW, Miller RG. Evidence of an abnormal intramuscular component of fatigue in multiple sclerosis. Muscle Nerve 1995;18(12):1403–11. https://doi.org/10.1002/mus.880181210.

[21] Tanigawa M, Stein J, Park J, Kosa P, Cortese I, Bielekova B. Finger and foot tapping as alternative outcomes of upper and lower extremity function in multiple sclerosis. Mult Scler J Exp Transl Clin 2017;3(1). https://doi.org/10.1177/2055217316688930. 2055217316688930.

[22] Thijssen DH, Paulus R, van Uden CJ, Kooloos JG, Hopman MT. Decreased energy cost and improved gait pattern using a new orthosis in persons with long-term

stroke. Arch Phys Med Rehabil 2007;88(2):181–6. https://doi.org/10.1016/J.APMR.2006.11.014.
[23] Wiszomirska I, Błażkiewicz M, Kaczmarczyk K, Brzuszkiewicz-Kuźmicka G, Wit A. Effect of drop foot on spatiotemporal, kinematic, and kinetic parameters during gait. Appl Bionics Biomech 2017;2017:3595461. https://doi.org/10.1155/2017/3595461.
[24] York G, Chakrabarty S. A survey on foot drop and functional electrical stimulation. Int J Intell Robot Appl 2019;3(1):4–10. https://doi.org/10.1007/s41315-019-00088-1.
[25] Zemková E, Kyselovičová O, Jeleň M, et al. The effect of 3 months aerobic and resistance training on step initiation speed and foot tapping frequency in the overweight and obese. Sport Sci Health 2017;13(2):331–9. https://doi.org/10.1007/s11332-017-0362-9.