Paper

Relationships between test characteristics and movement patterns, physical fitness, and measurement characteristics: suggestions for developing new test items for 2- to 6-year-old children

Takahiro IKEDA[*1, *2], Osamu AOYAGI[*3]

*1 Department of Early Childhood Education and Care, Saga Junior College
*2 Graduate School of Health and Sport Science, Fukuoka University
3-18-15, Kamizono, Saga-shi, Saga, 840-0806 Japan
ikeda@saga-jc.ac.jp
*3 Faculty of Health and Sport Science, Fukuoka University

Received November 26, 2007 ; Accepted August 29, 2008

The reliability and validity of any test are important. However, it is difficult to determine both the reliability and validity of tests for young children's motor ability. Therefore, the purpose of the present study was to elucidate issues in motor ability testing in order to develop new test items for young children. First, in order to assess the "Test Characteristics" of motor ability, reliability and validity were determined by administering a motor ability test to 304 young children (boys N=159, girls N=145) aged 2-6 years. Practicality was evaluated by teachers who actually conducted the motor ability tests. Next, the following five factors were extracted using defective principal component analysis with a Normal Varimax rotation: "reliability, validity and practicality", "practicality without familiarity of young children", "difficult for young children and general validity", "acceptance for young children and practicality" and "reflection of young children's daily life, practicality and validity". Finally, the relationship between factor scores and movement patterns, physical fitness and measurement characteristics were examined using Quantification Theory Type I. The test items related to flexibility and using the trunk had high reliably, validity and practicality; however, items related to stability did not. The test items related to power and manipulation had low practicality, but were interesting to the young children. The test items that used upper limbs and were related to coordination and stability were difficult for young children, but had high validity. Items related to agility and combination had the opposite characteristics. Test items related to the trunk, muscular endurance and manipulation were not accepted by the young children and had low practicality. Combination and using the lower limbs and trunk had high validity and practicality and reflected young children's daily life; however, agility showed contrary findings. For the development of new items for testing the motor ability of young children, the present findings are useful because few young children's motor ability tests satisfy all "Test Characteristics".

Key words : Analysis of Test Item, Factor Analysis, Quantification Theory Type I

Human Performance Measurement Vol. 5, 9-22 (2008)

1. Introduction

Generally, the reliability and validity of a test are important (Clarke, 1967; Matsuura, 1983). As young children have no understanding of time, no sense of rivalry (Matsuda, 1961), and do not notice differences in their abilities to perform motor tasks (Inoue, 1968; Ichimura et al., 1969), it is difficult to design a test of motor ability for young children that has sufficient reliability, validity, objectivity and practicality (Aoyagi, 2005). Matsuda (1961), Morris et al. (1985) and Nakamura & Matsuura (1979) conducted research concerning the reliability of tests for young children's motor abilities. Previously, Demura (1993), Matsui et al. (1955) and Takeuchi et al. (1968) identified reliable and valid motor ability tests for young children. Furthermore, Goshi et al. (1998) and Murase & Demura (1990) confirmed the reliability and validity of motor ability tests by demonstrating max

capability and objectivity on pass or fail tests. In terms of the practicality of test items, Ikai (1972) reported that tests for young children must be simple and economical. It is important that young children can perform the tasks safely (Matsuura, 1982); therefore, safe tasks were identified for motor ability tests. Shiba (1972) noted that when all other factors are considered equal, a less reliable item has more validity. Most studies on young children's motor abilities, reliability, validity, objectivity and the practicality of the tests themselves have addressed several of these factors at once; however, few studies have investigated the relationship between reliability, validity and practicality. As has been stated, it is difficult to design a test of motor ability for young children that has sufficient reliability, validity, objectivity and practicality. Therefore, it is necessary to confirm the relationships among these factors.

To date, no integrated test of motor ability has been established, although many researchers have used various tests (Nakamura et al., 1980). For example, the Japanese Ministry of Education, Culture, Sports, Science and Technology developed The Japanese Test of Physical Fitness (Ministry of Education, Culture, Sports, Science and Technology, 2000). However, this test was not employed for young children. Therefore, it is necessary to investigate the issues related to the development of a new test of motor ability for young children.

In this study, reliability, validity and practicality were defined as "Test Characteristics" and the relationships among these "Test Characteristics" were examined. In addition, the relationship between these "Test Characteristics" and movement patterns, physical fitness and "Measurement characteristics" were also examined. Based on our results, the present study clarifies issues related to the development of new items for testing the motor ability of young children in consideration of movement patterns, physical fitness and measurement characteristics.

2. Procedure and materials

2.1. Test items of motor abilities

Testing the motor ability of young children requires "pass or fail" evaluation or performance grading because young children have not yet develop the ability to display their full performance (Matsuura, 1983). However, such tests have problems with objectivity due to large fluctuation among evaluators and the difficulty of establishing clear points of view in order to grade performances.

In contrast, quantitative motor performance tests, which use CGS scales (e.g. time, distance, etc), have been used for generations in order to observe individual or group differences (Goshi et al., 1999). Therefore, the present study used quantitative motor performance tests with CGS scales measured by instruments and scales counting the number of times a task (e.g. Side jumps, Ball bouncing, etc) was performed. However, the objectivity of the tests was not analyzed, as these tests were selected in order to minimize fluctuation among evaluators and provided observers with clear points of view.

The motor ability test used in this study comprised the sphere of movement skills represented by Gallahue's fundamental movement skill model (Gallahue & Donnelly, 2003). The performance in each test item was considered to coincide with the construct validity (Matsuura, 1983) of this model.

The following motor ability tests were used in this study. First, tests that have been used in antecedent studies were selected. For example, "25 m run", "Standing broad jump", "Tennis ball throw", "Timed dipping" and "Stepping-stone jumps" were used in national research conducted by Matsuda & Kondo (1968), Matsuda et al. (1975), Kondo et al. (1987), Kondo et al. (1998) and Sugihara et al. (2004). These test items were compared with national norms for young children's motor abilities published by Sugihara et al. (2004). The "Jump over and crawl under" task (Kurimoto et al., 1981) has been used as a coordination test. The "Sit-and-reach" task (Ministry of Education, Culture, Sports, Science and Technology, 2000) was not employed for young children, but the problems associated with similar items used previously were improved. The "Vertical jump" (Nakamura & Matsuura, 1979), "Side jumps" (Matsui et al., 1974), "Trunk extension" (Nakamura & Matsuura, 1979), "One foot balance" (Matsuda, 1961; Aoyagi et al., 1980), "Stand up and dash" (Aoyagi et al., 1980) and "Arm hang" (Matsuda, 1961) tasks were found to be reliable test items. While the "Jumping with hands", "Potato race" and "Rapid crawling" tasks were found to be valid by Aoyagi & Matsuura (1982). In addition, the "Ball bouncing" task was considered to be reliable and valid by Matsuda (1961) and Takeuchi et al. (1968). In

the present study, these items were measured using CGS scales or scales counting the number of times a task was performed.

Next, the "Balance beam walk" (Aoyagi et al., 1980), "Jumping back and forth" (Katsube & Konishi, 1968), "V-sit" (Aoyagi et al., 1980; Matsuura & Aoyagi, 1985; Goshi & Demura, 1992; Demura et al., 1992) and "Side roll" (Aoyagi et al., 1980) tasks were measured as "Pass or Fail" tests or evaluated using a point system.

Ikeda & Aoyagi (2006) developed new items, including the "Hurdle run", "Throw with both hands", "Ken-ken-pa with small and large circles", "T-ball", "Mini-putt golf", "Kick for distance", "Ball bouncing with a stick", "Rolling hoop" and "Squat balance" tasks based on movements that appear in young children's play activity and have fundamental skills used in sports. In this study, a total of 30 kinds of tests were used. However, as the "Side roll" task was measured for both distance and time, and both small and large circles were used for the "Ken-ken-pa" task, the total number of tests was 32. Table 1 shows the method used for measuring each item. In addition, the "Squat balance", "Timed dipping", "T-ball", "Ball bouncing with a stick" and "Mini-putt golf" task procedures are shown in Figure 1-5.

Table 1. Motor ability test items

No.	Test Items	Unit	Measurement method
1	25 m run	1/10 seconds	Two young children run 30 m together (elapse time to 25 m was recorded).
2	Side jumps	Times	Number of times of side-jumping between 30 cm wide parallel lines.
3	Standing broad jump	Centimeters	Distance of jumping.
4	Throw with both hands	0.5 meters	Distance a ball thrown (overhead with both hands like a soccer throw-in). Dodge ball for school students (is 16 cm⌀) is used.
5	Tennis ball throw	0.5 meters	Distance a tennis ball thrown (overhand throw).
6	Balance beam walk	1/10 seconds	Time to walk between two stands located at opposite ends of a balance beam.
7	Squat balance	Second	Duration of time for a squatted posture (Figure 1).
8	One foot balance	Second	Duration of time for standing on a balance beam. Maximum time: 120 seconds.
9	Timed dipping	Second	Duration of time dipping between two desks about shoulder-width apart. Maximum time: 120 seconds (Figure 2).
10	T-ball	Meter	Distance of hitting a t-ball with a bat. (Figure 3) The ball (7.5 cm ⌀) is made from polyvinyl chloride.
11	Jumping back and forth	Times	Same as item 2, but jumping back and forth.
12	Ball bouncing with a stick	Times	Numbers of times to bounce a ball with a 65cm wooden stick (Figure 4). The same ball is used as item no. 4.
13	Arm hang	Second	Duration of time to hang from a monkey bar. Maximum time: 120 seconds.
14	Hurdle run	1/10 seconds	Duration of time to run and jump a 12 m course with 6 obstacles set at 2 m intervals.
15	V-sit	Second	Time for maintaining a V-figure (posture flexed at the hip). Maximum time: 60 seconds.
16	Mini-putt golf	0.5 meters	Distance obtained when putting towards a 10-meter- target. Maximum record: 6.0 meters (Figure 5). Tennis ball is used.
17	Trunk extension	Centimeters	Distance between the jaw and floor (lying prone and extending trunk).
18	Sit-and-reach	Centimeters	Distance reached while sitting on floor, legs extended, reaching towards toes.
19	Rolling hoop	0.5 meters	Distance reached when a 70 cm ⌀ hoop was rolled.
20	Kick for distance	0.5 meters	Distance reached by a kicked ball. The same ball is used as item no. 4.
21	Ball bouncing	Times	Number of times a ball was bounced in a 2 m ⌀ circle. Maximum record: 50 times. The same ball is used as item no. 4.
22	Rapid crawling	1/10 seconds	Duration of time to crawl (on all fours) to and around a 5 m target.
23	Ken-ken-pa with small circles	1/10 seconds	Duration of time to perform 3 double hops and jumps within 40 cm ⌀ circles.
24	Ken-ken-pa with large circles	1/10 seconds	Same as item 23, but using 70 cm ⌀ circles.
25	Jumping with hands	times/15 seconds	Number of sideway jumps over a 25 cm-high elastic string (stabilized hands on a stand).
26	Jump over and crawl under	1/10 seconds	Duration of time to jump over and crawl under a 30 cm-high elastic string (3 times).
27	Side roll for distance	Centimeters	Distance covered from starting line after side-rolling between 50 cm-wide parallel lines. Maximum distance: 150 cm.
28	Side roll for time	1/10 seconds	Duration of time required to roll 150 cm (action same as item 27).
29	Stepping-stone jumps	1/10 seconds	Duration of time to consecutively jump on 10, 50cm ⌀ circles located in line.
30	Vertical jump	Centimeters	Distance reached when jumping from a standing position.
31	Potato race	1/10 seconds	Duration of time to run and bring three bean bags one by one (3 m away).
32	Stand up and dash	1/10 seconds	Duration of time to stand up from a prone position, run to, run around, and return from a target 3 m away.

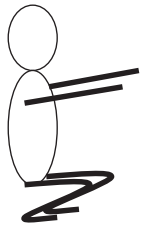

Figure 1. Squat balance
Measurement stopped when: 1) trunk tilted, 2) knees opened, 3) heels came up or 4) ground touched by hand or knee. Maximum time: 120 seconds.

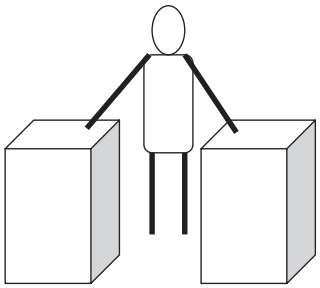

Figure 2. Timed dipping
Measurement stopped when arms flexed or when hands came off the desk.

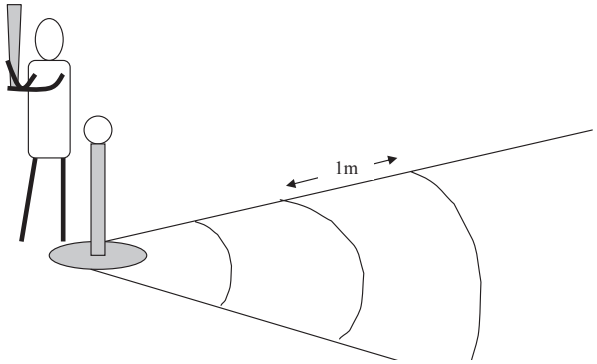

Figure 3. T-ball

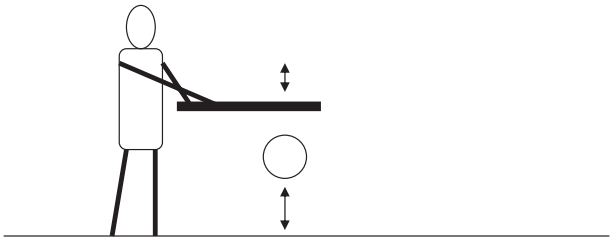

Figure 4. Ball bouncing with stick

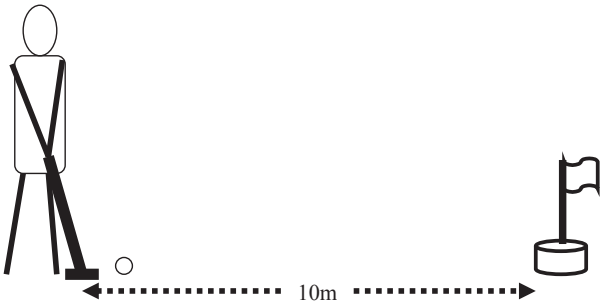

Figure 5. Mini-putt golf

A total of 304 young children in S-kindergarten of S-prefecture participated in this study. The children ranged from 2 to 6 years of age with a mean age of 5.63 and a standard deviation of 0.92. Table 2 shows the subjects by age and gender.

Table 2. Sample size

Gender	Age(yr)				
	2−3	4	5	6	Sub-total
Boys	6	25	62	66	159
Girls	8	31	50	56	145
Sub-total	14	56	112	122	304

S-kindergarten is affiliated with S-college, which trains nursery school teachers. This study was conducted as a part of the operations of the kindergarten. Parents received an explanation that this research would be performed in the kindergarten before the children were enrolled at the kindergarten. Parents provided consent in advance after receiving an explanation of the research protocol and safety measures. Measurements were conducted from March to June 2006. The height and body weight of each student were measured by the kindergarten teachers in March 2006 as part of a regular physical examination for the kindergarten.

2.2. Subjective ratings of young children's motor ability

In order to estimate young children's motor ability, subjective ratings of preschool teachers were used as in previous studies. For example, Demura et al. (1992) showed the structure of motor abilities of young children

using a pass or fail test while Murase et al. (1995) revealed the relationship and differences between observed and estimated values on motor performance. Furthermore, Goshi et al. (1999) reported a relationship between pass-or-fail criteria for teacher's observations and CGS scales in young children. Similarly, Goshi et al. (1991), Goshi and Demura (1992) and Murase et al. (1997) investigated teacher's observations. These previous studies demonstrated the efficacy of subjective ratings for estimating young children's motor abilities.

In this study, it was confirmed that the actual measurement of young children's motor ability is related to evaluations of typical young children's motor ability based on teachers' observations. The subjective rating was done using a five-point scale (1: extremely good to 5: extremely poor). Subjective ratings of young children's motor ability were performed by eleven kindergarten teachers in March 2006. The teachers had a mean of 4.9 years of experience with a standard deviation of 2.7 years.

2.3. Evaluation of the practicality of test items

Questionnaires and actual measurements have been used for the purpose of evaluating the practicality of test items for young children (Harada, 1974; Nakamura et al., 1980). In this study, the practicality of test items was defined as "Test Characteristics", and was evaluated from the following four perspectives: 1) Administration Feasibility (the test was or was not easy to administer), 2) Safety (no injuries or dangerous situations occurred during the testing), 3) Pervading Motor Movement in Daily Life (whether or not a motor task reflected young children's usual motor play patterns/activities), and 4) Familiarity and Acceptance (young children easily understood what to do and how to do it; there was no discomfort or endurance of hardship; interest and motivation were stimulated). Table 3 shows the factors, questionnaire items and alternatives corresponding to the questionnaires.

For Administration Feasibility, a three-point scale was used to evaluate Simplicity, Duration Time, Criteria, Tools and Apparatus, Place, Space and Weather, while Preparation Time and Human Power were rated on a five-point scale. Safety and Pervading Motor Movement in Daily Life did not include any sub-items, and were rated on a three-point scale. Interest was rated on a five-point scale and Understanding, Difficulty, and Discomfort or

Table 3. Criteria for practicality of test items, alternatives of questionnaire items and points

Domain	Sub-domain	Alternatives and points (number equals points)	
Administration feasibility	Simplicity	Was the measurement simple?	1) No. 2) Unknown. 3) Yes.
	Duration time	The duration time required for measurement including instruction and practice was	1) longer than 1 minute. 2) 30 seconds to 1 minute. 3) shorter than 30 seconds.
	Criteria	The criteria to decide was	1) unclear. 2) unknown. 3) clear.
	Tools and apparatus	The tools and apparatus needed for measurement were	1) special tools or apparatus which are not seen in kindergartens and can not be easily obtained. 2) special tools or apparatus which are not seen in kindergartens but can be obtained. 3) special tools or apparatus are not needed.
	Place	Can the measurement be conducted anywhere?	1) No. 2) Unknown. 3) Yes.
	Space	Space for measurement was	1) larger than $25m^2$. 2) $4m^2$ to $25m^2$. 3) smaller than $4\ m^2$.
	Weather	The measurement was	1) influenced by weather. 2) unknown. 3) not influenced by weather.
	Preparation time	Preparation time was	1) longer than 1 hour. 2) 30 minutes to 1 hour. 3) 15 to 20 minutes. 4) 5 to 15 minutes. 5) shorter than 5 minutes.
	Human power	The number of testers needed was	1) one. 2) one, but two was desirable. 3) two. 4) two, but three was desirable. 5) more than three.
Safety		Safety	1) Injury and dangerous situations happened. 2) No injury or dangerous situations happened, but were possible. 3) No injury or dangerous situations happened.
Pervading motor movement in daily life		The motor tasks did or did not include daily motor play?	1) Not included. 2) Unknown. 3) Included
Familiarity and Acceptance	Interest	The interest of the young children was	1) none. 2) weak. 3) unknown. 4) strong. 5) very strong.
	Understanding	The young children	1) hardly understood irrespective of age. 2) hardly understood by lower age groups. 3) easily understood.
	Difficulty	Performing the motor task was	1) difficult irrespective of age. 2) difficult for lower age groups. 3) easy.
	Discomfort or endurance	The young children	1) experienced discomfort or endured hardship. 2) unknown. 3) did not experience discomfort or hardship.

Endurance was rated on a three-point scale. As a result, practicality was evaluated using 9 items for Administration Feasibility, one item each for Safety and Pervading Motor Movement in Daily Life and four items for Familiarity and Acceptance, for a total of 15 items.

Evaluations of the practicality of test items were provided by five kindergarten teachers and a researcher who was involved in the physical education curriculum at the kindergarten. The teachers had a mean of 9.0 years of experience with a standard deviation of 4.6 years while the researcher had ten years of experience.

2.4. Reliability and validity of test items

Generally, reliability was estimated using the test-retest method and Pearson's product moment coefficients (Matsuura, 1983). This method was developed by Baumgartner & Jackson (1982) who used intra-class correlation coefficients to estimate reliability when there were more than two scores. This method can identify changes in mean and standard deviation from one set of measures to the next. In this study, both methods were used in order to evaluate reliability.

In this study, the validity of test items was evaluated according to internal validity (Shiba, 1972) and Criterion-related validity. Internal validity was calculated using the correlation coefficient between the averaged T-scores of all items and the T-score of each item. Criterion-related validity was obtained using the correlation coefficient between each item and age, physique (height and body weight) and subjective ratings of motor ability, which was assessed as significant. Age, physique and motor ability increased and developed with maturity. Thus, the criterion-related validity indicated that each test item measured ability that developed with maturity.

2.5. Movement patterns, physical fitness and measurement characteristics of test items

The 17 items shown in Table 4 were used to determine movement patterns, physical fitness and the measurement characteristics of test items. First, movement patterns were analogous to a test item from among the following: 1) locomotion, 2) manipulation, 3) stability and 4) compound movement. These judgments were made based on Gallahue's model (Gallahue & Donnelly, 2003) in order to ensure validity. Next, the element of physical fitness related to specific test items from among the following: 5) power, 6) muscular endurance, 7) dexterity, 8) agility, 9) flexibility, 10) balance and 11) coordination. These judgments were based on antecedent studies (Matsuura,1983; Murase & Demura,1990; Nakamura et al.,1980; Takeuchi et al,1968) in order to ensure validity.

The part of the body that was mainly used while the movement was being performed was identified from among the following: 12) upper limbs, 13) lower limbs and 14) trunk. The remaining items were evaluated using the following questions: 15) did the test item record the number of times the task was performed? 16) was a larger value considered better? and 17) was the test item performed at full power? These final 3 items were each determined to be applicable or not applicable. To ensure objectivity, these judgments were provided by two researchers who are specialists in the motor ability of young children.

Table 4. Movement pattern, physical fitness and measurement characteristics

Background	Items
Movement pattern	Locomotion movement
	Manipulation movement
	Stability movement
	Compound movement
Physical fitness	Power
	Muscular endurance
	Dexterity
	Agility
	Flexibility
	Balance
	Coordination
Measurement characteristics	Upper limbs
	Lower limbs
	Trunk
	Recorded by the number of times
	Larger value is better
	Full power

2.6. Data analysis

To investigate the general trend of measurement characteristics, factor analysis was applied to 22 items that included Z-transformed variables (see note at end of paper), two reliability coefficients (Pearson's correlation coefficient and intra-class correlation coefficient) and 5 validity coefficients (correlation coefficients for internal validity, age, height, body weight and subjective ratings); a total of 15 items were evaluated for practicality. Factor

analysis was performed using a defective principal component analysis with Normal Varimax criterion.

In order to examine the relationship between each item and factor, factor scores were computed using the regression method.

The relationships between the score for each factor and 'movement patterns, physical fitness and measurement characteristics of each item' were analyzed using Hayashi's Quantification Theory Type I. Using this theory, it was possible to confirm the contribution of each category weight's size (large or small) and value (positive or negative), to the "Test Characteristics" (e.g. reliability or validity).

All data were processed using SPSS 15.0J for Windows and the Quantification Theory program (GUI 2.2.3).

3. Results and discussion

3.1. Reliability, validity and practicality of test items

The reliability and validity coefficients of 32 test items are shown in Table 5. Thirteen product moment correlation coefficients and 21 intra-class correlation coefficients were equal to or greater than 0.8. Twenty-five product moment correlation coefficients and 29 intra-class correlation coefficients were equal to or greater than 0.6. For validity, 30 items showed internal validity, 19 items correlated with age, 20 items correlated with body weight and 19 items correlated with subjective ratings. All 32 items were significant at the alpha level of 0.05. We demonstrated that these test items had sufficient reliability and validity and were considered typical sample items for testing motor ability for young children. Our results are similar to previously published studies (Aoyagi, 1990; Aoyagi et al., 1980; Matsuda, 1961; Matsui et al., 1974; Nakamura & Matsuura, 1979).

To investigate the practicality of the test items, the mean T-score was computed for the following four factors: 1) Administration Feasibility, 2) Safety, 3) Pervading Motor Movement in Daily Life and 4) Familiarity and Acceptance. Twenty items had remarkably high T-scores equal to or greater than 50. The highest score in the twenty items was observed for "Kick for Distance (T-score

Table 5. Reliability and validity of 32 test items

No.	Test Items	Reliability		Validity [1]				
					Criterion-related validity			
		Pearson's correlation coefficients	Intra-class correlation coefficients	Internal validity	Age	Height	Body weight	Subjective rating
1	25m run	.736	.866	.688 **	.053 ns	.291 **	.141 ns	.433 **
2	Side jumps	.675	.712	.457 **	.181 ns	.092 ns	-.069 ns	.167 ns
3	Standing broad jump	.797	.887	.801 **	.062 ns	.204 *	.122 ns	.417 **
4	Throw with both hands	.623	.768	.680 **	.026 ns	.221 *	.272 **	.229 **
5	Tennis ball throw	.800	.887	.737 **	.014 ns	.144 ns	.152 ns	.328 **
6	Balance beam walk	.785	.864	.822 **	.406 *	.419 **	.288 ns	.375 *
7	Squat balance	.823	.902	.693 **	.701 **	.562 **	.468 **	.451 **
8	One foot balance	.550	.686	.505 **	.586 **	.520 **	.342 *	.191 ns
9	Timed dipping	.785	.829	.718 **	.615 **	.540 **	.408 **	.416 **
10	T-ball	.308	.309	.172 ns	.074 ns	.043 ns	-.027 ns	-.072 ns
11	Jumping back and forth	.904	.935	.658 **	.741 **	.493 **	.337 *	.252 ns
12	Ball bouncing with a stick	.841	.734	.568 **	.476 **	.425 **	.325 *	.246 ns
13	Arm hang	.297	.470	.612 **	.166 ns	.010 ns	-.029 ns	.676 **
14	Hurdle run	.893	.945	.819 **	.638 **	.597 **	.413 **	.218 ns
15	V-sit	.749	.826	.678 **	.382 **	.320 *	.244 ns	.281 *
16	Mini-putt golf	.256	.363	.403 **	-.065 ns	-.166 ns	-.185 ns	.259 ns
17	Trunk extension	.941	.979	.554 **	.340 *	.247 ns	.330 ns	.245 ns
18	Sit-and-reach	.835	.932	.531 **	.175 ns	.234 ns	.190 ns	.307 *
19	Rolling hoop	.569	.700	.582 **	.318 *	.357 *	.274 ns	.183 ns
20	Kick for distance	.666	.824	.675 **	.465 **	.458 **	.481 **	.294 *
21	Ball bouncing	.744	.816	.260 ns	.483 **	.264 ns	.145 ns	.069 ns
22	Rapid crawling	.729	.844	.809 **	.671 **	.522 **	.339 ns	.423 *
23	Ken-ken-pa with small circles	.949	.969	.898 **	.620 **	.508 **	.295 ns	.320 ns
24	Ken-ken-pa with large circles	.565	.699	.824 **	.469 **	.435 *	.408 *	.632 **
25	Jumping with hands	.845	.918	.822 **	.751 **	.558 **	.511 *	.616 **
26	Jump over and crawl under	.826	.889	.889 **	.772 **	.673 **	.293 ns	.332 ns
27	Side roll for distance	.685	.795	.475 *	.073 ns	.290 ns	.201 ns	.055 ns
28	Side roll for time	.866	.930	.520 *	.585 **	.651 **	.684 **	-.028 ns
29	Stepping-stone jumps	.749	.829	.856 **	.434 **	.286 *	.134 ns	.571 **
30	Vertical jump	.881	.927	.646 **	.197 ns	.190 ns	.078 ns	.366 **
31	Potato race	.825	.898	.803 **	.072 ns	.178 ns	.232 ns	.367 **
32	Stand up and dash	.481	.636	.787 **	.047 ns	.112 ns	.059 ns	.505 **

1) * $p < .05$, ** $p < .01$ (Validity)

= 57.1)", followed by the scores for "One Foot Balance (56.7)", "25 m Run (55.4)", "Vertical Jump (54.3)" and "Sit-and-Reach (54.1)", respectively. All items had high T-scores (greater than 50) for Safety and Familiarity and Acceptance. In particular, One Foot Balance had T-scores greater than 50 in all four factors.

Even for the 12 items that had total T-scores less than 50, scores of greater than 50 were recorded in at least one of the four factors. There were no items with scores below the mean in all four factors. This data clearly confirmed that no selected test items presented unmanageable practicality issues. In fact, all items exceeded the criteria for practicality.

3.2. Tendency of "Test Characteristics" analyzed by reliability, validity and practicality

Factor analysis was applied to 32 items, including the Z-transformed reliability and validity variables and the items for evaluating practicality.

As a result, five factors were derived that accounted for 70.1% of the total variance. Factor loading, contribution, percent contribution and communalities are shown in Tables 6 and 7, which also show the factor score for each item.

The following features were observed for each factor. The first factor had significant factor loading for Pearson's product moment coefficients and intra-class correlation coefficients, internal validity, criterion related validity [age, height and body weight] and practicality. Thus, the factor was interpreted as "reliability, validity and practicality". Factor scores were greater than 1.0 on "Ken-ken-pa with small circles", "Vertical jumps", "Hurdle run", "Jumping with hands" and "Jump over and crawl under". As "T-ball", "Arm hung", "Mini-putt golf", "Rolling hoop" and "Ball bouncing" were negative scores, these items do not have "reliability, validity and practicality". Takeuchi et al. (1968) showed that "Successive jumping on one leg", which is similar to the "Ken-ken-pa" task in the present study, was reliable for boys. Similar findings were observed in this study.

The second factor had significant factor loading for

Table 6. The factor loading after rotation

Criteria of item analysis			F1	F2	F3	F4	F5	communalities
Reliability	Pearson's correlation coefficients		.841					.800
	Intra-class correlation coefficients		.794					.815
Validity	Internal validity		.583		.612			.760
	Criterion-related validity	age	.532				.591	.779
		height	.579				.704	.889
		body weight	.490				.752	.822
		subjective rating			.780			.659
Administration feasibility	Simplicity					.675		.527
	Duration time						.374	.284
	Criteria				.324	.686		.592
	Tools and apparatus						.697	.551
	Place		.634					.546
	Space			.890				.829
	Weather		.775	.315				.757
	Preparation time		.395	.764				.743
	Human power		.493	.672				.779
Safety					-.752			.660
Pervading motor movement in daily life						.476	.552	.427
Familiarity and Acceptance	Interest			-.832				.721
	Understanding					.694		.569
	Difficulty				-.503	.699		.803
	Discomfort or endurance			-.700	-.416			.746
Contribution			4.15	3.38	2.76	2.64	2.48	15.42
Percent contribution			18.85	15.37	12.56	12.00	11.29	70.07

Note: The factor loadings under 0.300 are excluded.
F1: Reliability, validity and practicality
F2: Practicality without familiarity of young children
F3: Difficult for young children and general validity
F4: Acceptance for young children and practicality
F5: Reflection of young children's daily life, practicality and validity

Table 7. The factor score of 32 items

No.	Test Items	F1	F2	F3	F4	F5
1	25m run		-1.752		1.230	
2	Side jumps					
3	Standing broad jump					
4	Throw with both hands		-1.228			
5	Tennis ball throw		-1.473		1.056	
6	Balance beam walk			1.391		
7	Squat balance		1.088		-3.307	
8	One foot balance		1.401			1.061
9	Timed dipping					1.043
10	T-ball	-1.451	-1.092			-1.302
11	Jumping back and forth					
12	Ball bouncing with a stick				-2.161	
13	Arm hang	-2.686	2.326	2.178		
14	Hurdle run	1.150				
15	V-sit					
16	Mini-putt golf	-1.703	-1.081		-1.464	-2.191
17	Trunk extension		1.535	-1.846	1.758	
18	Sit-and-reach		1.096		1.328	-1.100
19	Rolling hoop	-1.606	-1.172			1.418
20	Kick for distance		-1.698			1.308
21	Ball bouncing	-1.139				1.159
22	Rapid crawling					
23	Ken-ken-pa with small circles	1.723				
24	Ken-ken-pa with large circles			2.073		
25	Jumping with hands	1.106				
26	Jump over and crawl under	1.014				
27	Side roll for distance			-2.251		
28	Side roll for time			-1.434		1.794
29	Stepping-stone jumps			1.478		
30	Vertical jump	1.450				-1.869
31	Potato race					-1.016
32	Stand up and dash					-1.251

Note: The factor scores under 1.000 are excluded.
F1: Reliability, validity and practicality
F2: Practicality without familiarity of young children
F3: Difficult for young children and general validity
F4: Acceptance for young children and practicality
F5: Reflection of young children's daily life, practicality and validity

Practicality, including Space, Weather, Preparation time and Human power, but negative loadings were observed for interest and discomfort or endurance. The factor was interpreted as "practicality without familiarity of young children". Items with factor scores greater than 1.0 included "Squat balance", "One foot balance", "Arm hang", "Trunk extension" and "Sit-and -reach". According to Akimaru et al. (2001) and Nakamura et al. (1980), these items are difficult to motivate young children because showed they are related to resolution. However, these items are simple to measure; therefore it is considered difficult to ensure both usability for the measurer and interest for the young children. In contrast, "25 m run", "Throw with both hands", "Tennis ball throw", "T-ball", "Mini-putt golf", "Rolling hoop" and "Kick for distance" had negative scores; therefore, these items do not have practicality, but are familiar to young children.

The third factor had significant loadings for internal validity, criterion related validity for subjective ratings and measurement criteria, which is necessary for practicality. However, this factor had low loadings for "safety", "Difficulty" and "Discomfort and endurance" for young children. Thus, the third factor was interpreted as "difficult for young children and general validity". The factor scores for "Balance beam walk", "Arm hang", "Ken-ken-pa with large circles" and "Stepping-stone jumping" were greater than 1.0. In contrast, scores for "Side roll for distance", "Side roll for time" and "Trunk extension" were negative.

The fourth factor had good practicality, such as simplicity and criteria, and high acceptance by young children. This analysis demonstrated "acceptance for young children and practicality". High factor scores were observed for "25 m run", "Tennis ball throw", "Trunk

extension" and "Sit-and-reach". Both "25 m run" and "Tennis ball throw" required the movement skills of running and throwing and have been widely used since Matsuda (1961). The items related to flexibility were not interesting for young children, but were simple movements. Conversely, the "Squat balance", "Ball bouncing with stick" and "Mini-putt golf" items showed negative factor scores.

The fifth factor had significant factor loading for Criterion-related validity, age, height and body weight and practicality; "duration time" and "tool and apparatus". Thus, this factor was interpreted as "reflection of young children's daily life, practicality and validity". Factor scores that were greater than 1.0 were observed for "One Foot Balance", "Timed dipping", "Rolling hoop", "Kick for Distance", "Ball bouncing" and "Side roll for time". These items require static balance or body control in a narrow area. As stability, movement and muscular endurance were widely related to resolution for performing movement (Akimaru et al., 2001), these items are considered closely related to not only physique but also mental maturity. In contrast, the items that showed negative factor scores less than -1.0 were "T-ball", "Mini-putt golf" and "Sit-and-reach". Young children do not often play sports such as baseball or golf. Furthermore, Matsuda & Kondo (1968) reported that flexibility, as measured by the "Sit-and -reach" task, did not develop with growth. Therefore, the results of the present study are considered to be valid.

3.3. Relationship between "Test Characteristics" and motor patterns, physical fitness and measurement characteristics of test items

Quantification Theory Type I was performed using factor scores as the dependent variable and motor pattern, physical fitness and measurement characteristics as the independent variables. Multiple correlation coefficients, category weights and partial correlation coefficients were calculated as shown Table 8.

This theory predicted the relationship between each factor and movement pattern, physical fitness and measurement characteristics by category weight and sign (positive or negative) [partial correlation]. The multiple correlation coefficient for the "Factor of reliability, validity and practicality" was 0.798 and this formula had middle relation. The elements with positive relationships were "Locomotion" 0.955 [0.394], "Combination" 0.925 [0.274]

Table 8. Relationship between factor scores and movement pattern, physical fitness and measurement characteristics

	Movement pattern, physical fitness and measurement characteristics	Reliability, validity and practicality		Practicality without familiarity of young children		Difficult for young children and general validity		Acceptance for young children and practicality		Reflection of young children's daily life, practicality and validity	
		Category weight	Partial correlation	Category weight	Partial correlation	Category weight	Partial correlation	Category weight	Partial correlation	Category weight	Partial correlation
1	Locomotion	.955	.394	.529	.321			-.781	.317	-.433	.132
2	Manipulation	.566	.185	-.490	.226			-1.802	.495	-.455	.103
3	Stability	-1.281	.395	.809	.362	.551	.175			-.685	.156
4	Combination	.925	.274	-.375	.162	-.482	.141			.486	.102
5	Power	1.930	.395	-.717	.222						
6	Muscular endurance	3.207	.464					-1.647	.248		
7	Dexterity	-.629	.201	.375	.172			-.930	.277		
8	Agility					-.276	.130	-.258	.121	-.583	.193
9	Flexibility	4.949	.580								
10	Balance	.208	.076	.217	.113						
11	Coordination					.224	.144	-.285	.181		
12	Upper limbs	.331	.249	.227	.244	.592	.403				
13	Lower limbs	.359	.273			.376	.275	-.224	.167	.252	.135
14	Trunk	1.138	.507	-.221	.161	-.452	.219	-.656	.307	.422	.148
15	Recorded by the number of times	2.165	.431			.604	.127				
16	Larger value is better	-.512	.284	.438	.340	-.325	.178				
17	Full power	-.234	.297			.152	.191	.206	.253		
Multiple correlation coefficient			.798		.907		.779		.775		.474

Note: Partial correlations under .100 and their category weights are excluded.

and “Manipulation” 0.566 [0.185] in movement patterns, “Flexibility” 4.949 [0.580], “Muscular endurance” 3.207 [0.464] and “Power” 1.930 [0.395] in physical fitness, “Trunk 1.138” [0.507] and “Recorded by the number of times” 2.165 [0.431]. Therefore, these elements are considered important for establishing reliability, validity and practicality. Takeuchi et al. (1968) reported that test items related to flexibility are high in reliability, validity and practicality in a test battery for young children. The present study supported these previous findings. The elements that showed negative relationships were “Stability” -1.281 [0.395], “Dexterity” -0.629 [0.201], “Full power” -0.234 [0.297] and “Larger value is better” -0.512 [0.284]. These items are considered difficult for determining reliability, validity and practicality.

The multiple correlation coefficient for the “Factor of practicality without familiarity of young children” was 0.907; thus, this regression equation showed a strong relationship with the 17 test items measuring motor patterns, physical fitness and measurement characteristics. Positive category weights [and partial correlation] were observed for “Locomotion” 0.529 [0.321], “Stability” 0.809 [0.362], “Dexterity” 0.375 [0.172], “Balance” 0.217 [0.113] and “Larger value is better” 0.438 [0.340]. However, the elements of “Manipulation” -0.490 [0.226], “Combination” -0.375 [0.162] and “Power” -0.717 [0.222] were negative, thus these elements have problems with practicality, but are familiar to young children.

In the “Factor of difficult for young children and general validity”, the multiple correlation coefficient was 0.779; thus, this regression equation showed a moderate relationship with the 17 test items measuring motor patterns, physical fitness and measurement characteristics. The elements that showed positive relationships were “Stability” 0.551 [0.175], “Coordination” 0.224 [0.144], “Upper limbs” 0.592 [0.403], “Lower limbs” 0.376 [0.275] and “Recorded by the number of times” 0.604 [0.127]. In contrast, “Combination” -0.482 [0.141], “Agility” -0.276 [0.130], “Trunk” -0.452 [0.219] and “Larger value is better” -0.325 [0.178] showed negative scores, indicating that they are easy for young children but have low validity and practicality.

“Factor of acceptance for young children and practicality” showed a multiple correlation coefficient of 0.775. Therefore, this regression equation also showed a moderate relationship with the 17 test items measuring motor patterns, physical fitness and measurement characteristics. A positive relationship was observed only for “Full power” 0.206 [0.253]. Conversely, negative relationships were observed for “Locomotion” -0.781 [0.317], “Manipulation” -1.802 [0.495], “Muscular endurance” -1.647 [0.248], “Dexterity” -0.930 [0.277], “Coordination” -0.285 [0.181], “Agility” -0.258 [0.121] and “Trunk” -0.656 [0.307]. Accordingly, these elements are considered difficult for young children to understand, and for measurers to use.

The last factor was “Reflection of young children’s daily life, practicality and validity”. Positive elements were “Combination” 0.486 [0.102], “Lower limbs” 0.252 [0.135] and “Trunk” 0.422 [0.148]. Negative relationships were observed for “Locomotion” -0.433 [0.132], “Manipulation” -0.455 [0.103], “Stability” -0.685 [0.156] and “Agility” -0.583 [0.193]. However, as the multiple correlation coefficient of this formula was 0.474, it is difficult for movement patterns, physical fitness and measurement characteristics to explain this factor.

Based on the present results, “Test Characteristics” could be assumed by movement patterns, physical fitness and measurement characteristics. For example, 1) items that use the trunk and are related to flexibility may have the potential to have high reliability, validity and practicality 2) items which are related to stability and in which “a larger value is better” will not be able to have high reliability, validity and practicality and 3) items related to manipulation and require power will have low practicality, but young children will be interested in these test items. 4) The items that are related to stability and coordination using upper limbs have the potential for validity and clear criteria, but may have problems associated with safety and difficulty for young children. 5) In contrast, the items that are related to combination with agility will be safe and easy for young children, but have low validity and unclear criteria. 6) The test items related to the trunk, muscular endurance and manipulation will not be accepted by young children and will have low practicality. 7) Combination and using the lower limbs and trunk will have high validity and practicality and reflect young children’s daily life. However, agility may not have high validity and practicality because it does not reflect young children’s daily life.

Murase (2005) note that it is important to consider the validity and reliability of motor ability tests according to

the characteristics of young children. However, due to the characteristics of young children, it is difficult for young children's tests to incorporate all "Test Characteristics", including reliability, validity and practicality (Matsuda, 1961). Therefore, we must select and conduct tests for investigation or education, even if reliability, validity or practicality are sacrificed (Matsuura, 1983). In motor ability tests for young children, it is difficult to satisfy all "Test Characteristics"; therefore, it is important to note the present findings when developing new test items.

Summary

In order to determine "Test Characteristics" for young children's motor ability, 32 motor ability test items were administered to 304 young children (aged 2 to 6 years) and their reliability and validity were examined. In addition, the practicality of items was evaluated by the preschool teachers who measured the young children's motor ability. Five factors were extracted by defective principal component analysis and Normal Varimax rotation, using reliability, validity and practicality. Finally, Quantification Theory Type I was conducted in order to examine the relationship between factor scores, movement patterns, physical fitness and measurement characteristics. The present results can be summarized as follows.

1) Five factors were: "reliability, validity and practicality", "practicality without familiarity of young children", "difficult for young children and general validity", "acceptance for young children and practicality" and "reflection of young children's daily life, practicality and validity".

2) The test items related to flexibility and using the trunk had high reliably, validity and practicality, while items related to stability did not have high reliably, validity or practicality. And the test items related to power and manipulation had low practicality, but were interesting to the young children.

3) The test items that used the upper limbs and were related to coordination and stability were difficult for young children, but had high validity. Items related to agility and combination were easy for young children, but had low validity. And the test items related to the trunk, muscular endurance and manipulation were not accepted by young children and had low practicality.

4) Combination and using the lower limbs and trunk had high validity and practicality and reflected the young children's daily life. However agility had low validity and practicality.

5) When developing new items of testing the motor ability of young children, these findings are important to note because few young children's tests satisfy all "Test Characteristics".

Note) Z-transformation $Z = \frac{1}{2}\log_e \frac{1+r}{1-r}$,

where r is the correlation coefficient.

References

Akimaru, T., Nonaka, H., Hanai, T., Murase, T. and Fujii, K. (2001) [The annual change of physique and motor ability in Aichi prefecture]. Journal of Humanities and Social Sciences, Nagoya City University 11: 127-145. (in Japanese)

Aoyagi, O. (1990) [How to measure motor ability]. Yonetani, M. (ed.) [Health: practice version]. Sankou-shobo: Tokyo, pp.82-86. (in Japanese)

Aoyagi, O. (2005) [Test Theory and Measurement of Physical Fitness]. Toka-shobo: Fukuoka, p.51. (in Japanese)

Aoyagi, O. and Matsuura, Y. (1982) On the structure of motor ability in childhood. Japan Journal of Physical Education, Health and Sport Science 26: 291-303. (in Japanese with English abstract)

Aoyagi, O., Matsuura, Y., Demura, S., Pasau, M. A., Hattori, T. and Tanaka, K. (1980) A factor analytical study on balancing coordination in childhood; For selection of the test items validated for estimating the balancing coordination. Japan Journal of Physical Education, Health and Sport Science 25:197-206. (in Japanese with English abstract)

Baumgartner, T. A. and Jackson, A. S. (1982) Measurement for Evaluation in Physical Education. Wm. C. Company Publishers: Dubuque, pp.91-145.

Clarke, H. H. (1967) Application of Measurement to Health and Physical Education, Fourth Edition. Prentice-hall: Englewood, pp.24-34.

Demura, S. (1993) Contribution of physical fitness and throw form to ball-throw distance and sex difference

in preschool children. Japan Journal of Physical Education, Health and Sport Science 37: 339-350. (in Japanese with English abstract)

Demura, S., Goshi, F., Tatsunami, M. and Katsuki, T. (1992) The factorial structure of motor ability and the developmental trend of these factors in preschool children; by taking motor achievement tests based on pass or fail criteria. Japanese Journal of School Health 34: 229-239. (in Japanese with English abstract)

Gallahue, D. L. and Donnelly, F. C. (2003) Developmental Physical Education for All Children 4th Edition. Human Kinetics: Champaign, p.15.

Goshi, F., Demura, S. and Kasuga, K. (1998) The development of various motor achievement abilities and their relation in infants. Japanese journal of school health 40: 3-18.

Goshi, F. and Demura, S. (1992) The construction of a test battery on motor achievement based on daily behavior observation for preschool children; The use as a screening test. Japan Journal of Physical Education, Health and Sport Science 37: 123-134. (in Japanese with English abstract)

Goshi, F., Demura, S., Kasuga, K., Kobayashi, H., Sato, S. and Minami, M. (1999) Relationship between two types of motor ability tests based on different scales in preschool children: Measurements by pass-or-fail criteria and CGS scales. Japan Journal of Physical Education, Health and Sport Science 44: 345-359. (in Japanese with English abstract)

Goshi, F., Demura, S., Tada, N., Yoshimura, Y. and Nojima, T. (1991) Examination of motor achievement tests based on pass or fail criteria in preschool children: on observed values of motor performance; reliability, objectivity, and agreement with measured values. The Journal of Education and Health Science 36: 263-270. (in Japanese with English abstract)

Harada, S. (1974) [Measuring motor ability for children]. The Journal of Educational Medicine 20 (4): 38-41. (in Japanese)

Ichimura, S., Kamoshita, R. and Ochi, M. (1969) [A study on structure of physical fitness in preschool children]. Japan Journal of Physical Education, Health and Sport Science 13: 235. (in Japanese)

Ikai, M. (1972) [Physical fitness and measurement of children]. Journal of Health, Physical Education and Recreation 20:480-483. (in Japanese)

Ikeda, T. and Aoyagi, O. (2006) A selection of test items for children's physical and motor fitness: classification and investigation of test items based on Gallahue's model of fundamental movement skills. Kyushu Journal of Physical Education and Sports 21 (2): 1-16. (in Japanese with English abstract)

Inoue, K. (1968) [On tests of motor ability for children]. Journal of Health, Physical Education and Recreation 18: 122-126. (in Japanese)

Katsube, A. and Konishi, Y. (1968) [A study of children's motor ability 1]. Bulletin of Nagoya Junior College 8: 30-46.

Kondo, M., Matsuda, I. and Sugihara, T. (1987) [Motor ability on children; The national research in 1986]. Journal of Health, Physical Education and Recreation 37: 551-554. (in Japanese)

Kondo, M., Sugihara, T., Mori, S. and Yoshida, I. (1998) [Motor ability on recent children]. Journal of Health, Physical Education and Recreation 48: 851-859. (in Japanese)

Kurimoto, E., Asami, T., Shibukawa, R., Matsuura, Y. and Katsube, A. (1981) [The final version of the field test of motor coordination by Research Center in Physical Education; The report of Committee of Motor Coordination]. Report of Research Center in Physical Education 9: 207-212. (in Japanese)

Matsuda, I. (1961) Studies on motor development during early childhood. Bulletin of the Faculty of Physical Education, Tokyo University of Education 1: 38-53. (in Japanese with English abstract)

Matsuda, I. and Kondo, M. (1968) A study on the motor abilities of preschool children; a composition of a standard of the development of the motor abilities of preschool children. Bulletin of the Faculty of Physical Education, Tokyo University of Education 7: 33-45. (in Japanese with English abstract)

Matsuda, I., Kondo, M., Sugihara, T. and Minami, S. (1975) [The study on the motor abilities of preschool children and their annual changes]. Bulletin of the Faculty of Physical Education, Tokyo University of Education 14: 31-46. (in Japanese)

Matsui, M., Matsuda, I. and Mori, K. (1955) Studies on motor ability tests for children. Japan Journal of Physical Education, Health and Sport Science 9: 523-533. (in Japanese)

Matsui, H., Katsube, A. and Kajita, K. (1974) A study on construction measuring gross motor coordination (1); Examination on the tests for preschool children. Report of Research Center in Physical Education 2: 290-299. (in Japanese with English abstract)

Matsuura, Y. (1982) [Development of Physical Fitness]. Asakura-shoten: Tokyo, p.47. (in Japanese)

Matsuura, Y. (1983) [Measuring method of Physical Fitness]. Asakura-shoten: Tokyo, pp.58-90. (in Japanese)

Matsuura, Y. and Aoyagi, O. (1985) A study on the classification of growth and developmental curves with metrical procedure; coordination concerned balancing performance in childhood. Bulletin Health and Sports Sciences, University of Tsukuba 8: 193-203. (in Japanese with English abstract)

Ministry of Education, Culture, Sports, Science and Technology (2000) [The Japanese Test of Physical Fitness; for useful utility]. Gyosei: Tokyo. (in Japanese)

Morris, A. M., Williams, J. M., Atwater, A. and Wilmore, J. H. (1982) Age and sex differences in motor performance of 3 through 6 year old children. Research Quarterly for Exercise and Sport 53 (3): 214-221.

Murase, T. (2005) [Scientific Approach of Physical Fitness and Motor Ability on Children; Theory and Practice of Test and Measurement]. NAP: Tokyo, p.82. (in Japanese)

Murase, T. and Demura, S. (1990) The relationship between motor ability and pass or fail tests in preschool children aged 4.0 to 6.5 years. Japan Journal of Physical Education, Health and Sport Science 35: 207-217. (in Japanese with English abstract)

Murase, T., Demura, S., Goshi, F., Kasuga, K. and Ishimura, U. (1995) Relationship and differences between observed and estimated values on motor performance in preschool children aged 4 to 6 years. The Journal of Education and Health Science 41: 195-201. (in Japanese with English abstract)

Murase, T., Demura, S., Kasuga, K. and Goshi, F. (1997) Relationship and differences between observed and estimated results on motor performance on pass-or-fail test in preschool children aged 4 to 6 years. The Journal of Education and Health Science 42: 267-273. (in Japanese with English abstract)

Nakamura, E. and Matsuura, Y. (1979) The development of fundamental motor ability of boys and girls aged four to eight. Japan Journal of Physical Education, Health and Sport Science 24: 127-135. (in Japanese with English abstract)

Nakamura, S., Tsunoda, S., Sugiyama, M., Funato, N., Yoshimura, T. and Miyazaki, M. (1980) A study on teaching physical education in kindergarten, part 5.The Annual Report of Physical Education, Waseda University 12: 16-27. (in Japanese with English abstract)

Shiba, Y. (1972) [Item Analysis]. Hinoda, N. (ed.) [Research Method of Psychology 7, Test 1]. University of Tokyo Press: Tokyo, pp.53-91. (in Japanese)

Sugihara, T., Mori, S. and Yoshida, I. (2004) Annual Change in Preschool Children's Motor Ability and Structural Analysis of Environmental Factors. Report of Grant-in-Aid for Scientific Research B 2002-2003: Tokyo. (in Japanese)

Takeuchi, H., Kawabata, A. and Matsuura, Y. (1968) A study on the motor ability test battery for kindergarten children. Japan Journal of Physical Education, Health and Sport Science 13: 49-57. (in Japanese with English abstract)