quickly and precisely as possible, not to move the body, and to move the left foot smoothly on the training board. The subjects performed the tasks two times for each course. After finishing each task, the result was not informed of the subject.

We used data from the three courses, with two repetitions for each person. The *ICC*(1,1) of the speed *t* and the accuracy *S* were calculated using the Eq. (1).

$$ICC(1,1) = \frac{BMS - WMS}{BMS + (k-1)WMS} \quad (1)$$

$$BMS = \frac{\sum_{i=1}^{N} k(\bar{x}_i - \bar{x})^2}{N-1}, \quad WMS = \frac{\sum_{i=1}^{N}\sum_{j=1}^{k}(x_{ij} - \bar{x}_i)^2}{N(k-1)}$$

where *BMS* is the variance between subjects, *WMS* is the variance in each subject, *k* is the number of measurements, *N* is the number of subjects, $\bar{x}$ is the mean value of the data from *N* subjects and *k* times, $\bar{x}_i$ is the mean value of the data from *k* times of Subject *i*, and x_{ij} is the data from Subject *i* at time *j*. In this study, $k = 2$ and $N = 14$.

The criterion methods for *ICC*(1,1) have been reported [3][4][6]. In Ref. [3], the numbers were set such that 0.00-0.20: Slight, 0.21-0.40: Fair, 0.41-0.60: Moderate, 0.61-0.80: Substantial, and 0.81-1.00: Almost perfect. On the other hand, in Ref. [4], the numbers were reported such that more than 0.6: Possible, more than 0.7: Normal (OK), more than 0.8: Good, and more than 0.9: Great.

Previous studies showed that if the *ICC*(1,1) is more than 0.7, then the reliability is defined as good.

3.2. *Results and discussion*

The *BMS* and the *WMS* of the speed *t* and accuracy *S* for the three courses are listed in Table 1. From these values, the *ICC*(1,1)s were calculated. The *ICC*(1,1)s of the speed *t* were 0.951, 0.982, and 0.984 for the straight course, the circular course, and the star-shaped course, respectively. The *ICC*(1,1)s of the accuracy *S* were 0.265, 0.924, and 0.711 for the straight course, circular course, and star-shaped course, respectively.

Table 1. Results of *BMS* and *WMS*.

		Three evaluating courses		
		(a) Straight course	(b) Circular course	(c) Star-shaped course
speed *t*	*BMS*	4.01×10^{1}	2.10×10^{2}	2.97×10^{2}
	WMS	1.01×10^{0}	1.95×10^{0}	2.42×10^{0}
accuracy *S*	*BMS*	1.17×10^{7}	1.29×10^{8}	6.01×10^{7}
	WMS	6.79×10^{6}	5.07×10^{6}	1.01×10^{7}

On the basis of these results, the reliabilities of the speed *t* and the accuracy *S* were confirmed for the circular and star-shaped courses. On the other hand, the reliability of the accuracy *S* was not confirmed in the straight course. The reason for this can be determined from Table 1 and Eq. (1). In the case of the accuracy S in the straight course, the ratio of *BMS* and *WMS* was small compared to the other cases. This means that the variance between the subjects was small and the variance in each subject was large. It is probable that the reason for this was that the straight course was easy for the healthy subjects.

As a result, it was confirmed that the developed measurement system has sufficient reliability in both the circular and star-shaped courses.

4. Conclusion

We developed a measurement system to assess the skillfulness of the lower extremities with a straight course, a circular course, and a star-shaped course. The quantitative evaluation method for the lower extremities in both the circular and star-shaped courses has sufficient reliability.

Future work will involve the verification of the reliability of the proposed quantitative evaluation method for elderly persons and subjects with reduced skillfulness of the lower extremities.

References

1. Shrout PE, Fleiss JL, *Intraclass Correlations: Uses in Assessing Rater Reliability*, Psychological Bulletin, **86** (1979) 420-428
2. Yamazaki K, Morita Y, Sato N, Ukai H, Nagaya M, Hayashi T, Ito T , Sakai Y, *Quantitative Evaluation System of Sensorimotor Integration Function and Proprioception of Lower extremities -Verification Effectiveness for surgical patients with intramedullary spinal cord tumor-*, The Japan Society for Welfare Engineering, **15**(2) (2013) 44-49 (Article in Japanese)
3. Landis JR, Koch GG, *The measurement of observer agreement for categorical data*, Biometrics, **33** (1977) 159-174
4. Kuwabara Y, Saito T, Inagaki Y, *Evaluation of intra- and inter-observer reliability*, Kokyu To Junka, **41**(10) (1993) 945-952 (Article in Japanese)
5. Portney LG, Watkins MP, *Foundations lf clinical research-Applications to practice*, Appleton & Lange, USA (1993) 505-516