Intrarater and Interrater Reliability of Angular Measures Using TEMPLO Two-dimensional Motion Analysis Software

Andrea Greisberger, MSc, Brigitte Wolf, MSc, Klaus Widhalm, MSc, David Kollmitzer, BSc, Maximilian Arbesser, BSc, and Peter Putz, PhD

Abstract

Objective: The purpose of this study was to assess the intrarater and interrater reliability of marking 2 angles with the TEMPLO software and to provide relevant information for clinical practice.
Methods: A prospective test–retest study has been conducted. Four raters took measures on 2 days, with 2 weeks in between. Craniovertebral angle and trunk forward lean were drawn on 22 video frames using TEMPLO. Reliability was examined using intraclass correlation coefficients including standard errors of measurement and minimal detectable change values as measures of precision expressed in the unit of the test (°).
Results: Intraclass correlation coefficients for intrarater and interrater reliability ranged from 0.98 to 1.00. Standard errors of measurement and minimal detectable change values ranged from 0.4° to 0.8° and 0.8° to 2.3°, respectively.
Conclusion: These results indicate excellent reliability for craniovertebral angle and trunk forward lean assessed with TEMPLO software. Changes exceeding 2.3° may be expected to fall outside the test's variability. (J Manipulative Physiol Ther 2019;42:425-429)
Key Indexing Terms: *Posture; Reproducibility of Results*

Introduction

Inappropriate spinal postural alignment is associated with soft tissue stress and changes in the vertebral discs,[1,2] and it might be linked to pain and disability.[3-5] In addition, it has been hypothesized that a forward head posture might be associated with impaired recovery from dysphagia after stroke.[6] Several activities of daily living, such as carrying a backpack or sitting in front of a computer, can change posture significantly.[5,7] Therefore, clinicians routinely assess posture by means of visual inspection because this is feasible with little effort and associated with low costs; however, it has been shown that visual posture inspection might not be yielding reliable outcomes.[8]

Three-dimensional motion capture technology is the most accurate way to assess posture in laboratory conditions. However, such systems are expensive and limited to stationary use, and their use is time consuming and requires extensive training. Therefore, several settings, such as field assessments in research and clinical practice, require an alternative method. Several technical approaches and software-assisted systems exist to support simple and straightforward posture assessments.[9] Our laboratory is equipped with the 2-dimensional video analysis software TEMPLO 7.1 (Contemplas, Kempten, Germany). To our knowledge, the reliability of this method has not been reported previously; therefore, its usefulness for clinical practice and research is not comprehensively established.

Posture assessment usually includes evaluation of the lower extremity, the trunk, and the posture of the head. Trunk forward lean (TFL) and craniovertebral angle (CVA) are acknowledged parameters to assess trunk and head posture.[2,5]

The objective of this study was to assess the intrarater and interrater reliability of marking CVA and TFL with TEMPLO software. Furthermore, the standard errors of measurement (SEM) and the minimal detectable change (MDC) values as measures of precision expressed in the test's unit (°) were calculated.

Department of Health Sciences, FH Campus Wien, Vienna, Austria.

Corresponding author: Andrea Greisberger, MSc, FH Campus Wien, Physiotherapie, Bachelorstudium, Favoritenstraße 226, 1100 Wien. Tel.: +43 1 606 68 77 4717.
(e-mail: *Andrea.greisberger@fh-campuswien.ac.at*).

Paper submitted May 3, 2018; in revised form July 4, 2018; accepted November 2, 2018.
0161-4754

https://doi.org/10.1016/j.jmpt.2018.11.018

Methods

Study Design

A prospective test–retest study (fully crossed) was performed; all participants were scored by all raters on both days. Data collection took place at the university's movement laboratory. Four raters measured the TFL and the CVA on 2 days (2 weeks apart) on the initial frame of 22 videos using the 2-dimensional video analysis software TEMPLO 7.1. Videos were displayed in random order (regarding day and rater). Raters were blinded to the results of the other raters and to their own records of the first day.

Participants

We assumed that the study of Gadotti et al[8] to be transferrable to ours in terms of design and outcomes and hence took their intrarater sample size calculation as a landmark. Consequently, a convenient sample of 22 healthy school children (mean age, 8 ± 1.3 years) was tested between February and May 2015.

Markers were placed on the tragus (14-mm diameter) and the processus spinosus of TH1 and TH12 (25-mm diameter). CVA was measured as the angle between Tragus-TH1 and a horizontal line through TH1. TFL was defined as the angle between TH1-TH12 and a vertical line through TH12. Standing position was standardized with feet aligned on a cross mark attached to the floor. All participants and their parents gave their informed consent. Ethical approval was obtained from the Ethics Committee of the Medical University of Vienna.

Raters

A convenient sample of 4 raters (experienced physiotherapists with 15 to 28 years of practice, age ranging from 41 to 48 years) measured TFL and CVA. Raters 1 and 2 had experience using the software, whereas raters 3 and 4 had not used any software program for assessing posture before.

Instrumentation

Standing position was recorded with a high-speed camera (100 Hz), positioned at a 10° angle to the sagittal plane of the children (distance to the child, 2.95 cm) to improve visibility of the dorsal markers. The camera was attached to a tripod 75 cm above the floor level and was connected to a laptop with the corresponding software. After data collection, the videos were opened with TEMPLO the software, and the corresponding angles where marked manually by the raters by dragging dots to the reference points (TH1, TH12, and Tragus). The required vertical and horizontal lines were drawn automatically by the system.

Raters were instructed on how to use the program, and they completed the assessment on 3 training videos before the actual data collection. Raters were able to pose questions for clarification before they began their assessment. After marking the 2 angles in the first frame of the opened video, the score was noted on a separate sheet and the next video was opened.

Data Analysis

Statistical analysis was performed with SPSS version 24 (IBM Corp, Armonk, New York). Shapiro-Wilk tests and visual inspections of Q-Q plots were performed to check for normality of distribution and outliers. Intrarater and interrater reliability of measuring CVA and TFL were examined separately, by means of intraclass correlation coefficients (ICC) with 95% confidence interval (CI_{95}) expressed for absolute agreement. Agreements of day 1 and day 2 scores were expressed as ICC (3,1) to determine each rater's intrarater reliability. Interrater reliability was calculated from each rater's means of day 1 and day 2 scores, separated for experienced and inexperienced raters and for all raters (ICC3,k). The SEM ($SEM = SD\sqrt{1{-}ICC}$) was calculated to examine the precision of the measurement in the test's unit of measure (°), whereas the standard deviation for all test scores was derived from the total sum of squares of the ICC's analysis of variance ($SD = \sqrt{SS/(n{-}1)}$).[10] The MDC_{95} indicates the required change of a measurement that would, with a certainty of 95%, exceed the intrarater or interrater variability ($MDC_{95} = SEM \times 1.96 \times \sqrt{2}$).[11] Alpha was set at 0.05, although *P* values have been reported.

Results

The mean (±SD) CVA and TFL angles for all participants, separated by rater, are presented in Table 1. ICCs for the intrarater reliability for CVA and TFL ranged from 0.98 to 0.99 and 0.99 to 1.00, respectively. The SEM (and corresponding MDC) ranged from 0.7° to 0.8° (2.0°-2.3°) for CVA and from 0.3° to 0.6° (0.8°-1.8°) for TFL (Table 2).

All ICCs for interrater reliability (experienced and inexperienced raters, both angles) were greater than 0.99 and the overall interrater reliability was excellent (0.99). The SEMs (and MDCs) derived from interrater reliability ICCs ranged from 0.4° to 0.7° (1.1°-1.9°) for CVA and from 0.4° to 0.5° (1.0°-1.5°) for TFL (Table 2).

Discussion

To our knowledge, this is the first study providing data on psychometric properties of posture analysis assessed by multiple raters by means of the TEMPLO software. Intrarater and interrater reliability for measuring angles with TEMPLO was excellent (all ICCs ≥ 0.98). The precision of the instrument, expressed with the MDC, ranges from 0.8° to 2.3°, depending on the measured angle: measuring CVA requires drawing short lines, resulting in bigger MDC values, whereas for quantifying the TFL longer lines have to be drawn (resulting in smaller MDCs).

Comparable research on the reliability of photographic posture assessment with different software programs shows likewise results. Moradi et al[12] and Hazar et al[13] used the MB-Ruler software (freeware), which simply uses an on-screen goniometer to assess head posture. Although skin markers are not necessarily required when applying the MB-Ruler, Hazar et al[13] did so. They measured head position using similar reference points as we did in our

Table 1. *Descriptive Statistics (Means and Standard Deviations) of Day 1, Day 2, and Both Days*

	Craniovertebral Angle			Trunk Forward Lean		
Rater	Day 1 Mean (SD)	Day 2 Mean (SD)	Both days Mean (SD)	Day 1 Mean (SD)	Day 2 Mean (SD)	Both days Mean (SD)
1	41.8 (±6.1)	41.6 (±5.8)	41.7 (±5.9)	7.7 (±5.2)	7.9 (±5.3)	7.8 (±5.2)
2	41.6 (±6.2)	41.6 (±6.2)	41.6 (±6.2)	7.9 (±5.2)	7.8 (±5.3)	7.8 (±5.2)
3	41.5 (±6.1)	41.3 (±5.9)	41.4 (±6.0)	8.1 (±5.3)	7.9 (±5.2)	8.0 (±5.3)
4	41.3 (±5.8)	41.6 (±5.6)	41.5 (±5.7)	8.1 (±5.2)	7.9 (±5.1)	8.0 (±5.1)

SD, standard deviation.

Table 2. *Intrarater and Interrater Reliability*

	Craniovertebral Angle			Trunk Forward Lean		
Reliability	ICC (95% CI)	SEM	MDC	ICC (95% CI)	SEM	MDC
Intrarater						
Rater 1	0.98 (0.95–0.99)	0.8	2.2	0.99 (0.97–0.99)	0.6	1.8
Rater 2	0.98 (0.96–0.99)	0.8	2.3	0.99 (0.98–1.00)	0.4	1.2
Rater 3	0.99 (0.96–0.99)	0.7	2.0	1.00 (0.99–1.00)	0.3	0.8
Rater 4	0.98 (0.95–0.99)	0.8	2.2	0.99 (0.98–1.00)	0.5	1.4
Interrater						
Experienced	0.99 (0.97–0.99)	0.7	1.9	0.99 (0.98–1.00)	0.5	1.5
Not experienced	1.00 (0.98–1.00)	0.4	1.1	1.00 (0.99–1.00)	0.4	1.0
All	0.99 (0.98–1.00)	0.6	1.6	0.99 (0.98–1.00)	0.5	1.3

For all ICCs reported in this table, $P < .01$.
CI, confidence interval; *ICC,* intraclass correlation coefficient; *MDC,* minimal detectable change; *SEM,* standard error of measurement.

study; Moradi et al[12] used different reference points (chin, tragus, and manubrium; no skin markers). Despite those differences in methodology, the ICCs reported for intrarater and interrater reliability ranged from 0.94[12] to 0.98[13] and from 0.88[12] to 0.98[13], respectively. Because higher ICCs were obtained by Hazar et al,[13] it seems that placing skin markers improves reliability. Unfortunately, Hazar et al[13] did not provide data on SEM or MDC. Moradi et al[12] reported an SEM of 0.52° to 1.53°. Our calculated SEMs for the CVA (although measured with slightly different reference points) are similar (range, 0.4°-0.8°). Therefore, it can be hypothesized that these 2 measurement procedures are comparable.

Another freeware package evaluated in prior studies[14,15] is the Postural Assessment Software (PAS/SAPO) (Laboratory of Biomechanics and Motor Control, UFABC - Bloco Ômega , Sao Paolo, Brazil), for which skin markers need to be placed. A major difference compared with TEMPLO is that vertical or horizontal lines, or both, have to be plotted by the investigator. This might be a disadvantage in calculating head position as indicated by data reported by Ferreira et al[14] and Ruivo et al[15]; ICCs and SEMs for intrarater and interrater reliability ranged from 0.66 to 0.93 and from 3.5° to 0.7°, respectively. Interpreting these results compared with our data, it seems that plotting more than 1 line for the calculation of an angle decreases the accuracy of the measurement procedure.

An alternative to using a software program was investigated by Atkins et al[16]; they measured angles on photographs and videos using a smartphone with an integrated inclinometer. Skin markers were applied on reference points. In contrast to our study, joint angles of the lower extremity were captured. The ICCs reported for intrarater and interrater reliability ranged from 0.82 to 1.00, with SEMs ranging from 0.0° to 1.1°. This finding suggests a device of everyday use, like a mobile phone, could also be used for assessing angles reliably.

Cheung Lau et al[17] measured CVA without taking a photograph or video, but by using the electronic head posture instrument, which basically is an electronically operating inclinometer. Taking into account that the reliability of the measurement procedure in this setup also is threatened by the participant (every measurement was taken on a different time point) and by the rater, the obtained values are remarkable; ICCs ranged from 0.86 to 0.91. The resulting SEM is 1.2° and the MDC value is 3.3°. For assessing head posture, electronic head posture instrument might thus be seen as a good and time-efficient alternative to photographic posture analysis because taking pictures or videos and drawing angles with a software program to calculate the angles is not required.

Limitations

Because of our recording setup (the camera was positioned at a 10° angle to the sagittal plane of the children), measured angles in our study are subject to minor distortion and cannot be applied as reference values for different camera setups. However because the objective of this study was not the reliability of posture of the children, but the reliability of drawing angles on the recorded videos, this has not influenced our results. Furthermore, we did not perform an a priori sample size calculation, but we based our consideration on existing studies.[8] Relevant studies had analyzed 20 to 30 participants assessed by 2 to 6 raters.[8,13,16]

Conclusion

According to our results, the reliability of assessing posture with TEMPLO was excellent. When used as an instrument to evaluate change, the MDC value of up to 2.3° should be kept in mind. Freeware and other measuring procedures not requiring skin marker placement appear to be a little less reliable, although also good to excellent compared with the fee-based TEMPLO software.

Funding Sources and Conflicts of Interest

No funding sources or conflicts of interest were reported for this study.

Contributorship Information

Concept development (provided idea for the research): A.G., B.W., K.W.
Design (planned the methods to generate the results): A.G., B.W., D.K., M.A.
Supervision (provided oversight, responsible for organization and implementation, writing of the manuscript): A.G., B.W.
Data collection/processing (responsible for experiments, patient management, organization, or reporting data): B.W., K.W., D.K., M.A.
Analysis/interpretation (responsible for statistical analysis, evaluation, and presentation of the results): A.G., P.P.
Literature search (performed the literature search): A.G., B.W., D.K., M.A.
Writing (responsible for writing a substantive part of the manuscript): A.G., P.P.
Critical review (revised manuscript for intellectual content, this does not relate to spelling and grammar checking): B.W., K.W., D.K., M.A., P.P.
Other (list other specific novel contributions) technical support: K.W.

Practical Applications

- TEMPLO software was used to assess head and trunk posture.
- TEMPLO was a reliable tool for identifying craniovertebral angle and trunk forward lean.
- Changes exceeding 2.3° may be expected to fall outside the test's variability.

References

1. Wilke HJ, Neef P, Caimi M, Hoogland T, Claes LE. New in vivo measurements of pressures in the intervertebral disc in daily life. *Spine.* 1999;24(8):755-762.
2. Brackley HM, Stevenson JM, Selinger JC. Effect of backpack load placement on posture and spinal curvature in prepubescent children. *Work (Reading, Mass).* 2009;32(3):351-360.
3. Yip CHT, Chiu TTW, Poon ATK. The relationship between head posture and severity and disability of patients with neck pain. *Man Ther.* 2008;13(2):148-154.
4. Hwang U-J, Kwon O-Y, Yi C-H, Jeon H-S, Weon J-H, Ha S-M. Predictors of upper trapezius pain with myofascial trigger points in food service workers: The STROBE study. *Medicine.* 2017;96(26):e7252.
5. Kistner F, Fiebert I, Roach K. Effect of backpack load carriage on cervical posture in primary schoolchildren. *Work (Reading, Mass).* 2012;41(1):99-108.
6. Kim SK, Mo SJ, Moon WS, Jun PS, Kim CR. Effects of cervical kyphosis on recovery from dysphagia after stroke. *Ann Rehabil Med.* 2016;40(5):816-825.
7. Brink Y, Louw Q, Grimmer K, Jordaan E. The spinal posture of computing adolescents in a real-life setting. *BMC Musculoskelet Disord.* 2014;15:212.
8. Gadotti IC, Armijo-Olivo S, Silveira A, Magee D. Reliability of the craniocervical posture assessment: visual and angular measurements using photographs and radiographs. *J Manipulative Physiol Ther.* 2013;36(9):619-625.
9. Porto de Freitas Camelo EM, Uchôa DM, Uchoa Santos-Junior FF, de Vasconcelos TB, Macena RHM. Use of softwares for posture assessment: integrative review. *Coluna/Columna.* 2015;14(3):230-235.
10. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res.* 2005;19(1):231-240.
11. Haley SM, Fragala-Pinkham MA. Interpreting change scores of tests and measures used in physical therapy. *Phys Ther.* 2006;86(5):735-743.

12. Moradi N, Maroufi N, Bijankhan M, et al. Intrarater and interrater reliability of sagittal head posture: a novel technique performed by a physiotherapist and a speech and language pathologist. *J Voice.* 2014;28(6):842.
13. Hazar Z, Karabicak GO, Tiftikci U. Reliability of photographic posture analysis of adolescents. *J Phys Ther Sci.* 2015;27(10):3123-3126.
14. Ferreira EAG, Duarte M, Maldonado EP, Burke TN, Marques AP. Postural assessment software (PAS/SAPO): Validation and reliabiliy. *Clinics (Sao Paulo).* 2010;65(7): 675-681.
15. Ruivo RM, Pezarat-Correia P, Carita AI. Intrarater and interrater reliability of photographic measurement of upper-body standing posture of adolescents. *J Manipulative Physiol Ther.* 2015;38(1):74-80.
16. Atkins LT, James CR, Sizer PS, Jonely H, Brismee J-M. Reliability and concurrent criterion validity of a novel technique for analyzing hip kinematics during running. *Physiother Theory Pract.* 2014;30(3):210-217.
17. Cheung Lau HM, Wing Chiu TT, Lam T-H. Clinical measurement of craniovertebral angle by electronic head posture instrument: a test of reliability and validity. *Man Ther.* 2009;14(4):363-368.