Journal of Physical Education and Sport ® (JPES),Vol 21 (Suppl. issue 3), Art 266 pp 2087 – 2096, July.2021
online ISSN: 2247 - 806X; p-ISSN: 2247 – 8051; ISSN - L = 2247 - 8051 © JPES

Original Article

Development of observational indicators for evaluating handstand posture in the mat exercise in physical education class: validity and reliability

MASANORI KOJIMA[1], YOSHINORI KINOMURA[1], KENJI KUZUHARA[2]
[1] Department of Human Health, Faculty of Human Health, Aichi Toho University, Nagoya, JAPAN
[2] Department of Athletic Training and Conditioning, School of Health and Sport Sciences, Chukyo University, Aichi, JAPAN

Published online: July 31, 2021
(Accepted for publication July 15, 2021)
DOI:10.7752/jpes.2021.s3266

Abstract:
Problem Statement: Evaluating skill in gymnastics in physical education classes is required to attain a higher level of skill. However, elementary and junior physical education teachers may not have the necessary background to identify performance errors, and guidance for improvement may be limited. Especially in gymnastics, this presents a significant challenge when inexperienced teachers teach gymnastics classes in physical education. **Purpose:** Handstand is an essential technique related to advanced techniques such as handstand to forward roll, backward roll to handstand, cartwheel, and forward handspring, which are used in elementary and junior high school gymnastics classes. This study aimed to develop an observational evaluation index of handstand posture during mat exercise in physical education classes. **Materials and Methods:** A total of 23 university students (14 males and 9 females) participated in this study. Participants were asked to perform a handstand posture and were photographed from the side. An experienced gymnastics instructor and an experienced track and field instructor evaluated the participants' handstand posture using the images. The gymnastics instructors scored handstand posture using E-scores created by technical and aesthetic evaluation items related to handstand posture extracted from gymnastics scoring rules. Observational evaluation indices corresponding to the evaluated technical and aesthetic/implementation perspectives were created. In addition, the posture of the whole body and joint angles were calculated from the images. Further, the number of walking steps, the angle at which the leg was opened, and whether the participant fell were assessed. The validity and inter-and intra-observer reliability of all observational evaluation indices were verified. **Results:** Significant correlations were found between the evaluation scores of each posture model and the E-score of the whole body ($r=0.65-1.00$, $p<0.001$), and between the evaluation scores of each joint angle and the joint flexion of body parts ($r=0.58-0.95$, $p<0.001$), indicating the validity of these evaluation scores. The evaluation scores of each posture model and each joint angle (intra-rater kappa coefficient: 0.73-1.00, inter-rater kappa coefficient: 0.68-0.92) and walking, leg opening, and falling (intra-rater kappa coefficient: 0.88-1.00, inter-rater kappa coefficient: 0.84-0.92) were highly reliable. The intra-rater and inter-rater kappa coefficients were also reliable. **Conclusions:** Our results showed that the item that strongly affected the E-score was falling, while overall body posture and hip bend strongly affected performance. Therefore, these results suggest that it is crucial to avoid these defects to improve handstand performance.
Key Words*:* **gymnastic, novice, qualitative evaluation, inverted balance, physical education teacher**

Introduction

When evaluating handstand posture in gymnastics, judges follow specific gymnastics scoring rules (International Gymnastics Federation, 2017). In the gymnastics scoring method, the combination of the D-score (the total score of difficulty) and the E-score (the total deduction of technical and compositional, and aesthetic and execution errors) is evaluated for the performance (Onishi, 2017). However, in curriculum guidelines for junior high school and high school physical education in Japan, the skill performance in gymnastics is assessed as "Understanding a rational solution so that the technique can be performed better" (Ministry of Education, Culture, Sports, Science, and Technology, 2018). Thus, learning content is designed to increase the qualitative value (technical value and aesthetic and execution value) of the technique in line with the values of gymnastics in physical education. Evaluating the skill of gymnastics in physical education is similar to that of gymnastics at the international level, and evaluation is required to be able to perform the skill better" (Ministry of Education, Culture, Sports, Science and Technology, 2018). It is important to pay close attention to evaluation from a qualitative point of view (Tanaka, Murata, & Gushiken, 2015). However, in physical education gymnastics, there is a tendency to observe and evaluate quantitatively, such as "high, strong, and fast", rather than qualitatively observing and evaluating techniques such as in the skill performance. In addition, physical education teachers

Corresponding Author YOSHINORI KINOMURA, E-mail: kinomura.yoshinori@aichi-toho.ac.jp

may not have the necessary background in all sports to identify performance errors (Khong & Kong, 2016), and, for those sports in which the teacher is inexperienced, it is assumed that guidance will be given with a quantitative perspective, which is easier to observe and evaluate. Thus, especially in gymnastics, it becomes impossible to provide a qualitative perspective from the motor characteristics, which presents a significant challenge when inexperienced teachers teach gymnastics classes in physical education.

In other sports, evaluation methods have been developed, allowing many physical education teachers and instructors to quantitatively assess qualitative perspectives on performance(Aburano, Ogata, Sekioka, Nagai, & Shimizu, 1995; Aglioti, Cesari, Romani, & Urgesi, 2008; Amara et al., 2015; Fujita, Ikeda, Chin, & Takeda, 2010; Hiruma, Mori, & Ogata, 2014; Kaji, Tomozoe, Yoshinaga, & Suzuki, 2017; Knudson & Morrison, 1996; Kokudo, 2012; Nakamura et al., 2011; Roberton & Konczak, 2001; Suzuki, Tomozoe, Yoshinaga, Kaji, & Hirayama, 2016; Takamoto, Ogata, & Idei, 2003; Takizawa & Tomozoe, 2017), and students are able to learn from this qualitative point of view (Knudson & Morrison, 2007). Since observation evaluation depends on the sports experience of the physical education teacher, it is necessary to use methods that more teachers can apply and to consider the priority of which method to emphasize (Adachi & Suzuki, 2009; Hosogoshi, Nakamura, Yonemura, & Takahashi, 2009; Noda, Asaoka, Hasegawa, & Kato, 2009).

To the best of our knowledge, studies conducting an observational evaluation of gymnastics in physical education other than forward and backward rolls (Kato, Kawamoto, Ae, & Morioka, 2014) and horizontal bar movement (Harigai & Kondoh, 2017) have not been verified. Handstand is an essential technique related to advanced techniques such as "handstand to forward roll, backward roll to handstand, cartwheel, and forward handspring", which are used in elementary and junior high school gymnastics classes. This study aimed to develop an observational evaluation index of handstand posture during mat exercise in physical education classes.

Material & methods

Participants and trials

Participants were 23 college students, including 14 males and nine females, who had no experience in gymnastics and took gymnastics classes at University A. The class content was "to straighten the posture of the whole body in handstand and to acquire an accurate handstand posture". Regarding handstand posture, two trials (before and after instruction) were analyzed.

In recording the trials, video for the first handstand trial was obtained at the beginning of the first lesson and that of the second trial was obtained at the end of the second lesson. In addition, in order to ensure safety, the trials were performed on a competition landing mat (ngc system mat, AM920280 Landing Mat, Senoh Corp., Japan) with spotting support assistance. We provided support assistance when risk was present (handstand falls with excessive force). In addition, we instructed the students to try the handstands while considering their own abilities.

As an ethical consideration, to obtain the participants' consent to participate, the Declaration of Helsinki was observed and the purpose, method, and risk of the experiment were fully explained in advance. We also explained that it is possible to cancel or suspend participation at any time, and obtained consent for voluntary participation. This study was approved by the research ethics committee of Aichi Toho University.

Observation evaluators

Two observation evaluators (A and B) were selected. Evaluator A was a researcher in the field of physical education with 20 years of experience in gymnastics, more than 15 years of coaching in gymnastics, and a certified coach at international level accredited by the Japan Sport Association. Evaluator B was a researcher in the field of coaching, with 22 years of track and field experience, and 12 years or more of coaching in track and field. The E-score was scored by the Evaluator A, the observational evaluation index was scored by the Evaluators A and B, and the validity and reliability of the observational evaluation score were examined.

Measuring methods

Observation evaluation was performed using video and still images. Since the participants had no experience in gymnastics and did not have the skill to stop in a handstand position, the image of handstand posture in which the participant was closest to the vertical line was used. For all trials, a digital video camera (EX-ZR3100, CASIO) was installed on the side of the trial (approximately 10 m away within 20°), and shooting was performed at 60 fps. Video analysis software, Dart Trainer Pro (Dartfish Japan Co., Ltd., Tokyo), was used for angle analysis, and seven body parts (center of the wrist joint, center of the elbow joint, center of the shoulder joint, center of the hip joint, center of the knee joint, lateral malleolus of the ankle joint, fifth metatarsal of the foot) were digitized to calculate the posture angle and each joint angle (Kinomura, Kojima, & Kuzuhara, 2018).

Scoring methods of E-score

Gymnastics competitions are scored based on "difficulty, beauty," and the combination of the D-score and the E-score is evaluated as the performance of the technique (International Gymnastics Federation, 2017). The D-score evaluates the difficulty of the technique and the E-score evaluates the aesthetic, execution, and technical performance. In particular, the E-score evaluates performance by deducting points when there are defects. Therefore, the performance (P) of gymnastics can be expressed by the following formula:

P = D-score + E-score

In this study, the evaluation perspective is based on the scoring rules of gymnastics competitions (D-score and E-score). Therefore, when trying to evaluate the performance of handstand posture in gymnastics class, the D-score becomes a constant and the E-score becomes a variable to evaluate the performance of handstand posture. Then, P in handstand was evaluated. When evaluating handstand in the E-score, the technical viewpoint and the aesthetic and execution viewpoints are used in the scoring rules, so the E-score can be expressed by the following formula:

E-score = Technical perspective + Aesthetic and Execution perspective

In the technical defect, "correct stationary posture" is a straight handstand posture, which is also shown in the difficulty table in the scoring rules (International Gymnastics Federation, 2017). When observing and evaluating a handstand, the technical defects of "angle deviation from the correct stationary posture, wobbling or falling" and the aesthetic and execution defects of "bending arms or legs, opening of legs, walking", based on the scoring rules of the Japan Gymnastics Association (2017), were used in this study. "Wobbling or falling" is an item to evaluate the stability of the handstand, but "wobbling" is an expression that assumes that the handstand can be stopped. It is difficult for general students in physical education classes to "wobble", so "falling" was used alone in this study. Regarding "walking", we evaluated participants moving their hands at handstand as aesthetic and execution defects. Thus, the following formula was used to calculate the E-score:

E-score = [Posture of the whole body, Falling] + [Bending arms or legs, Opening of legs, Walking]

Development of observational evaluation models

First, from the general defects of the scoring rules, two technical viewpoints related to handstand ("angle deviation from the correct stationary posture" and "falling from handstand") and three aesthetic and execution viewpoints ("bending of arms or legs", "opening of legs", and "walking or jumping") were extracted. Next, we devised three posture models (Fig. 1) for observational evaluation of handstand posture, and created an observational evaluation index that can be evaluated quantitatively. For these three posture models, the posture angle of the whole body is evaluated, and the "straight handstand posture" indicated by the scoring rules was used as the evaluation index.

The evaluation judgment was made by Evaluator A who was familiar with gymnastics. A vertical line was drawn from the center of the wrist joint, the angle formed by the line segment (reference value 0°) was used as a reference, and the angle difference between the vertical line and the straight line passing through the center of each joint was analyzed in each model. Regarding the joint angle, in order to evaluate the bending of each joint in handstand posture, the angle formed by the line segment connecting the centers with each joint was analyzed. The definitions of the calculation items are shown below.

Posture Model 1: A posture model connecting the center of the wrist joint to the center of the hip joint.
Posture Model 2: A posture model connecting the center of the wrist joint to the center of the knee joint.
Posture Model 3: A posture model connecting the center of the wrist joint to the center of the lateral malleolus of the ankle.
In evaluating these models, the angle formed by the vertical line was used (Fig. 1) as described below.

Elbow joint angle: The angle formed by the line segment connecting the center of the wrist joint, the center of the elbow joint, and the center of the shoulder joint.
Shoulder joint angle: The angle formed by the line segment connecting the center of the elbow joint, the center of the shoulder joint, and the center of the hip joint.
Hip joint angle: The angle formed by the line segment connecting the center of the shoulder joint, the center of the hip joint, and the center of the knee joint.
Knee joint angle: The angle formed by the line segment connecting the center of the hip joint, the center of the knee joint, and the lateral malleolus of the ankle joint.
Ankle joint angle: The angle formed by the line segment connecting the center of the knee joint, the center of the lateral malleolus of the ankle, and the fifth metatarsal bone.

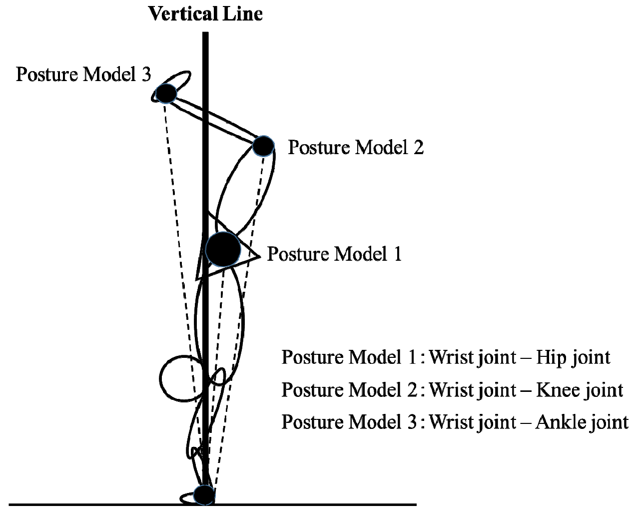

Figure 1. Three posture models

Evaluation index scoring

Postural angles, joint angles, walking, opening of legs, and falling in handstand were scored for each item. The evaluations were set to 3.0 points for posture angles and joint angles from 0° to 15° (small defects), 2.0 points from 16° to 30° (medium defects), 1.0 point from 31° to 45° (major defect), and 0 points for ≥46° (major negligence). Walking was 3.0 points for no steps, 2.0 points for 1 step, 1.0 point for 2 steps, and 0.0 points for 3 steps or more. Leg opening was 3.0 points from 0° to 15°, 2.0 points from 16° to 30°, 1.0 point from 31° to 45°, and 0.0 points for 46° or more. Fall from the handstand was 0.0 points for falling and 3.0 points for not falling.

Important points for evaluating E-score

At the time of scoring, an observational evaluation index consisting of the two viewpoints and five items above was used. From a technical point of view, the item "deviation of angle from correct stationary posture" evaluated the angle of deviation of the posture of the whole body from the vertical line at handstand. For "falling", we evaluated falling after the handstand was stopped. From aesthetic and execution points of view, the item "walking or jumping" evaluated only walking in handstand because no participants jumped in handstand in this study. The item "bending arms or legs" evaluated the bending of body parts. Since the item of "opening of legs" is observed from the side of all attempts, we decided to use only "opening of legs to the front and rear" that can be observed from the side of the body.

Validity and reliability of observational evaluation

The validity of the observational evaluation index was examined by analyzing the correlation between the actual angles (posture angle and each joint angle) and the corresponding evaluation points. The reliability within and between the evaluators was examined. For reliability within the evaluator, the same evaluator evaluated handstand posture of the same participant twice using the re-examination method for the calculation items of the two viewpoints consisting of the posture angle of each model and each joint angle at handstand. We verified the consistency of the evaluation. Since it is necessary to consider the motor observation ability of the evaluator for the reliability between the evaluators, and the motor observation affects the motor experience of the evaluator (Asaoka, 2019, Hiruma et al., 2014; Pizzera, A., Möller, C., & Plessner, H., 2018), we decided to select different evaluators in consideration of their motor experience and skills.

Statistical analysis

Spearman's rank correlation coefficient was used to examine the validity of the observational evaluation index. The weighted Kappa coefficient (κ coefficient) was calculated in order to examine the reliability within and between the evaluators of the observational evaluation. Furthermore, in order to examine the factors that strongly influenced the evaluation of handstand posture, we used the forward-backward stepwise method of multiple regression analysis using the E-score and observational evaluation. The level of significance level was set to less than 5%.

Results

Posture angle of the whole body and each joint angle

Table 1 shows the posture angle of each posture model's whole body, and the mean and standard deviation of each joint angle of the body part. Posture Model 1 had a posture angle of 6.0 ± 6.0°, Posture Model 2 had a posture angle of 4.5 ± 4.8°, and Posture Model 3 had a posture angle of 6.6 ± 6.1°. The elbow joint angle was 143.6 ± 14.5°, the shoulder joint angle was 136.5± 9.2°, the hip joint angle was 168.9 ± 7.2°, the knee joint angle was 171.0 ± 11.6°, and the ankle joint angle was 150.5 ± 11.8°.

Table 1. Posture angle of the whole body and each joint angle

		Mean	*SD*
Postural angle of the entire body (deg)	Posture Model 1	6.0	6.0
	Posture Model 2	4.5	4.8
	Posture Model 3	6.6	6.1
Joint angles (deg)	Elbow	143.6	14.5
	Shoulder	136.5	9.2
	Hip	168.9	7.2
	Knee	171.0	11.6
	Ankle	150.5	11.8

Mean value of each item of E-score and reliability within the evaluator

Table 2 shows the results of each item of the E-score regarding handstand posture and reliability within each evaluator. From a technical point of view, "angle deviation from the correct stationary posture" was -0.13 ± 0.14 (κ = 0.92, p <0.05) for both the first and second trials, and "falling from handstand" was -0.31 ±. 0.34 in the first trial and-0.32 ± 0.36 (κ = 0.96, p <0.05) in the second trial, "walking from handstand" from the aesthetic and execution points of view was -0.11 ± 0.04 (κ = 1.00, p <0.05) for both the first and second trials, "bending arms or legs, opening legs" was -0.14 ± 0.08 in the first trial and -0.13 ± 0.07 (κ = 0.85, p <0.05) in the second trial. High reliability was demonstrated in all evaluation viewpoint items.

Table 2. Mean value of each item of E-score and reliability within evaluator

	Evaluation Items	First evaluation		Second evaluation		
		Mean	*SD*	*Mean*	*SD*	*κ*
Technical elements	Deviation of angle from a correct handstand posture	-0.13	0.14	-0.13	0.14	0.92
	Falling at a handstand posture	-0.31	0.34	-0.32	0.36	0.96
Aesthetic elements	Walking or jumping at a handstand posture	-0.11	0.04	-0.11	0.04	1.00
	Bending arms, legs, or opening legs	-0.14	0.08	-0.13	0.07	0.85
	E-score	-0.69	0.41	-0.69	0.40	0.96

Relationship between the angle of each posture model and the E-score

Table 3 shows the relationship between the angle of each posture model and the E-score, which indicates the posture of the whole body. Significant correlations were found between the E-score and the posture model (Posture Model 1, r = -0.44, p <0.01; Posture Model 2, r = -0.40, p <0.01; Posture Model 3, r = -0.40, p <0.01).

Table 3. Relationship between the angle of each posture model and the E-score

Posture Model 1	-0.44	**
Posture Model 2	-0.40	**
Posture Model 3	-0.40	**

**: p<0.01

Evaluation points of observational evaluation indexes

Table 4 shows the evaluation of each posture model's observational evaluation indexes scored by each evaluator. Evaluator A had the highest evaluation points for the first and second trials in Posture Model 1 and Posture Model 2, and Evaluator B had the highest evaluation points for both the first and second trials in Posture Model 1.

Regarding the evaluation points of the joint angles, both Evaluators A and B had high evaluation points for the hip joint (2.74 to 2.78) and knee joint (2.76 to 2.78) in the first and second trial. Both evaluation points for the shoulder joint (1.30 to 1.59) and ankle joint (1.72 to 1.91) were lower than two points. Regarding other items, the evaluation points for walking (2.74) were the highest for both Evaluators A and B, and the evaluation points for falling (1.70 to 1.76) were the lowest for both Evaluators A and B.

Relationship between evaluation points of each posture model and E-score

Table 5 shows the correlation coefficient between each posture model's evaluation points and the E-score (angle deviation from the correct stationary posture). Both Evaluators A and B showed a significantly high correlation between the evaluation points of all posture models and the E-score (p <0.001), of which Posture Model 1 showed the highest correlation coefficient.

Reliability within and between evaluators at the evaluation points of each posture model

Table 6 shows the reliability within and between the evaluators at the evaluation points of each posture model. The reliability within and between the evaluators of Posture Model 1 was 0.73-0.85 and 0.85, respectively. The reliability within and between the evaluators of Posture Model 2 was 0.91-0.92 and 0.92, respectively. The reliability within and between the evaluators of Posture Model 3 was 1.00 and 0.68, respectively, and high reliability was observed in all posture models.

Table 4. Evaluation points of observational evaluation indexes

Evaluator	Items	First evaluation		Second evaluation	
		Mean	*SD*	*Mean*	*SD*
Evaluator A	Evaluation points for Posture Model 1	2.91	0.28	2.93	0.25
	Evaluation points for Posture Model 2	2.91	0.35	2.93	0.33
	Evaluation points for Posture Model 3	2.87	0.40	2.87	0.40
	Evaluation points of elbow joint angle	2.24	0.71	2.24	0.71
	Evaluation points of shoulde joint angle	1.30	1.15	1.30	1.11
	Evaluation points of hip Joint Angle	2.76	0.48	2.74	0.49
	Evaluation points of knee joint angle	2.78	0.63	2.76	0.71
	Evaluation points of ankle joint angle	1.72	0.78	1.74	0.77
	Evaluation points for walking while handstand	2.74	0.80	2.74	0.80
	Evaluation of leg opening during handstand	2.59	0.65	2.61	0.65
	Evaluation of falling during handstand	1.76	1.49	1.76	1.49
Evaluator B	Evaluation points for Posture Model 1	2.93	0.25	2.89	0.31
	Evaluation points for Posture Model 2	2.89	0.38	2.91	0.35
	Evaluation points for Posture Model 3	2.74	0.61	2.74	0.61
	Evaluation points of elbow joint angle	2.35	0.74	2.33	0.70
	Evaluation points of shoulde joint angle	1.50	1.09	1.59	1.05
	Evaluation points of hip Joint Angle	2.78	0.47	2.78	0.47
	Evaluation points of knee joint angle	2.76	0.60	2.78	0.51
	Evaluation points of ankle joint angle	1.91	0.78	1.89	0.67
	Evaluation points for walking while handstand	2.74	0.80	2.74	0.80
	Evaluation of leg opening during handstand	2.57	0.69	2.59	0.65
	Evaluation of falling during handstand	1.70	1.50	1.76	1.49

Table 5. Relationship between evaluation points of each posture model and E-score

Evaluator	Items	Deviation of angle from a correct handstand posture	
Evaluator A	Evaluation points for Posture Model 1	1.00	***
	Evaluation points for Posture Model 2	0.86	***
	Evaluation points for Posture Model 3	0.65	***
Evaluator B	Evaluation points for Posture Model 1	0.86	***
	Evaluation points for Posture Model 2	0.74	***
	Evaluation points for Posture Model 3	0.70	

***: $p<0.001$

Table 6. Reliability within and between evaluators at the evaluation points of each posture model

		κ
Posture Model 1	Intra-rater reliability (Evaluator A)	0.85
	Intra-rater reliability (Evaluator B)	0.73
	Inter-rater reliability	0.85
Posture Model 2	Intra-rater reliability (Evaluator A)	0.91
	Intra-rater reliability (Evaluator B)	0.92
	Inter-rater reliability	0.92
Posture Model 3	Intra-rater reliability (Evaluator A)	1.00
	Intra-rater reliability (Evaluator B)	1.00
	Inter-rater reliability	0.68

Relationship between each joint angle and the evaluation point of joint bending

Table 7 shows the relationship between each joint angle and the evaluation points of joint bending. A significant high correlation was found between each joint angle and all evaluation points for joint bending (Evaluator A, r = 0.58-0.95, p <0.001; Evaluator B, r = 0.58-0.87, p. <0.001).

Table 7. Relationship between each joint angle and the evaluation point of joint bending

	Evaluation points for the corresponding joints	
	Evaluator A	Evaluator B
Elbow	0.81 ***	0.87 ***
Shoulder	0.95 ***	0.85 ***
Hip	0.72 ***	0.66 ***
Knee	0.58 ***	0.58 ***
Ankle	0.80 ***	0.74 ***

***: p<0.001

Reliability within and between evaluators in the evaluation points of each joint angle, walking, opening of legs, and falling

Table 8 shows the reliability within and between evaluators in the evaluation points of each joint angle, walking, opening of legs, and falling. Regarding the reliability within the evaluator at each joint angle, Evaluator A had κ = 0.88-0.98, Evaluator B had κ = 0.82-0.94, and the reliability between evaluators was κ = 0.70-0.89. Regarding the reliability within the evaluator at walking, opening of legs, and falling, Evaluator A had κ = 0.97-1.00, Evaluator B had κ = 0.93-1.00, and the reliability between evaluators was κ = 0.93-1.00. High reliability was shown at all evaluation points.

Table 8. Reliability within and between evaluators at each joint angle, walking, opening of legs, and falling evaluation points

	Ankle	Knee	Hip	Elbow	Shoulder	Walking	,eg openin	Falling
Intra-rater reliability (Evaluator	0.982	0.876	0.953	0.911	0.965	1.000	0.974	1.000
Intra-rater reliability (Evaluator	0.855	0.823	0.898	0.936	0.942	1.000	0.926	0.956
Inter-rater reliability	0.700	0.795	0.852	0.768	0.887	1.000	0.926	0.956

Relationship between E-score and each evaluation point (multiple linear regression analysis)

Table 9 shows the results of multiple linear regression analysis with the E-scores as the objective variable and each evaluation score as the explanatory variable. At this time, since Posture Model 1 showed the highest validity, Posture Model 1 was selected for the posture of the whole body. Furthermore, using all other evaluation points, the model for each evaluator was examined using the forward-backward stepwise method of multiple linear regression analysis. As a result, the items of falling were significant for both Evaluators A and B.

Table 9. Multiple linear regression analysis on E-score (Upper: Evaluator A, Lower: Evaluator B)

Evaluator A

Independent variable	β	SEB	p
Posture Model 1	0.33	0.15	0.03 *
Hip	0.30	0.09	0.00 ***
Falling	0.15	0.03	0.00 ***
Intercept	-2.76	0.41	0.00 ***
SEE	R^2	adiR^2	
0.26	0.62	0.60	

Evaluator B

Independent variable	β	SEB	p
Posture Model 1	0.33	0.18	0.07 #
Walking	-0.20	0.08	0.02 *
Falling	0.24	0.04	0.00 ***
Hip	0.16	0.09	0.08 #
Knee	-0.15	0.08	0.07 #
Intercept	-1.64	0.62	0.01 *
SEE	R^2	adiR^2	
0.26	0.65	0.60	

***: p<0.001, **: p<0.01, *: p<0.05, #: p<0.1

Discussion

In this study, we examined an evaluation index for handstand posture in order to enable the observational evaluation of handstand posture in gymnastic physical education classes. We also examined the validity and reliability of the evaluation index and the hierarchy of technical issues for the handstand.

Validity and reliability of posture model evaluation points and E-scores

By analyzing the correlation between each posture model's evaluation points and the E-score (angle deviation from the correct stationary posture), the validity of the posture model of the whole body at handstand was examined. As a result, a significant correlation ($r = 0.65$-1.00) was recognized between the evaluation points of each posture model and the E-score. In particular, the correlation coefficient ($r = 0.86$-1.00) of Posture Model 1 connecting the wrist joint and the hip joint showed the highest relationship ($p < 0.001$).

Next, the reliability within the evaluator and between the evaluators was examined for the evaluation points of each posture model. As a result, the κ coefficient within the evaluator and between the evaluators was high for all posture models (within the evaluator, 0.73-1.00; between the evaluators, 0.68-0.92). This result indicates that the reliability within and between the evaluators was high in all posture models, and that the most valid posture model was Posture Model 1.

Matsuyama (2012) proposed that when a beginner is inverted, the target image should be a handstand posture in which the whole body is bent backwards to ensure stability and reduce the difficulty of the lesson. However, if Posture Models 2 and 3 are adopted, the body is bent backward and the knee and ankle joints are directly above. Although the hip joint is out of the vertical line, it is assumed that the evaluation score will be high even though the posture of the whole body is not straight. Therefore, when evaluating handstand posture, Posture Model 1 is considered to be the most appropriate posture model of the whole body because handstand posture in which the hip joint deviates from the vertical line can be evaluated step by step.

Validity and reliability of joint angle and evaluation points of joint bending

The validity and reliability of each joint angle and the evaluation points of joint bending were examined. As a result, the correlation coefficient showed a significantly high relationship of 0.58-0.95 in both evaluators. The validity of all evaluation points in the bending of body parts was ensured ($p < 0.001$). In addition, the κ coefficient showed a significantly high relationship of 0.70-0.98 both within the evaluator and between the evaluators, so the reliability was also ensured.

We examined the reliability within and between evaluators regarding falling from the technical viewpoint and walking and opening legs from the aesthetic and execution viewpoints. As a result, the within evaluator κ coefficient of falling was 0.95. The between evaluator κ coefficient was 0.84, which was significantly high. The κ coefficient of walking and opening legs within the evaluator was 0.88-1.00 and between the evaluators was 0.89-0.92, showing a significantly high relationship. Therefore, our results indicate that this evaluation index is highly reliable and that handstand posture can be observed and evaluated to the same extent as a gymnastics specialist from any evaluation viewpoint.

Examination of the hierarchical structure of handstand posture using each evaluation point (multiple linear regression analysis)

In the previous section, the validity and reliability of the posture models were verified, and Posture Model 1 was adopted for the posture of the whole body. In addition, the validity and reliability of the evaluation points for falling, joint bending in body parts, walking, and leg opening were also verified. In order to examine the relationship between the E-score, which is regarded as the performance of an athletic skill, and each evaluation item, a multiple regression analysis with the E-score as the objective variable was conducted for each evaluator using the forward-backward stepwise method (Table 9). As a result, the coefficients of determination were generally around 0.6 in all models, and the evaluation score on falling was identified as a significant variable. When the significance was set at 10%, the items selected for both models were posture model, falling, and hip joint evaluation scores. If a person falls during handstand, he or she loses a large number of points. Therefore, the most important factor related to the E-score is to maintain a stable handstand posture. In order to increase the E-score while maintaining stability of handstand posture, straightening the posture of the whole body and reducing bending of the hip joint contribute to improvement. Based on these results, when dealing with handstand in the mat exercise class, it is necessary to first check whether the student has the stability to perform without falling, and, then, proceed with learning to secure handstand posture of the whole body in order to improve the performance of handstand.

The evaluation indices developed in this study are variables extracted from the scoring criteria. Straightening the posture of the whole body and reducing bending of the joints itself may lead to an increase in the difficulty of maintaining handstand posture, and, if a person is about to lose his balance, he may try to maintain his balance by walking or opening his legs. Variables not selected for the multiple linear regression equation may still be important for instruction, but falling and body posture as a whole are essential priorities for immediate assessment of handstand posture.

In this study, only handstand posture was evaluated, and not the preparatory movements leading up to handstand posture. In addition, a standard for evaluating and judging handstand posture step by step based on the evaluation index created in this study is required. For this purpose, we must verify the validity and reliability of the evaluation index widely by adding junior high school and high school teachers as new evaluators and expanding the evaluation to junior high school and high school students.

Conclusions

The purpose of this study was to develop an observational evaluation index of handstand posture based on the criteria of gymnastics scoring rules for handstand posture in the mat exercise. By creating observational evaluation indexes for handstand posture and verifying their validity and reliability and the hierarchy of technical issues, we have developed observational evaluation indexes for handstand posture that can be used generally in physical education gymnastic classes.

The evaluation of the whole body posture, the correlation coefficient of Posture Model 1 connecting the wrist joint and the hip joint showed the highest validity, while the κ coefficient of all posture models within and between evaluators showed the highest reliability. Thus, reliability within and between evaluators was also ensured.

Validity was demonstrated by the significant correlation between each joint angle and the evaluation point of joint bending. The κ coefficients of falling from a technical point of view and all items from aesthetic and execution points of view showed significantly higher relationships within and between evaluators, ensuring high reliability.

Finally, as a result of multiple regression analysis of the relationship between the E-score, which is regarded as the skill performance of gymnastics, and each evaluation item, no falling in handstand was most related to the E-score, suggesting that it is important to reduce the deviation of the posture of the whole body and the bending of the hip joint.

Conflicts of interest - There is no conflict of interest.

References

Aburano, T., Ogata, M., Sekioka, Y., Nagai, J., & Shimizu, S. (1995). Study on observational evaluation of throwing for adult women. *Sports Pedagogy Research*, 15 (1): 15-24.

Adachi, M., & Suzuki, H. (2009). A study on the acquisition of techniques in elementary school gymnastics classes. *Physical Education Pedagogy Research (in Japanese)*. 25 (1): 15-38.

Aglioti, S.M., Cesari, P., Romani, M., & Urgesi, C. (2008). Action anticipation and motor resonance in elite basketball players. *Nature Neuroscience*, 11(9): 1109-1116.

Amara, S., Mkaouer, B., Nassib, S.H., Chaaben, H., Hachana, Y., & Ben Salah, F.Z. (2015). Effect of video modeling process on teaching/learning hurdle clearance situations on physical education students. Advances in Physical Education, 5: 225-233.

Asaoka, M. (2019). Introduction to the aesthesiological movement theory of sports. *Taishukan Publishing (in Japanese)*, pp.171-174.

Fujita, I., Ikeda, N., Chin, Y., & Takeda, Y. (2010). A study of technical elements contributing to achievement of the target record in high jump (scissors jump) performance: Derivation of criteria based on observational evaluation and longitudinal practice in upper-grade elementary school children. *Japan Journal of Physical Education, Health and Sport Sciences*, 55 (2): 539-552.

Harigai, M., & Kondoh, T. (2017). A study on derivation of criteria for observational evaluation of horizontal bar movement: focusing on the basic movement for first to fourth grade elementary school children. *Physical Education Pedagogy Research*, 33 (2): 19-34.

Hiruma, K., Mori, K., & Ogata, M. (2014). Research on the observation points of the batting motion which can be used by physical education class -the batting motions in the technical college students-. *Sports Pedagogy Research,* 34 (1): 23-32.

Hosogoshi, J., Nakamura, T., Yonemura, K., & Takahashi, T. (2009). A study on movement tasks as analogon to acquire the skills of the astride vault. *Sports Pedagogy Research*, 21 (2): 81-92.

International Gymnastics Federation. (2017). 2017 Code of points: men' 's artistic gymnastics. Retrieved from https://www.gymnastics.sport/publicdir/rules/files/en_MAG CoP 2017 - 2020.

Japan Gymnastics Association. (2017). Gymnastics men's scoring rules. *Japan Gymnastics Association*, pp.23-31.

Kaji, M., Tomozoe, H., Yoshinaga, T., & Suzuki, K. (2017). Creation of an observational motion evaluation method for sprinting motion in the middle grades of elementary school. *Physical Education Pedagogy Research (in Japanese)*, 33 (2): 49-64.

Kato, K., Kawamoto, C., Ae, M., & Morioka, Y. (2014). Validity of their observational evaluation in forward and backward rolls for elementary school children. *Japan Journal of Human Growth and Development Research*, 64: 1-10.

Khong, S.W.J., & Kong, P.W. (2016). A simple and objective method for analyzing a gymnastics skill. *European Journal of Physical Education and Sport*, 12(2): 46-57.

Kinomura, Y., Kojima, M., & Kuzuhara, K. (2018). Reliability and validity of upper and lower limb joint angles calculated using DARTFISH: A two-dimensional motion analysis of inverted motion as an example. *Strength & Conditioning Journal Japan (in Japanese)*, 25 (4): 12-18.

Knudson, D., & Morrison, C. (1996). An integrated qualitative analysis of overarm throwing. *Journal of Physical Education, Recreation & Dance,* 67(6): 31-36

Knudson, D.V., & Morrison, C.S. (2007). Qualitative analysis of human movement (2nd ed.). (M. Ae, Trans.). Nap.

Kokudo, S. (2012). An investigation of evaluation points on overhand throw form for children considering the motion causality. *Japan Journal of Human Growth and Development Research*, 55: 1-10.

Matsuyama, N. (2012). A study of the movement instruction example -handstand training compared with standing training-. *Bulletin of Biwako Seikei Sport College*, 9: 83-91.

Ministry of Education, Culture, Sports, Science and Technology. (2015). School physical education practical guidance material: 10th edition gymnastics guidance. *Toyokan Publishing (in Japanese)*, pp.16-20.

Ministry of Education, Culture, Sports, Science and Technology. (2018). Junior high school learning guidelines (Notification in 2017): health and physical education. *Higashiyama Publishing (in Japanese)*, p.63, pp.75-76.

Ministry of Education, Culture, Sports, Science and Technology. (2019). High school learning guidelines (Notification in 2018): health and physical education. *Higashiyama Publishing (in Japanese)*, p.58, p.67.

Nakamura, K., Takenaga, R., Kawaji, M., Kawazoe, K., Shinohara, T., Yamamoto, T., Yamagata, Z., & Miyamaru, M. (2011). Development of basic movement patterns of infants by observational evaluation methods. *Japan Journal of Human Growth and Development Research*, 51: 1-18.

Noda, T., Asaoka, M., Hasegawa, K., & Kato, S. (2009). Effects of differences in methods of presenting visual information on the understanding of movement processes: Observation of movements in apparatus gymnastics. *Japan Journal of Physical Education, Health and Sport Sciences*, 54 (1): 15-28.

Onishi, N. (2017). Changes and trends in scoring rules for the vault in men's gymnastics from 2017 to 2020. The Annual Reports of Health, Physical Education and Sport Science. 36: 83-89.

Pizzera, A., Möller, C., & Plessner, H. (2018). Gaze behavior of gymnastics judges: Where do experienced judges and gymnasts look while judging?. Research Quarterly for Exercise and Sport, 89(1), 112-119.

Roberton, M.A., & Konczak, J. (2001). Predicting children's overarm throw ball velocities from their developmental levels in throwing. *Research Quarterly for Exercise and Sport*, 72(2): 91-103.

Suzuki, K., Tomozoe, H., Yoshinaga, T., Kaji, M., & Hirayama, K. (2016). Study on observational motion evaluation method of sprinting motion-Examination of evaluation criteria for 5th and 6th graders-. *Physical Education Pedagogy Research (in Japanese)*, 32 (1): 1-20.

Takamoto, E., Ogata, M., & Idei, Y. (2003). Development of running, jumping and throwing movements in elementary school children: For all grades. *Sports Pedagogy Research*, 23 (1): 1-15.

Takizawa, Y., & Tomozoe, H. (2017). Study on observational evaluation criteria for throwing movements-For movements of children in all grades of elementary school-. *Physical Education Pedagogy Research*, 33 (2): 1-17.

Tanaka, R., Murata, Y., & Gushiken, K. (2015). A study on educational significance in pursuit of gymnastic beauty: Based on the causality between violence, beauty, and education. *Bulletin of Nippon Sport Science University*, 45(1): 1-11.