Clinical Neurophysiology Practice 8 (2023) 16–23

Contents lists available at ScienceDirect

Clinical Neurophysiology Practice

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

Research paper

Investigating the intra-session reliability of short and long latency afferent inhibition

Ravjot S. Rehsi [a], Karishma R. Ramdeo [a], Stevie D. Foglia [b], Claudia V. Turco [c], Faith C. Adams [a], Stephen L. Toepp [a], Aimee J. Nelson [a,b,*]

[a] Department of Kinesiology, McMaster University, Canada
[b] School of Biomedical Engineering, McMaster University, Canada
[c] Faculty of Medicine and Dentistry, University of Alberta, Canada

ARTICLE INFO

Article history:
Received 24 October 2022
Received in revised form 24 November 2022
Accepted 5 December 2022
Available online 21 December 2022

Keywords:
Afferent inhibition
Transcranial Magnetic Stimulation
Reliability
SAI
LAI

ABSTRACT

Objective: To establish the intrasession relative and absolute reliability of Short (SAI) and Long-Latency Afferent Inhibition (LAI). These findings will allow us to guide future explorations of changes to these measures.
Methods: 31 healthy individuals (21.06 ± 2.85 years) had SAI and LAI obtained thrice at 30-minute intervals in one session. To identify the minimum number of trials required to reliably elicit SAI and LAI, relative reliability was assessed at running intervals of 5 trials.
Results: SAI had moderate–high, and LAI had high-excellent relative reliability. Both SAI and LAI had high amounts of measurement error. LAI had high relative reliability when only 5 frames of data were included, whereas SAI required ~20–30 frames of data for the same. For both SAI and LAI, individual smallest detectable change was large but was reduced at the group level.
Conclusions: SAI and LAI can be used for both diagnostic purposes and to assess group level change but have limited utility in assessing within-individual changes.
Significance: These results can be used to inform future work regarding the utility of SAI and LAI, particularly in terms of their ability to identify particularly high or low values of afferent inhibition.

1. Introduction

Afferent Inhibition is a phenomenon reflecting decreases in Transcranial Magnetic Stimulation (TMS)-induced motor output following peripheral nerve stimulation. Afferent inhibition is subdivided into two separate phenomenon elicited at distinct interstimulus intervals (ISI); short-latency afferent inhibition (SAI) is elicited at ISIs of 18–28 ms (Ni et al., 2011; Tokimura et al., 2000) and long-latency afferent inhibition (LAI) is elicited at ISIs of 200–1000 ms (Chen et al., 1999).

Alongside reflecting the integrity of sensorimotor systems, afferent inhibition is also associated with neural functioning in neurological disorders. Particularly, SAI is reduced in various disorders of impaired cognition including Alzheimer's Disease (AD) (Di Lazzaro et al., 2004, 2002), Parkinson's Disease (PD) with Dementia (Celebi et al., 2012; Yarnall et al., 2013), and Mild Cognitive Impairment (MCI) (Nardone et al., 2012; Peter et al., 2016) whereas LAI is reduced in disorders of the sensory-motor system such as PD (Sailer et al., 2003). The differentiation between the two types of afferent inhibition may lie in their molecular basis, with SAI reflecting both cholinergic and gamma-aminobutyric acid type A (GABA) receptor activity, compared to LAI which appears to be solely GABAergic (Di Lazzaro et al., 2007, 2005b; Turco et al., 2018).

Importantly, despite their emerging use in the sensorimotor research, the reliability of afferent inhibition ranges from low to strong in most cases (Brown et al., 2017; Toepp et al., 2021; Turco et al., 2019). As such, it is difficult to interpret changes in afferent inhibition within an experimental setting as real (i.e. induced by the intervention) or simply due to inherent variability within the measurement. Determining the reliability of both SAI and LAI would allow experimenters to assess whether changes in these measures are due to actual physiological change.

Reliability can be partitioned into two forms: relative and absolute reliability. Relative reliability reflects the ability of a measure to allow for individuals to maintain their position relative to each other, by providing the correlation between repeated measures (Beaulieu et al., 2017a; Šerbetar, 2015). Intraclass Correlation Coefficients (ICC) are a common form of assessing relative reliability,

* Corresponding author at: Department of Kinesiology, McMaster University, 1280 Main Street West, Hamilton, ON L8S 4K1, Canada.
E-mail address: nelsonaj@mcmaster.ca (A.J. Nelson).

https://doi.org/10.1016/j.cnp.2022.12.001
2467-981X/

reflecting the between subject variability as a function of total error (Šerbetar, 2015; Weir, 2005). Higher ICC values indicate a greater ability of the measure to identify differences between individuals, an essential trait for the diagnostic usage of a measure. Absolute reliability reflects the total variability of a measure across time (Beaulieu et al., 2017a; Šerbetar, 2015). Unlike relative reliability, absolute reliability is not dependent on the characteristics of the sample (Turco et al., 2019; Weir, 2005). Traditionally, absolute reliability can be assessed through the Standard Error of Measurement (SEMeas), reflecting the standard deviation of all errors in the measure (Beckerman et al., 2001; Hopkins, 2000; Šerbetar, 2015). Smaller SEMeas values result in a measure that is less likely to change with variability. The SEMeas can be further used to calculate the Smallest Detectable Change (SDC) in the measure at both the group and individual level. The $SDC_{Individual}$ reflects the smallest amount of change required to be seen in an individual across time that can be attributed to sources other than measurement error (Beaulieu et al., 2017a; Šerbetar, 2015). SDC_{Group}, similar to $SDC_{Individual}$, reflects the smallest amount of change needed to be seen at the group level which can be attributed to real change (Beaulieu et al., 2017a; Šerbetar, 2015).

Previous work has only explored the intersession reliability of SAI and LAI, with the three studies reporting the relative reliability of both measures as ranging from low to strong (Brown et al., 2017; Toepp et al., 2021; Turco et al., 2019). Even more limited are quantifications of the absolute reliability of these measures, with a single study reporting low absolute reliability as indicated by high SEMeas and high SDC values, indicating large deviations in the measure are needed in order to be significantly different from measurement error (Turco et al., 2019). The literature therefore seems to suggest that, for both SAI and LAI, while individuals scores tend to be consistent relative to the spread of the data, the variability in true scores across time is high.

To date, no study has examined the absolute and relative reliability of SAI and LAI when collected at multiple time points within a single session. This is important in informing whether the measures are stable enough to reflect changes within a session when an intervention is administered. Furthermore, the literature exploring reliability of other TMS measures often include a frame-by-frame analysis in order to determine the smallest number of frames required to produce a reliable measure (Biabani et al., 2018; Goldsworthy et al., 2016) and this has yet to be explored for afferent inhibition. In the present study, we explore both the absolute and relative intrasession reliability of SAI and LAI evoked by stimulation of the Median Nerve (MN) at three time points separated by 30 min in 31 healthy, young adults. Given that our past work has indicated higher intersession relative reliability for LAI compared to SAI, we predict the same relationship in our intrasession exploration (Turco et al., 2019). These findings can help to inform research exploring the impact of various short-term interventions on neuroplasticity as assessed by afferent inhibition.

2. Experimental procedure

2.1. Participants

31 healthy, young, right-handed participants took part in this experiment (18 females; age = 21.06 ± 2.85 years). Participants attended one 3-hour session. All individuals participated in the experiment after 12 PM in order to account for rapid fluctuations in diurnal cortisol levels that may influence TMS measures (Milani et al., 2010). Participants were screened for contraindications to TMS prior to taking part in the study, and handedness was confirmed using a modified version of the Edinburgh Handedness Questionnaire. All individuals provided informed written consent prior to study onset. This research was approved by the Hamilton Integrated Research Ethics Board and conformed to the Declaration of Helsinki.

2.2. Electromyography (EMG)

Surface electrodes (9 mm, Ag-AgCl) were used to record activity from the first dorsal interosseous (FDI) muscle of the right hand. The active electrode was placed over the muscle belly, and the reference electrode was placed on the metacarpal joint of the first digit. A grounding electrode was placed at the styloid process at the wrist. EMG signals were first magnified x1000 and were then band pass filtered between 20 and 2.5 kHz (Intronix Technologies Corporation Model 2024F, Bolton, Canada). Data were digitized at 5 kHz by a digital to analog converter (Power1401; Cambridge Electronics Design, Cambridge, UK), and were then analyzed through commercial software (Signal v7.02; Cambridge Electronics Design, Cambridge, UK).

2.3. Electroencephalography (EEG)

To determine the latency of the N20 potential, Somatosensory Evoked Potentials (SEPs) were recorded by placing EEG electrodes on the scalp at C3′ over the postcentral gyrus and referencing activity to Fz (International 10–20 System). A ground electrode was placed on the clavicle ipsilateral to the stimulated nerve. Electrical stimulation (Digitimer DS7AH, 200 μs square wave pulses) was delivered using a bar electrode at the right MN at the wrist. Stimulation was delivered at 3 Hz and at an intensity corresponding to the motor threshold (MT) of the participant. 500 stimuli were delivered and time-locked averaged to determine the latency of the N20 potential of the SEP.

2.4. Transcranial Magnetic Stimulation (TMS)

TMS was performed with a monophasic waveform using a figure of eight coil connected to a Magstim 200^2 stimulator. The coil was oriented at a 45-degree angle in the posterior-anterior direction over the motor representation of the right FDI over the left motor cortex. The motor hotspot was registered using Brainsight Neuronavigation Software and was defined as the location that elicited the largest Motor Evoked Potential (MEP) in the right FDI muscle. The same exact Magstim system and stimulation coil was used for all data collection in this study. TMS delivery was performed by author RSR for all participants and all time points.

2.5. Resting Motor Threshold (RMT)

Resting motor threshold is defined as the lowest intensity of TMS stimulation that can produce an MEP $\geq$ 50 μV in peak-peak amplitude, 50% of the time (Rossini et al., 2015). RMT was obtained using a predictive algorithm using ML-PEST, which continuously predicted the next TMS intensity at which to deliver stimulation. The starting intensity was set to 37% of the max stimulator output, with 20 stimuli being delivered in order to predict the RMT (Ah Sen et al., 2017).

2.6. Afferent inhibition

The intensity of the nerve stimulus, also known as the conditioning stimulus (CS), was set to the MT for a visual twitch in the right Abductor Pollicis Brevis (APB) muscle following MN stimulation. The MT of the MN has previously been found to correlate with 50% of the maximum sensory nerve action potential elicited by the nerve, which is the intensity at which maximum inhibition occurs

for both SAI and LAI (Bailey et al., 2016; Turco et al., 2017). TMS stimulation was delivered at the lowest intensity needed to elicit a 1 mV peak-peak MEP in the target muscle. The ISI between the nerve stimulus and TMS pulse was set to the latency of the N20 component + 4 ms for SAI (Di Lazzaro et al., 2005b; Tokimura et al., 2000; Turco et al., 2019), and 200 ms for LAI (Chen et al., 1999). SAI and LAI were collected in the same order across participants. Unconditioned pulses contain only the TMS pulse and are referred to as Test Stimulus (TS) trials. Conditioned pulses contain both the nerve stimulus (i.e., CS) and TMS pulse and are referred to as Conditioning Stimulus-Test Stimulus (CSTS) trials. For each circuit, 40 TS frames and 40 CSTS frames were delivered pseudo-randomly such that there were never 3 of any one type delivered in succession. SAI and LAI were collected at three time-points within the session, separated by 30-minute breaks (T1, T2, T3).

2.7. Statistical analysis

EMG trials were discarded if the peak-peak amplitude of the signal was greater than 50 μv in a 100 ms window directly before the TMS pulse. Outliers were identified and removed using Grubb's Test. Normality was assessed using Shapiro-Wilks tests, and heteroscedasticity was assessed using Bland-Altman plots (Damron et al., 2008; Schambra et al., 2015). Bland-Altman plots were created comparing the respective variables at T1-T2, T1-T3, and T2-T3. Afferent inhibition was calculated as a ratio of the mean peak-peak MEPs amplitudes of the CSTS to TS MEPs. Repeated measure ANOVAs were used to compare SAI and LAI across the three time points to discover if systematic error was present. Significance was set to alpha ≤ 0.05, and Bonferroni corrections were used for multiple comparisons.

2.8. Reliability assessment

Relative reliability was assessed on running averages of the MEP amplitudes, defined as the average of all preceding trials. ICCs were evaluated for both SAI and LAI using both the overall measure, as well as at intervals of every 5 trials. As all data was collected by a single experimenter, a two-way random effects model was used (ICC 2,k). To aid in the interpretation of the measures, coefficients of variation (CV) were also calculated for each measure. ICCs were categorized based on suggested cut off points where ICC with 95% CI above 0.9 is Excellent; 0.75 < ICC < 0.9 is High; 0.5 < ICC < 0.75 is Moderate; and ICC < 0.5 is considered Low (Koo and Li, 2016; Portney and Watkins, 2009). Absolute reliability was assessed using SEMeas values obtained over the whole dataset. SEMeas was converted to be expressed as a percentage of the mean using the formula %SEMeas = (SEMeas/mean)*100%, in order to provide a unitless assessment of the measurement error (Weir, 2005). % SEMeas < 10% was used as a cut off to indicate low measurement error (Schambra et al., 2015). These SEM values were then used to calculate both $SDC_{Individual}$ and SDC_{Group}, which provide the minimum amount of change needed to be seen at the individual and group level to be considered real change and not due to measurement error (Schambra et al., 2015).

3. Results

All participants tolerated the experimental procedures well, with no reported adverse effects. For one participant, because more than 50% of the LAI dataset at T1 was removed due to the inability to relax the FDI muscle, this participant was excluded from the analysis. For another, LAI data could not be processed due to excessive EMG during collection. As well, Grubb's test necessitated the removal of a singular outlier from the LAI dataset at T1. The outlier datapoint was removed from the overall analysis of reliability only and was included in the frame-by-frame analysis method. Therefore, for overall assessments of reliability, SAI yielded n = 31 and LAI yielded n = 28. For the frame-by-frame analysis of relative reliability SAI yielded n = 31 and LAI yielded n = 29.

The datasets for both SAI and LAI were normally distributed. Bland-Altman plots indicated that homoscedasticity was maintained for SAI at all time points, but was violated for LAI comparing T1—T3, with $R^2 = 0.1$. Corrections for heteroscedasticity normally require a log-transformation on the dataset. However, given that log-transformations change the dataset to a ratio scale, we did not perform the log transformation, and analysed the data with an assumption of heteroscedasticity, as done previously (Liu and Au-Yeung, 2014; Ngomo et al., 2012; Sankarasubramanian et al., 2015).

TMS was delivered at the lowest intensity to elicit a ~1 mV peak-peak MEP, which equated to 134 ± 15 % of RMT averaged across participants when 1 mV was first assessed. The 1 mV peak-peak MEP had relatively large measurement error (% SEM = 14.74), while also having high relative reliability (ICC = 0.79, 95% CI [0.65 – 0.90]). The average latency for the N20 potential in the dataset was 18.21 ± 0.82 ms.

3.1. Short-Latency Afferent Inhibition (SAI)

The mean ± SD of SAI across the three timepoints is shown in Fig. 1A I. A repeated measures ANOVA indicated a main effect of State ($F_{(1, 30)} = 63.279$, $p < 0.001$, $\eta_p^2 = 0.678$) with no other significant main or interaction effects. Therefore, we can conclude a significant difference in peak-peak MEPs between states, such that CSTS was suppressed relative to TS. For SAI ratio, a repeated measures ANOVA indicated no main effect of Time on SAI magnitude ($F_{(2,60)} = 0.702$, $p = 0.499$, $\eta_p^2 = 0.023$) indicating that SAI was not different across timepoints, and that there was no detectable systematic error in the dataset. Overall, 30 out of 31 participants displayed inhibition at a minimum of one time point (Fig. 1B). Compared to the TS, the CSTS was reduced by an average of 26.78% at T1, 31.03% at T2, and 29.96% at T3.

Overall, high relative reliability was observed for the SAI inhibition ratio (ICC = 0.81; 95% CI [0.66 – 0.90]), and moderate reliability for the conditioned frames (ICC = 0.74; 95% CI [0.52 – 0.87]). The % SEMeas for SAI was 20.77, indicating large amounts of measurement error. For SAI, the $SDC_{Individual}$ indicates that a minimal change of 40.73 is needed to be seen over time to be considered physiological change on an individual level. Our data also indicates that for our sample size of 31 individuals, a change of 7.32 is needed to be considered physiological change in the measure at a group level (Fig. 1A II.).

The ICCs as a function of the number of stimulus pairs are shown in Fig. 1C. For SAI, the first 5 collected datapoints indicated moderate reliability of the measure (ICC = 0.52). The 95% CI of the first 5 frames of data was also noticeably large, with a width of ~0.6 (Fig. 1C). Increasing the number of frames included in the analysis steadily increased the relative reliability of the measure, with high reliability being achieved with 20 stimulus pairs (ICC = 0.77).

3.2. Long-Latency Afferent Inhibition (LAI)

The mean ± SD of LAI across the three timepoints is shown in Fig. 2A I. A repeated measures ANOVA indicated a main effect of State ($F_{(1,28)} = 37.907$, $p < 0.001$, $\eta_p^2 = 0.575$) and Time ($F_{(2,56)} = 3.225$, $p = 0.047$, $\eta_p^2 = 0.103$). However, following removal of the CSTS and TS datapoints that constituted the outlier LAI ratio as indicated by the Grubbs test, there was only a significant main effect of State ($F_{(1,27)} = 45.778$, $p < 0.001$, $\eta_p^2 = 0.629$). Therefore,

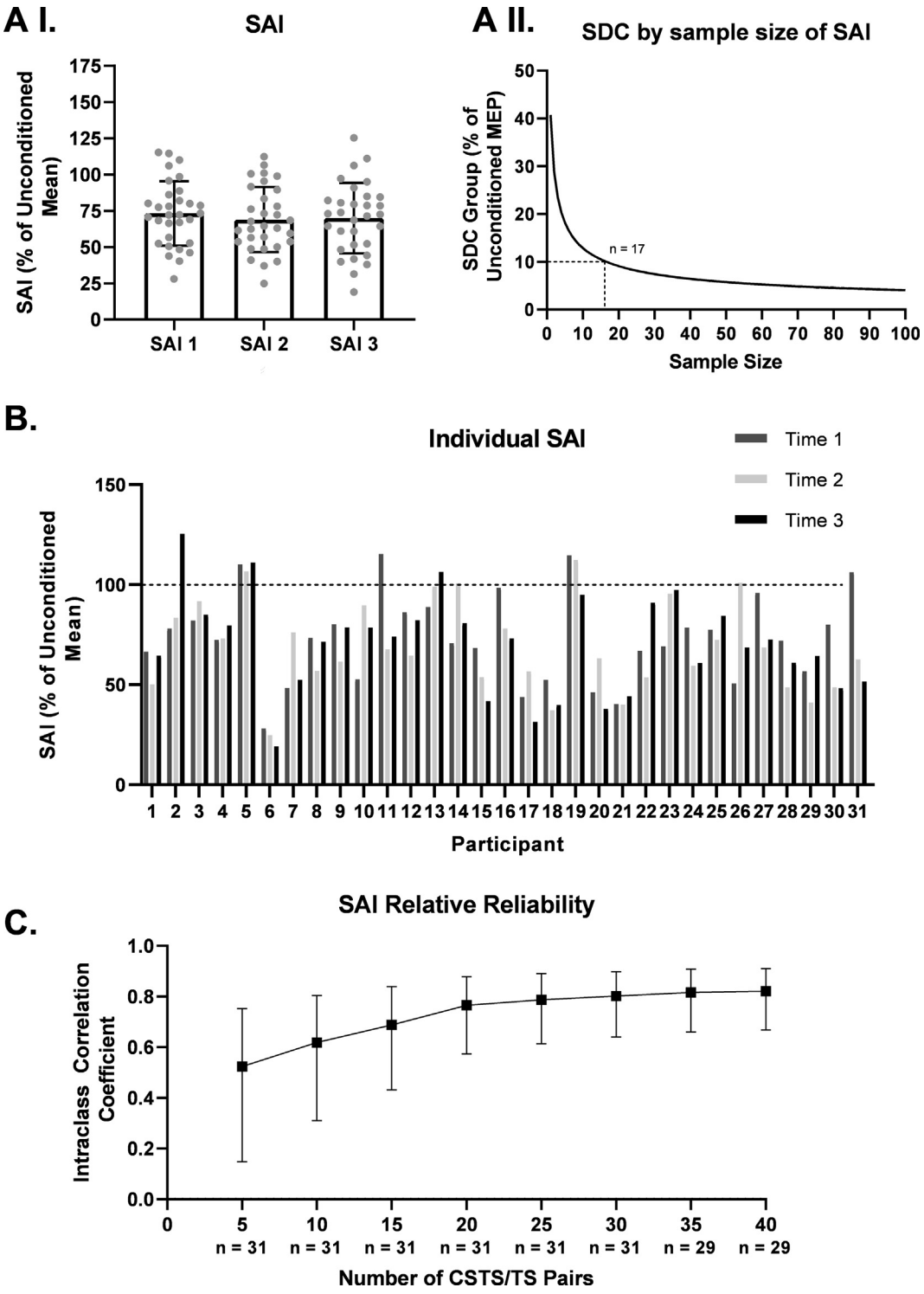

Fig. 1. Group Averaged and Individual SAI Data. A I. Average SAI, expressed as a % of the unconditioned mean, alongside Standard Deviation error bars, and individual scores. The CV was 30 at T1, 32 at T2, and 35 at T3. A II. SDC_{Group} presented as a function of the sample size. B. Individual scores on SAI by timepoint for each participant. C. ICC values depicted as a function of the number of CSTS/TS pairs included to determine SAI, along with 95% Confidence Intervals.

we can conclude that MEPs were significantly suppressed during CSTS compared to TS alone. A repeated measures ANOVA did not show a main effect of time on LAI magnitude ($F_{(2,\ 54)} = 0.377$, $p = 0.688$, $\eta_p^2 = 0.014$) indicating that LAI was not different across time points. All participants showed inhibition at a minimum of one time point (Fig. 2B). Compared to the TS, the CSTS was reduced by an average of 35.40% at T1, 38.85% at T2, and 38.06% at T3.

Overall, excellent relative reliability was observed for both the LAI inhibition ratio (ICC = 0.90, 95% CI [0.80 – 0.95]) and the conditioned frames (ICC = 0.91, 95% CI [0.83 – 0.96]). The %SEMeas for LAI was 24.92, indicating that there was large measurement error in the dataset. The $SDC_{Individual}$ indicates that a difference of 43.21 is required for the change in LAI to be considered significant in a given individual. Our calculations also indicate that for our sample

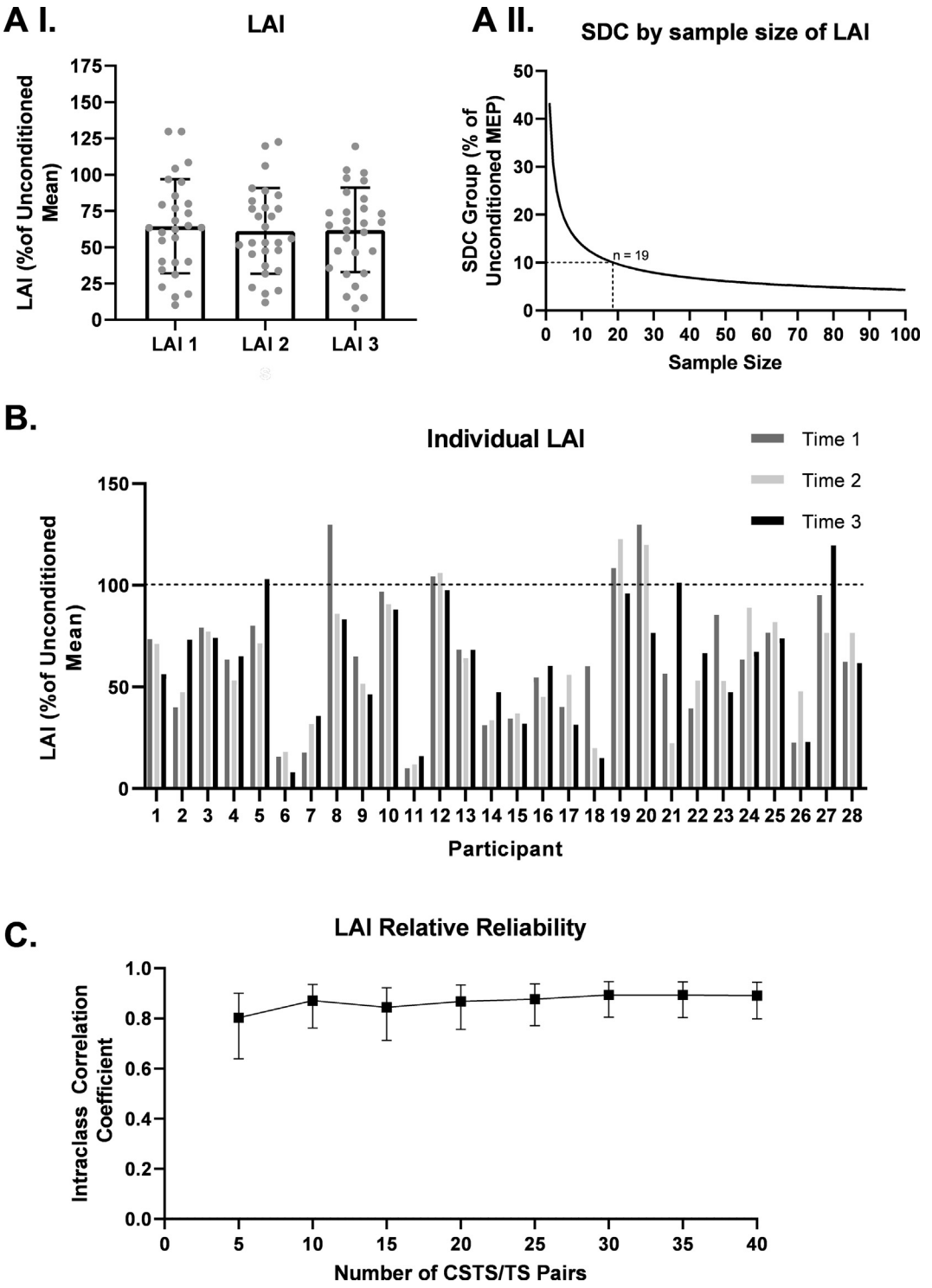

Fig. 2. Group Averaged and Individual LAI Data. A I. Average LAI, expressed as a % of the unconditioned mean, alongside Standard Deviation error bars, and individual scores. The CV was 50 at T1, 48 at T2, and 47 at T3. A II. SDC_{Group} presented as a function of the sample size. B. Individual scores on LAI by timepoint for each participant. C. ICC values depicted as a function of the number of CSTS/TS pairs included to determine LAI, alongside 95% Confidence Intervals.

size of 28 individuals, a considerably smaller change of 8.17% is needed in order to be considered physiological change at the group level analysis (Fig. 2A II).

The ICCs as a function of number of stimulus pairs is shown in Fig. 2C. The running-averages are pooled into groups of 5, and the ICC values are shown individually for the 8 created groups. For LAI, the first 5 collected datapoints indicated high relative reliability of the measures (ICC = 0.80), with a narrow 95% confidence interval (0.64 – 0.90) (See Fig. 2C). Increasing the number of CSTS/TS pairs continued to increase the ICC value until all 40 pairs of data were included, indicating high-excellent relative reliability of the measure (ICC = 0.89). The 95% confidence inter-

val also continued to narrow as the number of frames analysed increased.

4. Discussion

The goal of the present study was to assess the intrasession reliability of SAI and LAI. To do so, SAI and LAI were elicited by delivering peripheral nerve stimulation to the MN followed by TMS to M1. We assessed the relative reliability of these measures using ICCs overall and as a function of the total number of stimuli pairs. The absolute reliability of these measures was assessed through calculations of the %SEMeas and SDC.

4.1. Relative reliability

Relative reliability reflects the ability of a measure to consistently identify individuals within a dataset during repeated testing. Within this study, we found that SAI had moderate-high relative reliability, whereas LAI had high-excellent relative reliability, supporting our hypothesis that LAI would have higher relative reliability than SAI and is in line with previous findings (Turco et al., 2019). The differences in relative reliability between SAI and LAI observed both here and previously in the literature, may be due to differences in the between-subject variation of these measures. For SAI, the calculated CVs in our study ranged from 30% to 35%, whereas for LAI they ranged from 47% to 50%. This same differentiation was also seen in Turco et al.'s work, with LAI having higher levels of between subject variability when compared to SAI (Turco et al., 2019). Given that ICC values are a function of the between-subject variability in a sample, with large amounts of between-subject variability leading to higher estimates of relative reliability, this likely explains the trends in reliability seen across the literature and in the present study (Bruton et al., 2000; Weir, 2005). We recently published a large retrospective analysis of SAI and LAI data which also showed larger amounts of between-subject variability for LAI compared to SAI, with a plausible reasoning being the different neurological pathway that LAI may traverse (Toepp et al., 2021). We theorize that because LAI is evoked at longer ISIs, there is possible activation of additional brain regions including the basal-ganglio-thalamocortical loop, the posterior parietal cortex, and the secondary somatosensory cortex (Toepp et al., 2021). The potential activation of these various regions introduces several avenues of variability between individuals, which is not seen in SAI. Similar trends are seen with reliability comparisons of Short and Long Intracortical Inhibition (SICI & LICI) where LICI is evoked at longer ISIs compared to SICI and may involve more cortical regions, thus greater variation between individuals and higher levels of relative reliability (Schambra et al., 2015).

Previous work has quantified the reliability of these measures between sessions, allowing us the opportunity to compare intrasession to intersession reliability (Brown et al., 2017; Turco et al., 2019). However, as Brown et al. (2017) did not include CVs in their work, it is difficult to compare our ICC values with theirs, as ICC is a function of between-subject variability (Turco et al., 2019; Weir, 2005). Similar CVs between our work and Turco et al.'s allow for such comparisons and indicate that both SAI and LAI have better *intrasession* relative reliability compared to *intersession* relative reliability (Turco et al., 2019). Similar findings have been reported when comparing the relative reliability of other single and paired-pulse TMS measures, with intrasession reliability being higher than intersession (Biabani et al., 2018; Goldsworthy et al., 2016). Given that relative reliability reflects the ability to identify individuals on repeated testing, it is likely easier to do so in an intrasession format due to less opportunity for change in the various neurochemical and biological factors that govern afferent inhibition (Bruton et al., 2000; Di Lazzaro et al., 2005a; Turco et al., 2018). Furthermore, procedural factors are also unlikely to change with an intrasession model compared to intersession, as electrodes do not need to be removed and re-applied, as well as the hotspot staying consistent within a session. Therefore, for future work exploring the modulation of afferent inhibition across time, we recommend intrasession approaches if possible.

Analysing relative reliability as a function of the number of stimuli pairs indicated that, for LAI, the relative reliability is initially high and continues to stay high as more pairs are added, with the greater change being the narrowing of the 95% Confidence Intervals. This again, can likely be explained by the large amounts of between-subject variation in the LAI measure leading to overall higher values for the ICC. However, for SAI we found that the relative reliability is initially low when only 5 pairs of data are included and ICCs continue to increase as more pairs are included in the average calculation, eventually reaching a "high" level once 20 pairs are included. Similar trends have been explored with assessments of both, single and paired-pulse TMS measures, indicating that 20–30 frames are needed to produce reliable estimates (Biabani et al., 2018; Goldsworthy et al., 2016). Future studies exploring afferent inhibition may consider collecting between 20 and 30 stimuli pairs to ensure high relative reliability, without providing unnecessary stimulation to participants.

4.2. Absolute reliability

Absolute reliability assesses the repeatability of scores through repeated testing (Bruton et al., 2000). We found that both SAI and LAI had high levels of measurement error, quantified by the % SEMeas being greater than 10%. This high measurement error of SAI and LAI has also been documented previously in between-session explorations; however, while the error for SAI is largely the same as that reported here, the previously reported measurement error for LAI is ~20% greater than reported in our work (Turco et al., 2019). Given that SEMeas is considered to be a fixed characteristic of the measure, which is uninfluenced by the sample demographics, it stands to reason that for LAI particularly, there is more measurement error between sessions compared to within sessions. The lack of consistency may be attributed to the largely unknown biochemical underpinnings of LAI, as suggested elsewhere (Turco et al., 2019). This may be further explained by the longer ISI traversed by the LAI circuit as discussed earlier; similar to our SAI findings, previous work exploring the absolute reliability of SICI indicated no significant difference in intersession and intrasession measurement error, perhaps owing to the shorter ISI of this circuit (Samusyte et al., 2018). However, validation of this hypothesis would require investigations of the inter and intrasession absolute reliabilities of LICI as well. We currently recommend that future studies exploring interventional modulations of LAI consider intrasession models, due the decreased amount of measurement error present.

Calculations of SDC indicated that, at the individual level, large amounts of within-session change in both SAI and LAI is required in order for the change to be considered significant. This limits the utility of SAI and LAI to detect significant change on an individual level. Previous work has cited an SDC_{Group} level < 10% as a recommended threshold (Schambra et al., 2015). Given this recommended threshold, both SAI and LAI have utility at the group level, as sample sizes of $\geq$17 and $\geq$19 respectively lead to SDC_{Group} < 10%. This indicates that for adequately attainable sample sizes, SAI and LAI are able to establish whether interventions lead to changes across groups. High levels of $SDC_{Individual}$ have been reported for other paired-pulse TMS measures as well; however, single pulse TMS measures such as MT generally present with much higher levels of absolute reliability (Beaulieu et al., 2017b,

2017a; Samusyte et al., 2018; Schambra et al., 2015). For both SAI and LAI, there is limited utility of the measures as clinical biomarkers of change owing to the large $SDC_{Individual}$. However, the measures may be useful on the group level, for example to determine whether treatments or therapies are able to modulate SAI and LAI in patient populations through group statistics.

4.3. Limitations

Participants were recruited from the McMaster University student population and as such, the present study is limited to the sample demographics of healthy young adults. Given the inconsistent reports of modulations in SAI and LAI with age, our study may not be applicable to an older demographic (Bhandari et al., 2016; Degardin et al., 2011; Yarnall et al., 2016; Young-Bernier et al., 2012) or special populations. Specifically, older adults alongside patients with neurological disorders such as Alzheimer's Disease or Parkinson's Disease with and without dementia may present different reliabilities of SAI and LAI, as the measures are reduced in these populations. To date, there are no publications examining the reliability of SAI and LAI in special populations. Further work should be done to explore the reliability of SAI and LAI in these populations.

As well, the reliability statistics employed in this research only examine the reliability of SAI and LAI when obtained with the same experimenter, using the same equipment. However, reliability can differ between experimenters, and may be examined by modifying the equation used to calculate the ICC (Weir, 2005). As such, our results can only be applied to past and future work that explores the reliability of these circuits with a single experimenter.

5. Conclusions

This study is the first to explore the intrasession relative and absolute reliability of SAI and LAI, particularly as a function of the number of stimuli pairs. High ICC levels in our within-session design and in previous between-session explorations suggest that SAI and LAI can be used for identifying individuals with high or low values of reliability. For within-session explorations in particular, 20–30 CS-CSTS pairs should be collected to achieve this reliability in SAI and LAI. We also found that for both within- and between-session models, SAI and LAI have limited utility to identify differences between individuals, but can be used to assess group-level changes when appropriate sample sizes are collected. Alongside establishing the minimum number of CS-CSTS pairs required to achieve high levels of relative reliability in these measures, future work should also look to explore methods to improve reliability as well.

Funding

The authors gratefully acknowledge financial support from the Natural Sciences and Engineering Research Council of Canada to AJN (RGPIN-2020-06757).

Author Contributions

RSR, CVT & AJN conceptualized the study; RSR, KRR, SDF, FCA, and SLT collected the data; RSR processed and analysed the data; RSR and SLT created the figures; RSR wrote the final manuscript; all authors contributed to the revision and editing of the final manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Ah Sen, C.B., Fassett, H.J., El-Sayes, J., Turco, C.V., Hameer, M.M., Nelson, A.J., Tremblay, F., 2017. Active and resting motor threshold are efficiently obtained with adaptive threshold hunting. PLoS One 12 (10), e0186007. https://doi.org/10.1371/journal.pone.0186007.

Bailey, A.Z., Asmussen, M.J., Nelson, A.J., 2016. Short-latency afferent inhibition determined by the sensory afferent volley. J. Neurophysiol. 116, 637–644. https://doi.org/10.1152/jn.00276.2016.

Beaulieu, L.D., Flamand, V.H., Massé-Alarie, H., Schneider, C., 2017a. Reliability and minimal detectable change of transcranial magnetic stimulation outcomes in healthy adults: A systematic review. Brain Stimul. 10, 196–213. https://doi.org/10.1016/j.brs.2016.12.008.

Beaulieu, L.D., Massé-Alarie, H., Ribot-Ciscar, E., Schneider, C., 2017b. Reliability of lower limb transcranial magnetic stimulation outcomes in the ipsi- and contralesional hemispheres of adults with chronic stroke. Clin. Neurophysiol. 128, 1290–1298. https://doi.org/10.1016/j.clinph.2017.04.021.

Beckerman, H., Roebroeck, M.E., Lankhorst, G.J., Becher, J.G., Bezemer, P.D., Verbeek, A.L.M., 2001. Smallest real difference, a link between reproducibility and responsiveness. Qual. Life Res. 10, 571–578.

Bhandari, A., Radhu, N., Farzan, F., Mulsant, B.H., Rajji, T.K., Daskalakis, Z.J., Blumberger, D.M., 2016. A meta-analysis of the effects of aging on motor cortex neurophysiology assessed by transcranial magnetic stimulation. Clin. Neurophysiol. 127 (8), 2834–2845.

Biabani, M., Farrell, M., Zoghi, M., Egan, G., Jaberzadeh, S., 2018. The minimal number of TMS trials required for the reliable assessment of corticospinal excitability, short interval intracortical inhibition, and intracortical facilitation. Neurosci. Lett. 674, 94–100. https://doi.org/10.1016/j.neulet.2018.03.026.

Brown, K.E., Lohse, K.R., Mayer, I.M.S., Strigaro, G., Desikan, M., Casula, E.P., Meunier, S., Popa, T., Lamy, J.-C., Odish, O., Leavitt, B.R., Durr, A., Roos, R.A.C., Tabrizi, S.J., Rothwell, J.C., Boyd, L.A., Orth, M., 2017. The reliability of commonly used electrophysiology measures. Brain Stimul. 10 (6), 1102–1111.

Bruton, A., Conway, J.H., Holgate, S.T., 2000. Reliability: What is it, and how is it measured? Physiotherapy 86, 94–99. https://doi.org/10.1016/S0031-9406(05)61211-4.

Celebi, O., Temuçin, Ç.M., Elibol, B., Saka, E., 2012. Short latency afferent inhibition in Parkinson's disease patients with dementia. Mov. Disord. 27, 1052–1055. https://doi.org/10.1002/MDS.25040.

Chen, R., Corwell, B., Hallett, M., 1999. Modulation of motor cortex excitability by median nerve and digit stimulation. Exp. Brain Res. 129, 77–86. https://doi.org/10.1007/s002210050938.

Damron, L.A., Dearth, D.J., Hoffman, R.L., Clark, B.C., 2008. Quantification of the corticospinal silent period evoked via transcranial magnetic stimulation. J. Neurosci. Methods 173, 121–128. https://doi.org/10.1016/j.jneumeth.2008.06.001.

Degardin, A., Devos, D., Cassim, F., Bourriez, J.-L., Defebvre, L., Derambure, P., Devanne, H., 2011. Deficit of sensorimotor integration in normal aging. Neurosci. Lett. 498 (3), 208–212.

Goldsworthy, M.R., Hordacre, B., Ridding, M.C., 2016. Minimum number of trials required for within- and between-session reliability of TMS measures of corticospinal excitability. Neuroscience 320, 205–209. https://doi.org/10.1016/j.neuroscience.2016.02.012.

Hopkins, W.G., 2000. Measures of reliability in sports medicine and science. Sports Med. 30, 1–15. https://doi.org/10.2165/00007256-200030010-00001.

Koo, T.K., Li, M.Y., 2016. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J. Chiropr. Med. 15, 155–163. https://doi.org/10.1016/j.jcm.2016.02.012.

Di Lazzaro, V., Oliviero, A., Pilato, F., Saturno, E., Dileone, M., Marra, C., et al., 2004. Motor cortex hyperexcitability to transcranial magnetic stimulation in Alzheimer's disease. J. Neurol. Neurosurg. Psychiatry 75, 555–559. https://doi.org/10.1136/jnnp.2003.018127.

Di Lazzaro, V., Oliviero, A., Saturno, E., Dileone, M., Pilato, F., Nardone, R., Ranieri, F., Musumeci, G., Fiorilla, T., Tonali, P., 2005a. Effects of lorazepam on short latency afferent inhibition and short latency intracortical inhibition in humans. J. Physiol. 564 (2), 661–668.

Di Lazzaro, V., Oliviero, A., Tonali, P.A., Marra, C., Daniele, A., Profice, P., Saturno, E., Pilato, F., Masullo, C., Rothwell, J.C., 2002. Noninvasive in vivo assessment of cholinergic cortical circuits in AD using transcranial magnetic stimulation. Neurology 59 (3), 392–397.

Di Lazzaro, V., Pilato, F., Dileone, M., Profice, P., Ranieri, F., Ricci, V., Bria, P., Tonali, P. A., Ziemann, U., 2007. Segregating two inhibitory circuits in human motor cortex at the level of GABAA receptor subtypes: A TMS study. Clin. Neurophysiol. 118 (10), 2207–2214.

Di Lazzaro, V., Pilato, F., Dileone, M., Tonali, P.A., Ziemann, U., 2005b. Dissociated effects of diazepam and lorazepam on short-latency afferent inhibition. J. Physiol. 569, 315–323. https://doi.org/10.1113/jphysiol.2005.092155.

Liu, H., Au-Yeung, S.S.Y., 2014. Reliability of transcranial magnetic stimulation induced corticomotor excitability measurements for a hand muscle in healthy and chronic stroke subjects. J. Neurol. Sci. 341, 105–109. https://doi.org/10.1016/J.JNS.2014.04.012.

Milani, P., Piu, P., Popa, T., della Volpe, R., Bonifazi, M., Rossi, A., Mazzocchio, R., 2010. Cortisol-induced effects on human cortical excitability. Brain Stimul. 3 (3), 131–139.

Nardone, R., Bergmann, J., Christova, M., Caleri, F., Tezzon, F., Ladurner, G., Trinka, E., Golaszewski, S., 2012. Short latency afferent inhibition differs among the subtypes of mild cognitive impairment. J. Neural Transm. 119 (4), 463–471.

Ngomo, S., Leonard, G., Moffet, H., Mercier, C., 2012. Comparison of transcranial magnetic stimulation measures obtained at rest and under active conditions and their reliability. J. Neurosci. Methods 205, 65–71. https://doi.org/10.1016/j.jneumeth.2011.12.012.

Ni, Z., Charab, S., Gunraj, C., Nelson, A.J., Udupa, K., Yeh, I.-J., Chen, R., 2011. Transcranial Magnetic Stimulation in Different Current Directions Activates Separate Cortical Circuits. J. Neurophysiol. 105 (2), 749–756.

Peter, J., Lahr, J., Minkova, L., Lauer, E., Grothe, M.J., Teipel, S., et al., 2016. Contribution of the Cholinergic System to Verbal Memory Performance in Mild Cognitive Impairment. J. Alzheimer's Dis. 53, 991–1001. https://doi.org/10.3233/JAD-160273.

Portney, L.G., Watkins, M.P., 2009. Foundations of Clinical Research: Applications to Practice. 3rd edition. Pearson/Prentice Hall, Upper Saddle River, NJ.

Rossini, P.M., Burke, D., Chen, R., Cohen, L.G., Daskalakis, Z., Di Iorio, R., Di Lazzaro, V., Ferreri, F., Fitzgerald, P.B., George, M.S., Hallett, M., Lefaucheur, J.P., Langguth, B., Matsumoto, H., Miniussi, C., Nitsche, M.A., Pascual-Leone, A., Paulus, W., Rossi, S., Rothwell, J.C., Siebner, H.R., Ugawa, Y., Walsh, V., Ziemann, U., 2015. Non-invasive electrical and magnetic stimulation of the brain, spinal cord, roots and peripheral nerves: Basic principles and procedures for routine clinical and research application: An updated report from an I.F.C.N. Committee. Clin. Neurophysiol. 126 (6), 1071–1107.

Sailer, A., Molnar, G.F., Paradiso, G., Gunraj, C.A., Lang, A.E., Chen, R., 2003. Short and long latency afferent inhibition in Parkinson's disease. Brain 126, 1883–1894. https://doi.org/10.1093/brain/awg183.

Samusyte, G., Bostock, H., Rothwell, J., Koltzenburg, M., 2018. Short-interval intracortical inhibition: Comparison between conventional and threshold-tracking techniques. Brain Stimul. 11, 806–817. https://doi.org/10.1016/j.brs.2018.03.002.

Sankarasubramanian, V., Roelle, S.M., Bonnett, C.E., Janini, D., Varnerin, N.M., Cunningham, D.A., Sharma, J.S., Potter-Baker, K.A., Wang, X., Yue, G.H., Plow, E.B., 2015. Reproducibility of transcranial magnetic stimulation metrics in the study of proximal upper limb muscles. J. Electromyogr. Kinesiol. 25 (5), 754–764.

Schambra, H.M., Ogden, R.T., Martínez-Hernández, I.E., Lin, X., Chang, Y.B., Rahman, A., Edwards, D.J., Krakauer, J.W., 2015. The reliability of repeated TMS measures in older adults and in patients with subacute and chronic stroke. Front. Cell. Neurosci. 9. https://doi.org/10.3389/fncel.2015.00335.

Šerbetar, I., 2015. Establishing Some Measures of Absolute and Relative Reliability of a Motor Tests. Croat. J. Educ. Croat. J. Educ., 17 https://doi.org/10.15516/cje.v17i0.1484.

Toepp, S.L., Turco, C.V., Rehsi, R.S., Nelson, A.J., Schwenkreis, P., 2021. The distribution and reliability of TMS-evoked short- and long-latency afferent interactions. PLoS One 16 (12), e0260663. https://doi.org/10.1371/JOURNAL.PONE.0260663.

Tokimura, H., Lazzaro, V., Tokimura, Y., Oliviero, A., Profice, P., Insola, A., Mazzone, P., Tonali, P., Rothwell, J.C., 2000. Short latency inhibition of human hand motor cortex by somatosensory input from the hand. J. Physiol. 523 (2), 503–513.

Turco, C.V., El-Sayes, J., Fassett, H.J., Chen, R., Nelson, A.J., 2017. Modulation of long-latency afferent inhibition by the amplitude of sensory afferent volley. J. Neurophysiol. 118, 610–618. https://doi.org/10.1152/jn.00118.2017.

Turco, C.V., El-Sayes, J., Locke, M.B., Chen, R., Baker, S., Nelson, A.J., 2018. Effects of lorazepam and baclofen on short- and long-latency afferent inhibition. J. Physiol. 596, 5267–5280. https://doi.org/10.1113/JP276710.

Turco, C.V., Pesevski, A., McNicholas, P.D., Beaulieu, L.-D., Nelson, A.J., 2019. Reliability of transcranial magnetic stimulation measures of afferent inhibition. Brain Res. 1723. https://doi.org/10.1016/j.brainres.2019.146394.

Weir, J., 2005. Quantifying test-retest reliability using the intraclass correlation coefficient. J. Strength Cond. Res. 19, 231–240.

Yarnall, A.J., Ho, B.S.W., Eshun, E., David, R., Rochester, L., Burn, D.J., Baker, M.R., 2016. Short latency afferent inhibition: Effects of ageing. Clin. Neurophysiol. 127 (6), 2410–2413.

Yarnall, A.J., Rochester, L., Baker, M.R., David, R., Khoo, T.K., Duncan, G.W., Galna, B., Burn, D.J., 2013. Short latency afferent inhibition: A biomarker for mild cognitive impairment in Parkinson's disease? Mov. Disord. 28 (9), 1285–1288.

Young-Bernier, M., Kamil, Y., Tremblay, F., Davidson, P.S.R., 2012. Associations between a neurophysiological marker of central cholinergic activity and cognitive functions in young and older adults. Behav. Brain Funct. 8, 1–9. https://doi.org/10.1186/1744-9081-8-17.