Journal of Electromyography and Kinesiology 24 (2014) 593–600

Contents lists available at ScienceDirect

Journal of Electromyography and Kinesiology

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

Internal consistency and test–retest reliability of an instrumented functional reaching task using wireless electromyographic sensors

Rini Varghese, Christina W.Y. Hui-Chan, Edward Wang, Tanvi Bhatt *

Motor Behavior and Balance Rehabilitation Laboratory, Department of Physical Therapy, University of Illinois at Chicago, Chicago, USA

ARTICLE INFO

Article history:
Received 21 November 2013
Received in revised form 7 May 2014
Accepted 30 May 2014

Keywords:
Stand-arm reaching
Flexion
Abduction
Electromyography
Acceleration
Reliability

ABSTRACT

The purpose of this study was to establish the internal consistency and test–retest reliability of the electromyographic and accelerometric data sampled from the prime movers of the dominant arm during an antigravity, within-arm's length stand-reaching task without trunk restraint. Ten healthy young adults participated in two experimental sessions, approximately 7–10 days apart. During each session, subjects performed 15 trials of both a flexion- and an abduction-reaching task. Surface EMG and acceleration using wireless sensors were sampled from the anterior and middle deltoid. Reliability was established using Cronbach's alpha, intraclass correlation coefficients (ICC 2, k) and standard error of measurements (SEM) for electromyographic reaction time, burst duration and normalized amplitude along with peak acceleration. Results indicated high degrees of inter-trial and test–retest reliability for flexion (Cronbach's α range = 0.92–0.99; ICC range = 0.82–0.92) as well as abduction (Cronbach's α range = 0.94–0.99; ICC range = 0.81–0.94) reaching. The SEM associated with response variables for flexion and abduction ranged from 1.55–3.26% and 3.33–3.95% of means, respectively. Findings from this study revealed that electromyographic and accelerometric data collected from prime movers of the arm during the relatively functional stand-reaching task were highly reproducible. Given its high reliability and portability, the proposed test could have applications in clinical and laboratory settings to quantify upper limb function.

1. Introduction

Assessing upper limb function, both in the laboratory and clinic, can be challenging because of the complexities arising from its involvement in discrete tasks, multiple degrees of freedom and requirement of action preparation in the form of information processing and postural adjustments. Among the functions of the upper limb, reaching activities are the earliest to develop (Shumway-Cook and Woollacott, 1995), comprise a majority of its uses and are often followed by grasp and/or manipulation. Current methods to assess upper extremity dysfunction lack sensitive quantification, clinical translatability, functionality and/or repeatability. This limits the detection of early signs of improvement, the progression of disease, the level of structural impairment and how these influence the individual with respect to function and participation. Evaluating selective upper extremity tasks using appropriately designed, adequately controlled and reliable outcome measures is a defensible way of quantifying such changes (Long and Scott, 1994).

The advantages of clinical scales include their universal availability, cost-effectiveness and portability. Unfortunately, these measures can be highly subjective, lack sensitive quantification (necessary to measure the small changes that result from training, e.g. convalescing individuals) or have floor or ceiling effects, which limit their use in individuals who might be at the cognitive or autonomic stages (e.g. high-performing athletes) of performance, respectively, of a certain motor skill. Additionally, there are not as many tools available to assess general motor function of upper limbs in otherwise normal older adults, who are known to manifest rather silent symptoms of aging that contribute to poor performance in day-to-day activities.

With advances in portable, wireless technology, some laboratory-based measures can be conveniently used as sensitive clinical tools for the assessment of upper limb function. In the literature, many such measures have been developed and used for research purposes (Bertucco et al., 2013; Piovesan et al., 2013; Sciutti et al., 2012) yet have not been translated to the clinical setting. This could be possibly due to the infeasibility of having expensive, highly sophisticated measurement systems in the clinic, poor functionality of the testing design, lack of sufficient information to reproduce the given protocol, or unavailability of reliability estimates for the outcome variables. Especially, in the context of upper

* Corresponding author. Address: 514-E, College of Applied Health Sciences, 1919W. Taylor St. Chicago, IL 60612, USA. Tel.: +1 312 355 4443.
E-mail address: tbhatt6@uic.edu (T. Bhatt).

http://dx.doi.org/10.1016/j.jelekin.2014.05.011
1050-6411/

limb reaching measures, laboratory-developed reaching tasks lack one or more functional aspects such as the plane of reaching, starting position from which it is performed and/or use of physical trunk restraint. This adds to the inability of lab-designed measures to be replicated in clinical settings.

The functionality of the design affects the task and environmental constraints of the test and weigh largely on individual performance. With respect to reaching, the plane of and starting position from which it is performed address two different aspects of motor control underlying the skill. Reaching tasks have been predominantly measured in the horizontal plane (Almeida et al., 2006; Hollerbach and Flash, 1982; Piovesan et al., 2013; Trent and Ahmed, 2013). It is well known that movements, including reaching, in the anti-gravity plane are more resistive, necessitating greater muscle contraction in order to generate sufficient torque to overcome limb inertia, and thus more challenging. Nevertheless, they are more functional than reaching in a horizontal plane and are required by almost all activities of daily living. Despite this fact, only a few studies use an anti-gravity reaching measure (Papaxanthis et al., 1998a; Sciutti et al., 2012; Tyler and Karst, 2004).

It must be noted that most of these gravity-eliminated, horizontal plane reaching, and some other gravity-resisted reaching, tasks were performed from a seated position, sizably obviating the need for postural action preparation (Atkeson and Hollerbach, 1985; Chevalot and Xuguang, 2004; Klein Breteler et al., 1998; Nishikawa et al., 1999; Sciutti et al., 2012; Zhang and Chaffin, 2000). Performing a reaching task from a standing position is far more challenging than it is from a seated one (Christina et al., 1982; Christina and Rose, 1985). Most of the previously mentioned studies limited their scope to understanding simple or choice reaction times of the arm prime-mover without introducing any form of task complexities, and hence used a rather balanced seated position. On the other hand, some studies in the literature have reported the use of a stand-reaching task (Bertucco et al., 2013; Huang, 2009; Tyler and Karst, 2004), however the focus of these studies was limited to examining postural responses and did not consider kinetics or kinematics of the prime mover of the arm. Conversely, the few studies that explored upper limb reaching during standing (Paizis et al., 2008; Papaxanthis et al., 1998b; Tyler and Karst, 2004) did not report test–retest reliability of the lab-developed measure, impeding its use in other laboratory or clinical settings.

Keeping these functional traits in mind, we aimed to design a novel stand-reaching task that accounts and controls for some complexities of arm reaching, thus making it a sensitive yet functionally robust measurement tool. In order to ensure its utility as an outcome measure, it is imperative to know if such a task is feasible and repeatable over time under controlled circumstances. The purpose of this study, therefore, was to estimate the internal consistency and test–retest reliability of an antigravity, within-arm's length stand-reaching task of the dominant arm without trunk restraint.

2. Methods

2.1. Participants

Ten healthy, young and right-handed adults (8 females, 2 males; age = 23.6 ± 3.75 years; height = 162.6 ± 8.47 cm; weight = 127.1 ± 26.59 lb; arm length = 28.1 ± 1.63 in.) were recruited from the student pool of the University of Illinois at Chicago via flyers and volunteered to participate in the study. All participants visited the lab for testing two times within a span of 7–10 days. The Institutional Review Board of the University of Illinois at Chicago approved the study protocol. All subjects signed an informed consent form before participating in the study.

2.2. Design

The set-up primarily consisted of a custom-made arm reaching apparatus (Fig. 1) that comprised of a larger load-bearing clamp that was fixed to a stationary pole in the lab with the long-shaft of the clamp perpendicular to the length of the pole. This was adjusted to the shoulder height of the subjects. Another movable, smaller clamp was used to attach a 36-in. wooden ruler perpendicular to the shaft of the large clamp. This was adjusted to 90% of the subject's arm length. A 4-in. (diameter) circular foam with a distinctly marked smaller central target of 1-in. (diameter) was attached to one end of the wooden ruler. A passive marker, in line with the target, was taped to the top-end of the screw of the small clamp at a fixed distance of 3.5 in. to serve for gaze fixation. Electrical activity of the muscle was measured using Delsys® Trigno™ wireless electromyography sensors. These sensors also have tri-axial accelerometers embedded in them that sampled the rate of change of velocity.

2.3. Protocol

Subjects stood with shoulder-width base of support on a paper foot-mat with their arms at their side. Feet were marked on the paper mat in order to maintain a constant base of support throughout the period of testing. The target was set by adjusting the ruler to 90% of the subject's maximum arm-length-reach (defined as the distance between the acromion to the tip of the middle finger). Subjects were instructed before and intermittently throughout the period of testing to keep their back supported against the wall. They received two computer-generated auditory cues. The first cue (preparatory) – *"Get Ready"* – was given at 2 s, at which subjects visually focused their attention at the passive marker. The final cue – *"Go"* – was given at 4 s, at which subjects reached out and touched the target "as quickly and as accurately" as possible and returned to the starting position. On reporting to have missed the target during a certain trial, that trial was repeated to ensure achievement of the task each time. Subjects also received a short break after every five trials so as to avoid fatigue. Fifteen trials each of verbally cued forward reaching (through shoulder flexion) and sideways reaching (through shoulder abduction) were performed following three familiarization trials. Instructions to be "quick and accurate" were given from time-to-time to ensure continued adherence to the protocol. We tested the dominant arm due to its reported advantage for error correction and effective and consistent movement planning (Bagesteiro and Sainburg, 2002; R. Sainburg and Kalakanis, 2000).

2.4. Data recording and analysis

Delsys® Trigno™ wireless sensors were used to record surface electromyographic activity of the anterior (prime mover for shoulder flexion) and middle (prime mover for shoulder abduction) deltoid muscle of the dominant arm. The sensors were affixed to the skin surface using hypodermal tape, in line with the muscle over its belly, as recommended by Cram et al., 1998. Each sensor has four 5×1 mm contact surfaces made of 99.9% pure silver. Electromyographic signals were sampled at 2000 Hz, hardware band-pass filtered over a bandwidth of 20–450 Hz, using a common mode rejection ratio of >80 db. Tri-axial accelerometers embedded in these sensors rendered signals sampled at 148.1 Hz over a bandwidth of 50 Hz and amplitude range of ±1.5 g. To smooth the EMG data, signals were digitally high-pass filtered using a fourth order zero-lag Butterworth filter (MathWorks, Inc., MATLAB) with

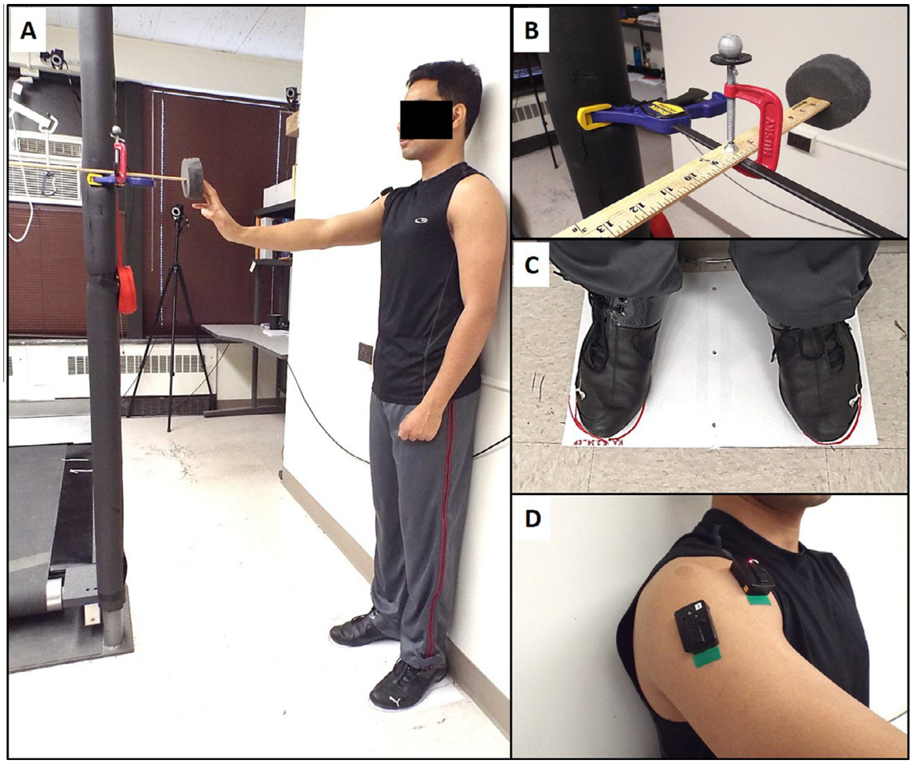


Fig. 1. (A) Experimental set-up showing custom-made apparatus, (B) including a ruler and eye-fixator marker held by clamp complex, attached to the stationary pole and adjustable in height and length. Subject stood with (C) shoulder width distance between feet, marked on a paper foot mat, reaching out to the target set at 90% of arm length. Trunk movement was controlled by instructing the subject to keep the shoulder blades in contact with the wall at all times. (D) EMG sensors were affixed to the anterior and middle deltoid muscles of the dominant arm.

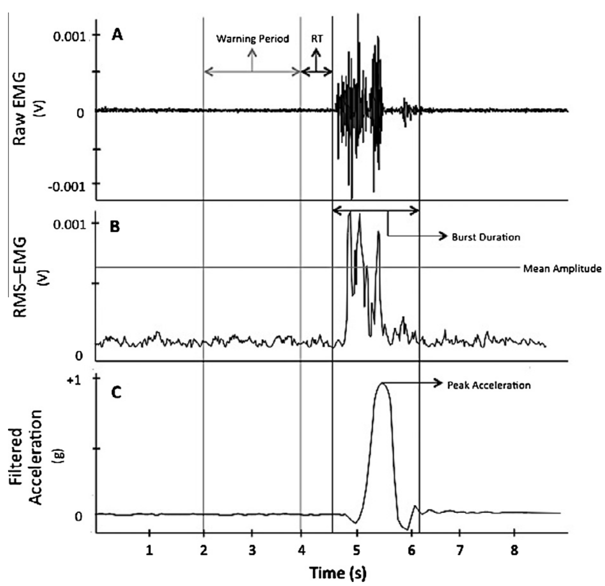


Fig. 2. (A) Raw and (B) processed (filtered and rectified) EMG and (C) acceleration, sampled from the anterior deltoid during flexion-reaching, depicting variables of interest. Reaction time (RT) is the times between final cue and EMG onset. Burst duration is the time between start and end of EMG signal. Normalized EMG is given as the root-mean-squared EMG divided by mean amplitude. Peak acceleration is the maximum amplitude of the signal along the X-axis.

a cut-off frequency of 20 Hz, full wave rectified, and then low-pass filtered with a cut-off frequency of 50 Hz. The acceleration signals were smoothed using a fourth order low-pass Butterworth filter with a cut-off frequency of 80 Hz.

Analysis of EMG signals was done to extract variables of interest using a customized MATLAB code (Mathworks, Inc., MATLAB). These variables included reaction time, burst duration, normalized EMG amplitude and peak acceleration for each of the flexion and abduction reaching trials. Fig. 2 shows raw and processed EMG and filtered acceleration trace depicting these variables of interest. ***Reaction time (RT)*** was defined as the time elapsed between the final cue, "Go" (at 4s), and the onset of EMG signal (calculated as ±1SD from baseline). ***Burst duration*** was defined as the time elapsed between the start and end of the EMG signal (calculated as the return of the signal to baseline). By convention, root-mean-squared (RMS) value of the EMG signal is the best representation of signal power. To compute ***normalized EMG***, we calculated RMS amplitudes for each of the 15 trials and used trial-wise mean amplitudes as the normalization value, thus represented as an ensemble average for each trial. We used this mean amplitude normalization approach instead of maximum voluntary isometric contraction (MVIC) to reduce inter-subject variability (Yang and Winter, 1984). ***Peak acceleration*** was defined as the maximum amplitude of acceleration signal along the X-axis of the sensor's coordinate system. We used the X-axis because the sensor was placed relatively closer to the joint rather than the end-point of the shaft, and hence movements occurred along the sagittal axis of the sensor.

2.5. Statistical analysis

All statistical analyses were performed using SPSS (version 21) for Mac. Descriptive statistics (mean ± 1SD) are presented in tables and text. We used paired t-tests to compare means of each of the response variables on the two sets of data (day-1 and day-2). Repeated-measures ANOVA was used to calculate variance among the 15 trials per session. Inter-trial reliability co-efficient was represented by Cronbach's Alpha for internal consistency. Intra-class correlation coefficients (ICC) and standard error of measurement (SEM) were used to capture day-to-day reliability of the data. Bland–Altman plots were constructed to compare each of the four variables for flexion and abduction reaching on the two days. A range of agreement was defined as mean bias ±1.96 SD. A one-sample *t*-test was used to analyze if the bias (difference) between the means of day-1 and day-2 were significantly different from zero. Additionally, to quantify the strength of the relationship between the means of day-1 and day-2, Pearson's product moment correlation (R) and co-efficient of determination (R^2) were also computed.

A two-factor random effects model, type consistency of ICC for average measures (Shrout and Fleiss, 1979) was used. This method was chosen in accordance with Case 2 model for k trials proposed by Shrout and Fleiss (1979). ICC ranges can be classified as good reproducibility: 0.80–1.0, fair reproducibility: 0.60–0.79 and poor reproducibility: <0.60 (Sleivert and Wenger, 1994). An ICC value >0.8 has been reported to be acceptable for clinical work (Currier, 1990).

The SEM is a representation of the error of measurement associated with the observed score for a given variable. It must be examined relative to the mean and standard deviation. SEM can also be determined from the square root of within-subject variance. Interpretation of the values of Cronbach's Alpha is similar to other correlation co-efficient in which scores range from 0 to 1, with score closer to 1 suggesting higher degree of consistency. A general agreement in psychometrics is that a score >0.70 implies high internal consistency (Lj, 1951).

3. Results

Repeated measures ANOVA revealed no significant differences among the fifteen trials for each of the four variables of flexion – RT ($F = 0.629$, $p = 0.840$), burst duration ($F = 0.967$, $p = 0.487$), normalized EMG amplitude ($F = 1.326$, $p = 0.191$) and peak acceleration ($F = 1.315$, $p = 0.198$) or abduction – RT ($F = 1.419$, $p = 0.144$), burst duration ($F = 0.599$, $p = 0.865$), normalized EMG amplitude ($F = 0.604$, $p = 0.861$) and peak acceleration ($F = 1.464$, $p = 0.125$). Paired *t*-tests showed no significant difference between day one and day two for any of the four variables for flexion – RT ($t = 0.698$, $p = 0.503$), burst duration ($t = 0.154$, $p = 0.881$), normalized EMG amplitude ($t = 0.769$, $p = 0.462$) and peak acceleration ($t = -0.795$, $p = 0.447$) or abduction – RT ($t = 1.411$, $p = 0.192$), burst duration ($t = 0.431$, $p = 0.677$), normalized EMG amplitude ($t = 1.153$, $p = 0.279$) and peak acceleration ($t = -1.062$, $p = 0.318$). Table 1 presents mean and standard deviation on day-1 and day-2, R as given by Pearson's product moment correlation and the co-efficient of determination R^2. Fig. 3 shows an overlaid means-plot demonstrating subject-wise means for all flexion and abduction variables on day-1 and day-2. A relatively high correlation was seen between day-1 and day-2 values for all variables. Fig. 5 shows a sample regression plot for anterior deltoid RT on day-1 and day-2.

3.1. Internal consistency

Table 2 presents Cronbach's alpha along with its 95% CI. A large degree of internal consistency was observed among trials. Cronbach's alpha ranged from 0.92–0.99 and 0.94–0.99 for response variables of flexion and abduction, respectively.

3.2. Relative reliability

Table 2 presents ICC with 95% CI for each of the variables. A high degree of reliability (observed ICC >0.80) was seen for all variables (EMG and acceleration) of both flexion and abduction arm reaching. The ICCs ranged from 0.823–0.916 for flexion (anterior deltoid) and 0.813–0.937 for abduction (middle deltoid) reaching of the arm. Fig. 4 shows Bland–Altman plots. There was no significant difference from 0 in the mean difference between the two days for flexion – RT (5.22 ± 23.64; $p = 0.503$), burst duration (2.49 ± 50.94; $p = 0.881$), normalized EMG amplitude (0.017 ± 0.068; $p = 0.462$) and peak acceleration (−0.016 ± 0.065; $p = 0.447$), or abduction – RT (16.17 ± 36.23; $p = 0.192$), burst duration (7.94 ± 58.21; $p = 0.677$), normalized EMG amplitude (0.022 ± 0.060; $p = 0.279$) and peak acceleration (−0.020 ± 0.062; $p = 0.316$).

3.3. Absolute reliability

Table 2 also shows SEMs for all variables. The SEM associated with response variables for flexion and abduction ranged from 1.55–3.26% and 3.33–3.95% of means, respectively. Absolute values for SEM for anterior deltoid EMG and accelerometer variables were relatively lower than that of middle deltoid, except for peak acceleration.

4. Discussion

Overall, results from this study indicate that electromyographic and accelerometric data were reproducible and showed a high level of internal consistency and test–retest reliability, with a relatively acceptable standard error of measurement. Consequently, it is recommended that the estimates presented herein be used as a guide for further exploration of this tool in other diverse settings.

4.1. Temporal variables

Current findings show that our set-up and protocol provide reliable estimates of electromyographic reaction time and burst

Table 1
Mean and standard deviation on both Day 1 and Day 2; Pearson's R and co-efficient of determination R^2 for all variables of both flexion and abduction.

Variable	Day 1		Day 2		R	R^2
	Mean	SD	Mean	SD		
Flexion						
RT (ms)	487.87	41.58	482.65	43.13	0.845	0.714
Burst duration (ms)	837.79	73.84	835.31	59.25	0.728	0.530
Normalized EMG amplitude (units)	1.36	0.10	1.34	0.08	0.773	0.597
Peak acceleration (g)	0.85	0.09	0.87	0.07	0.713	0.509
Abduction						
RT (ms)	506.21	54.72	490.04	43.64	0.751	0.564
Burst duration (ms)	862.46	116.90	854.52	122.57	0.883	0.779
Normalized EMG amplitude (units)	1.34	0.07	1.32	0.08	0.722	0.522
Peak acceleration (g)	0.89	0.07	0.91	0.07	0.685	0.469

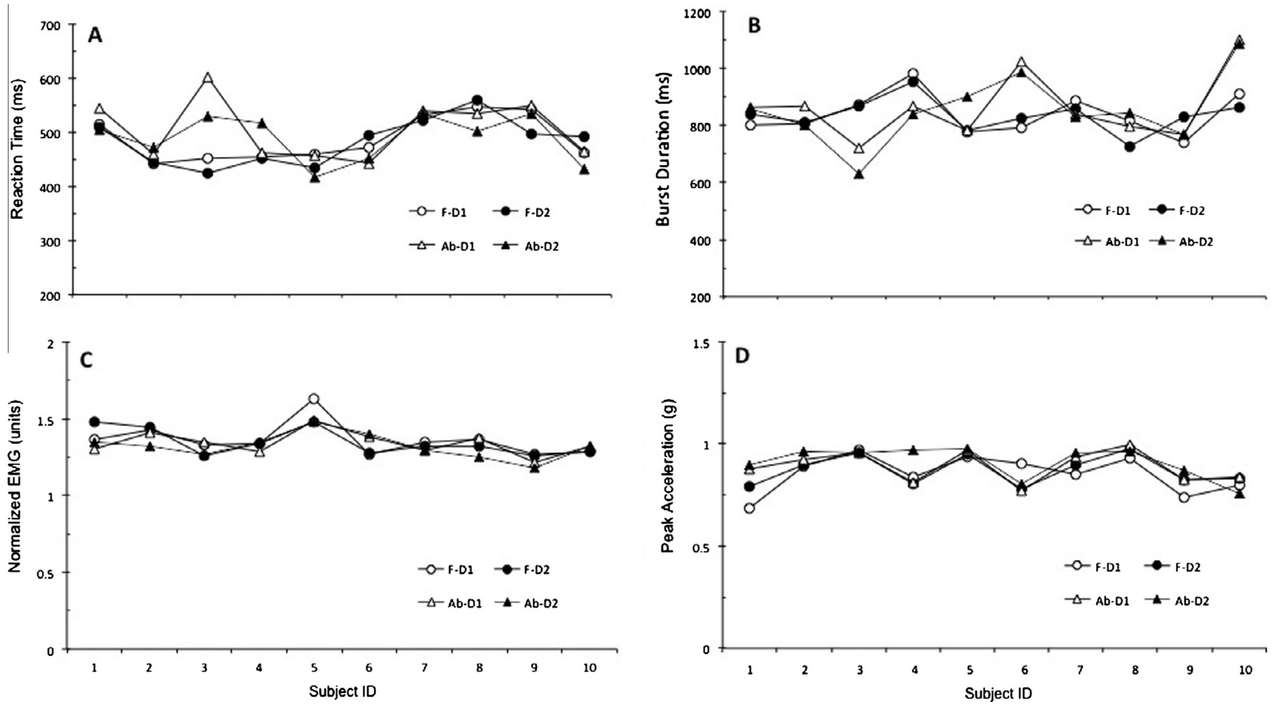


Fig. 3. An overlaid means-plot showing subject-wise means of temporal and amplitude variables for flexion and abduction on Day-1 and Day-2. (A) Reaction time (in milliseconds); (B) Burst duration (in milliseconds); (C) Normalized EMG (RMS/mean amplitude, in units); (D) Peak acceleration (in *g* or acceleration due to gravity). F-D1 – Flexion, Day-1; F-D2 – Flexion, Day-2; Ab-D1 – Abduction, Day-1; Ab-D2 – Abduction, Day-2. *Y*-axis represents subject IDs 1–10.

Table 2
Cronbach's alpha, ICC and SEM values for all variables of both flexion and abduction.

Variable	Cronbach's α	95% CI	ICC	95% CI	SEM
Flexion					
RT	0.920	0.858–0.963	0.916	0.660–0.979	12.27
Burst duration	0.928	0.871–0.967	0.831	0.318–0.958	27.35
Normalized EMG amplitude	0.947	0.906–0.976	0.860	0.437–0.965	0.021
Peak acceleration	0.991	0.983–0.996	0.823	0.286–0.956	0.034
Abduction					
RT	0.940	0.894–0.972	0.845	0.377–0.962	19.36
Burst duration	0.962	0.933–0.983	0.937	0.747–0.984	30.05
Normalized EMG amplitude	0.946	0.905–0.975	0.837	0.346–0.960	0.030
Peak acceleration	0.991	0.984–0.996	0.813	0.248–0.954	0.030

duration of the anterior and middle deltoid during flexion and abduction reaching, respectively. This finding is of particular importance because despite the effect of task complexity on the temporal variables of reaching, our design demonstrates substantial reproducibility in these outcomes. Simple reaction time, which ranges between 200 and 250 ms, has been traditionally measured and reported with the individual sitting and/or reaching performed in the horizontal plane. However, these tasks do not adequately represent real-world arm use. On the other hand, of the studies that used a seated reaching task in the anti-gravity plane, only one has reported the day-to-day reliability of the reaction time of arm prime mover (Kwok et al., 2010), which was approximately 280–320 ms for young adults, with an ICC of 0.70 (0.55–0.80). However, even seated tasks do not fully account for routinely performed reaching activities. Subsequently, as in the task discussed in this study, arm reaching from the standing position can be particularly challenging due to the additional postural demands of the task and can result in prolonged prime-mover latencies. A disproportional delay in the reaction time of arm prime-mover, with respect to the time required to sequentially activate postural muscles has been reported previously (Inglin and Woollacott, 1988). Whereas postural activation latencies form the primary focus of studies that observe the effect of stand arm reaching on anticipatory postural adjustments, reaction time of the prime mover of the arm has seldom been reported independently. Our findings provide reliable estimates of reaction time of the prime mover of the arm during a stand-reaching task, with a relative error of measurement as small as 2.53% and 3.88% of means for RT, and 3.26% and 3.50% of means for burst duration of flexion and abduction reaching, respectively. Internal consistency of temporal measures also indicates a steady, unchanging testing protocol, which accounts for complexities arising from multiple degrees of freedom of the arm and trial-induced fatigue.

4.2. Amplitude variables

Our findings also suggest reliability of normalized EMG amplitude. This can be attributed to a high degree of congruence in electrode position on day-1 and day-2, similar task and environmental conditions of testing and use of mean amplitude normalization method. Reliable method for EMG normalization is a topic of debate even to date. Yang and Winter (1984) were among the first

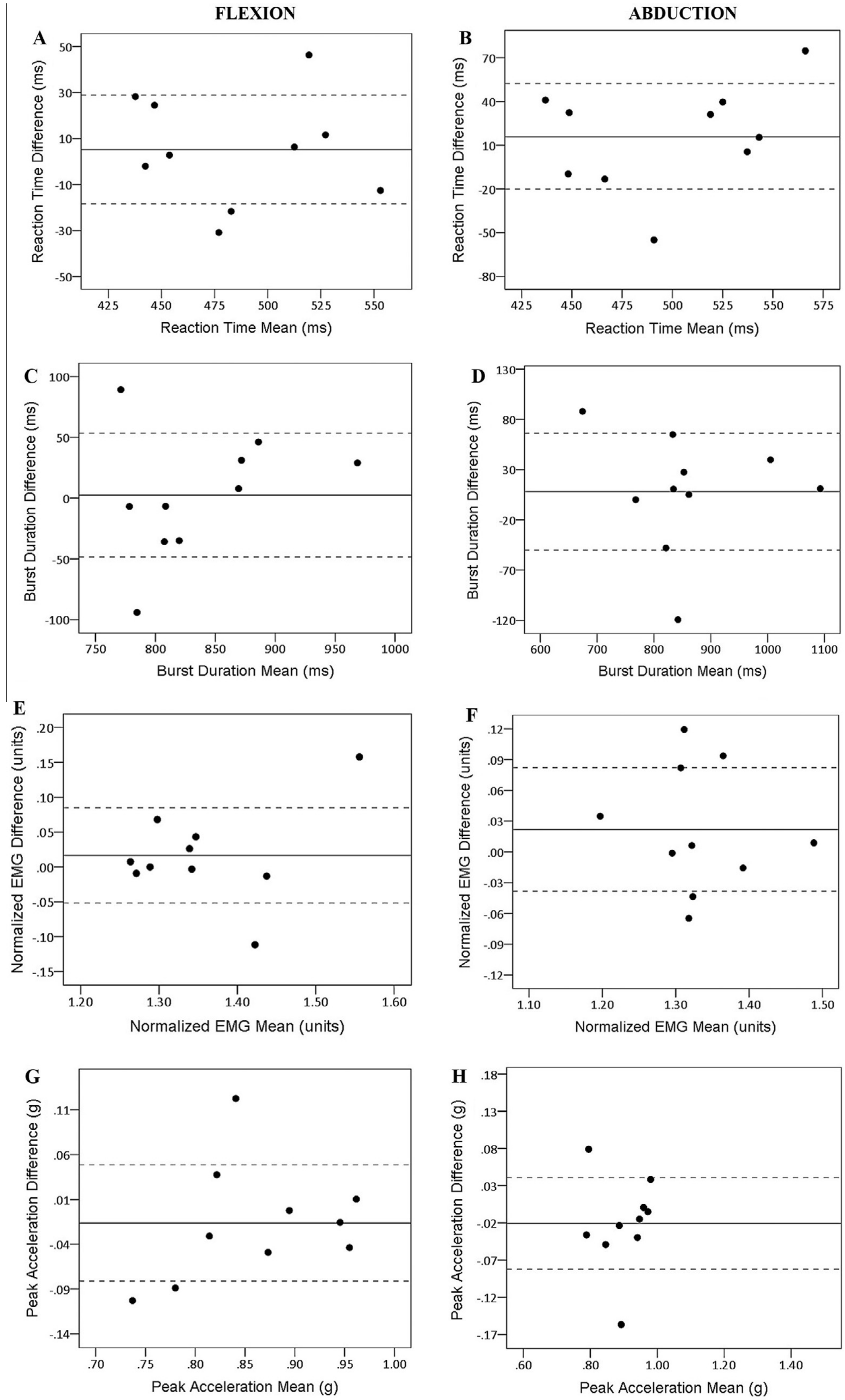


Fig. 4. Bland–Altman plots displaying limits of agreement; mean (solid line) ±1.96 SD (dotted line) of the difference between day-1 and day-2 values for each of the four variables: (A and B) reaction time; (C and D) burst duration; (E and F) normalized EMG amplitude, and (G and H) peak acceleration, for flexion (left column) and abduction reaching (right column) tasks.

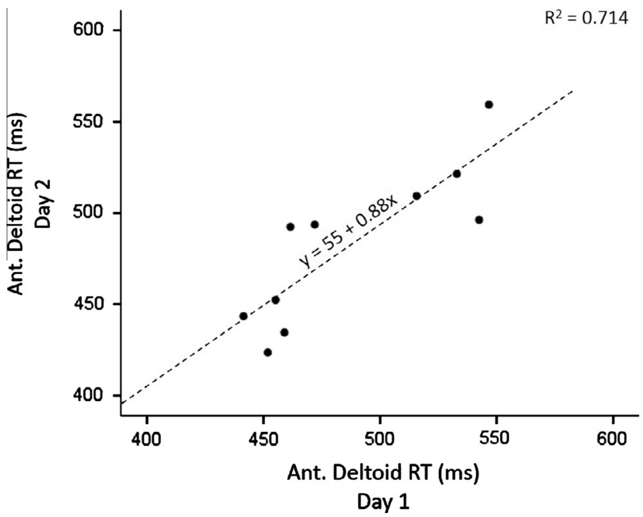


Fig. 5. A sample regression plot showing subject-wise means of reaction time (RT), for all subjects on Day-1 (*X*-axis) and Day-2 (*Y*-axis). Best-fit line along with associated slope equation is also shown. Coefficient of determination (R^2) shows 71.4% strength of correlation between means of Day-1 and Day-2.

ones to provide a guide to the use of amplitude normalization methods. We use this method in order to account for variability arising due to the dynamic nature of the task and to reduce the effects of inter-subject variability and fatigue. Other studies have reported reliability in EMG measures of lower extremities using the MVIC normalization technique, however, this method has been reported to be less reliable than mean or peak amplitude normalization for shoulder EMG (Morris et al., 1998).

High reliability was also observed in peak accelerations of the arm during the two reaching tasks. This could be attributed to the fact that subjects were asked to reach with their dominant arm, which, as previously mentioned, tends to manifest an error-minimized movement plan and a relatively robust control of limb dynamics. Moreover, the provision of instructional reminders to perform reaching "quickly and accurately", given from time to time, might have reinforced target reaching using almost similar limb strategies, therefore leading to less variability in the acceleration profile, especially peak acceleration. However, since no other kinematic outcomes were measured, a conclusive comment about segmental limb dynamics cannot be made. Finally, a very high degree of internal consistency was also noted in EMG amplitude and peak acceleration signals implying a strategic and consistently similar use of available degrees of freedom of deltoid muscles in a given session of the specific task. Our proposed protocol therefore seems to be a reliable tool to assess neuromuscular activation of deltoids during shoulder flexion and abduction with a fairly small relative error in measurement ranging from 1.55% and 2.25% for EMG amplitude, and 3.95% and 3.33% for peak accelerations of anterior and middle deltoid, respectively.

4.3. Reliability indices

Different schools of thought exist on the use of reliability statistics. In this study, we primarily chose to use intra-class correlation coefficients (ICC) and standard error of measurement (SEM) to capture reliability of the data. Whereas ICC is a measure of relative reliability (i.e. with respect to obtained scores), SEM provides a generalizable error factor that can be considered as the boundaries of the true score. Therefore, a narrower boundary (i.e. smaller SEM) suggests closer approximation of observed score to the true score for a given measure. Other statistical operations reported in the literature include Bland Altman technique (Martin Bland and Altman, 1986), which does not correct for inter-trial variability, however may provide useful information about any systematic bias in day-to-day measurement. ICC and SEM are particularly useful because they remain immune to inter-subject variability, which tend to be magnified by a small sample.

Furthermore, we used Cronbach's Alpha to determine the inter-trial reliability. As previously stated, there is a general agreement in psychometrics that a score >0.70 implies high internal consistency (Lj, 1951). However, this must be viewed in the context of electromyographic measurements (i.e. each of the 15 trials measured the construct in the same way). Since we do not expect a large difference among trials, due to steady electrode position and other task-related factors, the difference arising must be from individual variation; hence, it is better to raise the standard for this reliability index for EMG measurements to be greater than 0.70, as is otherwise generally used in qualitative statistics. Finally, we also used Pearson's R and the co-efficient of determination R^2 to support our observation of a close relationship in the scores between two test days, however, since Pearson's correlation assumes linearity, our results might be an underestimation of the true correlation.

Measuring upper extremity function can be challenging, and sensitive yet reliable tools are required for assessment. In the current study, we attempted to develop such a tool with a methodology that is novel and differs from existing upper extremity assessments in several ways. Firstly, due to the anti-gravity, within-arm's length, stand-reaching nature of the protocol, this measure is more functional than seated, horizontal-plane reaching and mimics daily-living activities more closely than traditional protocols, thus allowing better estimation of key features of intervention such as transfer of learning and generalizability in respective contexts. Moreover, studies that have used similar stand-reaching paradigms (Paizis et al., 2008; Papaxanthis et al., 1998b; Tyler and Karst, 2004) have not reported reliability estimates, limiting their use in other settings. Secondly, in addition to the relatively simple design and equipment, the reaching task was instrumented using wireless sensors that provide further portability; to the best of our knowledge this approach has not been reported in the past, at least in the area of upper limb movement control. Lastly, we report both amplitude and temporal variables for EMG and acceleration signals, which have also not been reported together for such a reaching task. It is a combination of these factors that make this methodology unique.

5. Conclusion

The reaching task developed in the current investigation is both internally consistent and repeatable for EMG and acceleration measures of prime movers. Given the high level of internal and external reliability of our measure in healthy young adults, it is recommended that future studies explore its feasibility and reliability in other populations ranging from neurological disorders to healthy older adults and the physically elite such as athletes. Reliability estimates and normative values from this study may be a useful guide for further evaluation of this tool in these fields and populations. Furthermore, due to the portability of this instrumented task, it could be easily translated to clinical, field and community-settings as well.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgement

Abdul Karim for his help with data analyses.

References

Almeida GL, Freitas SMSF, Marconi NF. Coupling between muscle activities and muscle torques during horizontal-planar arm movements with direction reversal. J Electromyogr Kinesiol 2006;16(3):303–11.

Atkeson CG, Hollerbach JM. Kinematic features of unrestrained vertical arm movements. J Neurosci 1985;5(9):2318–30.

Bagesteiro LB, Sainburg RL. Handedness: Dominant arm advantages in control of limb dynamics. J Neurophysiol 2002;88(5):2408–21.

Bertucco M, Cesari P, Latash ML. Fitts' law in early postural adjustments. Neuroscience 2013;231:61–9.

Chevalot N, Xuguang W. An experimental investigation of the discomfort of arm reaching movements in a seated position. SAE Trans 2004;113(1):98–103.

Christina RW, Rose DJ. Premotor and motor reaction time as a function of response complexity. Res Q Exerc Sport 1985;56(4):306–15.

Christina RW, Fischman MG, Vercruyssen MJ, Anson JG. Simple reaction time as a function of response complexity: memory drum theory revisited. J Mot Behav 1982;14(4):301–21.

Cram JR, Kasman GS, Holtz J. Introduction to surface electromyography. Aspen Publishers; 1998.

Currier DP. Elements of research in physical therapy. Williams & Wilkins Baltimore; 1990.

Hollerbach JM, Flash T. Dynamic interactions between limb segments during planar arm movement. Biol Cybern 1982;44(1):67–77.

Huang M. Age differences in the control of posture and movement during standing reach. Dissertation abstracts international: section B: the sciences and engineering. Retrieved from <http://search.proquest.com/docview/756306789?accountid=14552>. (756306789; 2010-99140-141); 2009.

Inglin B, Woollacott M. Age-related changes in anticipatory postural adjustments associated with arm movements. J Gerontol 1988;43(4):M105–13.

Klein Breteler MD, Meulenbroek RGJ, Gielen SCAM. Geometric features of workspace and joint-space paths of 3D reaching movements. Acta Psychol 1998;100(1–2):37–53.

Kwok JC, Hui-Chan CW, Tsang WW. Effects of aging and tai chi on finger-pointing toward stationary and moving visual targets. Arch Phys Med Rehabil 2010;91(1):149–55.

Lj C. Coefficient alpha and the internal structure of tests. Psychometrika 1951;16(3):297–334.

Long AF, Scott DL. Measuring health status and outcomes in rheumatoid arthritis within routine clinical practice. Br J Rheumatol 1994;33(7):682–5.

Martin Bland J, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. The Lancet 1986;327(8476):307–10.

Morris AD, Kemp GJ, Lees A, Frostick SP. A study of the reproducibility of three different normalisation methods in intramuscular dual fine wire electromyography of the shoulder. J Electromyogr Kinesiol 1998;8(5):317–22.

Nishikawa KC, Murray ST, Flanders M. Do arm postures vary with the speed of reaching? J Neurophysiol 1999;81(5):2582–6.

Paizis C, Papaxanthis C, Berret B, Pozzo T. Reaching beyond arm length in normal aging: adaptation of hand trajectory and dynamic equilibrium. Behav Neurosci 2008;122(6):1361–70.

Papaxanthis C, Pozzo T, Stapley P. Effects of movement direction upon kinematic characteristics of vertical arm pointing movements in man. Neurosci Lett 1998;253(2):103–6.

Piovesan D, Pierobon A, DiZio P, Lackner JR. Experimental measure of arm stiffness during single reaching movements with a time-frequency analysis. J Neurophysiol 2013. http://dx.doi.org/10.1152/jn.01013.2012.

Sainburg R, Kalakanis D. Differences in control of limb dynamics during dominant and nondominant arm reaching. J Neurophysiol 2000;83(5):2661–75.

Sciutti A, Demougeot L, Berret B, Toma S, Sandini G, Papaxanthis C, et al. Visual gravity influences arm movement planning. J Neurophysiol 2012;107(12):3433–45.

Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979;86(2):420.

Shumway-Cook A, Woollacott MH. Motor control: theory and practical applications. 2nd ed., Williams & Wilkins Baltimore; 1995. p. Section IV.

Sleivert GG, Wenger HA. Reliability of measuring isometric and isokinetic peak torque, rate of torque development, integrated electromyography, and tibial nerve conduction velocity. Arch Phys Med Rehabil 1994;75(12):1315–21.

Trent MC, Ahmed AA. Learning from the value of your mistakes: evidence for a risk-sensitive process in movement adaptation. Front Comput Neurosci 2013;2013(7):118.

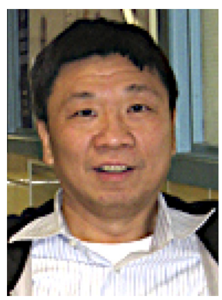
Tyler AE, Karst GM. Timing of muscle activity during reaching while standing: systematic changes with target distance. Gait Posture 2004;20(2):126–33.

Yang JF, Winter D. Electromyographic amplitude normalization methods: improving their sensitivity as diagnostic tools in gait analysis. Arch Phys Med Rehabil 1984;65(9):517–21.

Zhang X, Chaffin D. A three-dimensional dynamic posture prediction model for in-vehicle seated reaching movements: development and validation. Ergonomics 2000;43(9):1314–30.

Rini Varghese is a graduate student, currently pursuing her Masters in Rehabilitation Sciences in the Department of Physical Therapy at the University of Illinois at Chicago. Her thesis work concerns the study of arm movement control in older adults. She will be graduating her program in May 2014 and hopes to pursue doctoral education thereafter.

Christina W.Y. Hui-Chan, PT, PhD is Professor Emeritus in the Department of Physical Therapy, at the University of Illinois at Chicago, prior to which she served as its head (2008–11). She was also the Director of the School of Physical and Occupational Therapy at McGill (1986–91) and Head of the Department of Rehabilitation Sciences at the Hong Kong Polytechnic University (1995–2008). Her research interests are directed at merging the best of Chinese therapeutics with western rehabilitation practice, with a focus on Tai Chi and stroke rehabilitation over the past two decades.



Edward Wang, PhD is a Statistician and Research Associate Professor in the College of Applied Health Sciences at the University of Illinois at Chicago. He has served as a co-Investigator and statistical consultant in several NIH and CDC-funded research projects in health promotion and physical activity for persons with disabilities.

Tanvi Bhatt, PT, PhD is an Assistant Professor in the Department of Physical Therapy, at the University of Illinois at Chicago since 2011. She was a Research Assistant Professor and Post-doctoral fellow with the Department from 2006–2011. She is also the Director of the Motor Behavior and Balance Rehabilitation Laboratory, where her research focus includes investigating movement control and developing learning strategies in community-dwelling older adults and stroke survivors.