Development of Scales for Measuring the In-person Growth of Young Children in Japan, China and Korea

Xiangshan Gao¹, Mai Kominato², Takayuki Umezaki³, Yuichiro Yamagiwa⁴, Hajime Aoyagi⁵, Marie Otomo⁶

ABSTRACT
There are a lot of developmental scales around the world and they are used in different situations. However, these tests are not perfect for two reasons. The first reason is that most developmental scales, including intelligence tests, normally tend to rank the children in the groups of the same age as those in-growth groups. However, an examination to capture in-person growth is necessary. The second reason is that most conventional development tests are created independently by each country, translated and used in another country, so there is the problem of global standards not being satisfied. To cope with these problems, we tried to develop scales for 3 to 6-year-old children in a group of less than 4,000 in Japan, China and Korea. These countries have some common child rearing cultural aspects and declining birthrate problems. In our research, we examined the validity and reliability of the composition concept, created specific items, investigated, and selected practical items for child rearing. The relevance of the composition concept included content aspect, substantive aspect, structural aspect and external aspect.

Keywords: Development of Scales, In-person Growth, Validity, Reliability, Cross Culture Research

1. Introduction

Various developmental tests for infants have been made on the basis of the average value by age in order to grasp the position in the same age group. These tests are widely used to check how well children could adapt at school (Aoyagi et al., 2013). In addition, these scales tend to capture the characteristics of young children in the dimension of "able - not able". However, there is a need to clarify how children grow. Parents and guardians, who are the human environment surrounding the infant, and their caregivers could give appropriate support to the young child more effectively if they know the developmental character of each child. Furthermore, it can be used even when experts provide parenting support. Therefore, an examination to capture the developmental change within a child who is placed in a social and cultural environment such as the home, or nursery is necessary.

Since conventional developmental tests were created in one country and translated and used in other countries, there is the problem that global standards, including regionality and cultural characteristics, are not satisfied. For example, if a developmental...
test made in Europe and the United States is translated directly into each language and used throughout Asia, items that do not fit in each area will be generated. Therefore, our research group selected three countries in East Asia—Japan, China and South Korea—as the subjects to be surveyed. The reason is that these three countries are geographically close and are in a cultural area where parenting is carried out under the strong influence of Confucian thought.

2. Research Goal

Using the global standards in the three East Asian countries, we aim to create a developmental scale to measure the individual growth of young children and use it for child rearing support. For this purpose, this presentation mainly focuses on the validity of the composition concept and the reliability of the scale.

3. Methods

3.1 Preliminary Survey and Preparation of Questionnaires for Japanese Version, Chinese Version and Korean Version

The questions of this scale are based on "TK type infant development test" (Taken Co., 1979). "TK type infant development test" focuses on 5 realms (Health, Relationship, Environment, Language, Expression) to cherish in their childhood and is widely used for young children in Japan in order to capture their developmental aspect and character. This developmental scale has 2 viewpoints; (1) "Viability", such as social skills and competences, that are indispensable in our daily life and (2) "Lifestyle Habits" such as basic habits and acquisition of discipline. "Viability" includes questions about “Physical Activity”, “Verbal skill”, “Group Activity”, “Self-control” and “Spontaneity”. And "Lifestyle Habits” includes questions about “Cleanliness”, “Potty training”, “Grooming”, “Sleep” and “Eating”. Ishikawa et al. (2005), Kominato et al. (2005) and Sone et al. (2005) selected the questions about “Viability” and upon re-examination pointed out that some items may be unsuitable in today’s society. So, in our research, question items with inadequate content validity were deleted. Because "TK type infant development test" was written by Japanese researches, when we made the Chinese version and the Korean version, a psychology major from China and Korea living in Japan translated the questions into the Chinese version and Korean version. Subsequently, we asked psychology majors from China and Korea who live in Japan different from the translators for back translation. When the meaning differed between the Japanese version and the back-translation version, we asked the two translators from each country to work on their translation again and come up with the same meaning. Representatives of the research team visited each country and discussed with each psychologist several times to examine whether the created Chinese version and Korean version are appropriate for each country.

3.2 Investigation and Standardization Process of Globalization

The participants in this study consisted of 1179 Japanese mothers, 1304 Chinese mothers and 1349 Korean mothers with children aged 3-6. The investigation period was...
from January 2006 to April 2007. Mothers evaluated children's development by using the following words: "not able", "sometimes able", "often able" and "completely able". Sampling in Japan used the stratified random two-stage extraction method. To examine the validity of the scale, we referred to the effect size ($\eta^2_p$). Regarding the developmental difference, it was evaluated that the substantial difference in development was small if $\eta^2_p < .02$, evaluated that $\eta^2_p > .07$ was large, and those items were deleted, and 102 items remained (Table 1).

Factor analysis by weightless least squares method was performed on 102 items, and 11 factors were extracted by Promax rotation. However, items with high loads gathered up to the second factor, the number of items constituting the sixth factor to the ninth factor was small, and the factor correlation also exceeded a value of more than .5. In other words, a clear multifactorial structure could not be confirmed by factor analysis. Therefore, we carefully discussed again and decided to adhere to the above 11 factors. Therefore, the structural validity was examined for each subscale by using a reliability coefficient, I - T correlation, and a coefficient according to a convolution method.

Table 1. Scale in the Survey on Children

| Subscales and Items | 1. Fine motor skills | 2. Gross motor skills | 3. Language comprehension | 4. Language expression | 5. Group activity | 6. Play and Social development |
|---------------------|---------------------|---------------------|-----------------------|---------------------|---------------------|---------------------------|
|                     | (1) makes simple buildings with blocks | (1) makes slingshot | (1) speaks clearly | (1) can tell the story of a picture book | (1) can play with others | (1) understands the concept of sharing |
|                     | (2) draws a square properly by using a pencil | (2) jumps on a rope | (2) can express himself/herself fully | (2) can draw the picture book | (2) can play with others | (2) understands the concept of friendship |
|                     | (3) draws a circle properly by using a pencil | (3) runs on a line | (3) can express himself/herself fully | (3) can write a simple sentence | (3) can play with others | (3) understands the concept of cooperation |
|                     | (4) draws a triangle properly by using a pencil | (4) runs on a curve | (4) can express himself/herself fully | (4) can write a simple sentence | (4) can play with others | (4) understands the concept of taking turns |
|                     | (5) draws a diamond shape properly by using a pencil | (5) runs on a curve | (5) can express himself/herself fully | (5) can write a simple sentence | (5) can play with others | (5) understands the concept of taking turns |
|                     | (6) draws a heart properly by using a pencil | (6) runs on a curve | (6) can express himself/herself fully | (6) can write a simple sentence | (6) can play with others | (6) understands the concept of taking turns |
|                     | (7) draws a star properly by using a pencil | (7) runs on a curve | (7) can express himself/herself fully | (7) can write a simple sentence | (7) can play with others | (7) understands the concept of taking turns |
|                     | (8) draws a zig-zag properly by using a pencil | (8) runs on a curve | (8) can express himself/herself fully | (8) can write a simple sentence | (8) can play with others | (8) understands the concept of taking turns |
|                     | (9) draws a cross properly by using a pencil | (9) runs on a curve | (9) can express himself/herself fully | (9) can write a simple sentence | (9) can play with others | (9) understands the concept of taking turns |
|                     | (10) draws a circle properly by using a pencil | (10) runs on a curve | (10) can express himself/herself fully | (10) can write a simple sentence | (10) can play with others | (10) understands the concept of taking turns |

Table 1. Scale in the Survey on Children

| Subscales and Items | 7. Morality | 8. Self-control | 9. Self-motivation | 10. Attention | 11. Help |
|---------------------|------------|---------------|-----------------|-------------|---------|
| (1) can tell people what he/she wants to do (1) can make promises (1) can motivate himself/herself (1) can focus on something (1) can do it even if something goes wrong |
| (2) can make promises (2) can make promises (2) can motivate himself/herself (2) can focus on something (2) can do it even if something goes wrong |
| (3) can make promises (3) can make promises (3) can motivate himself/herself (3) can focus on something (3) can do it even if something goes wrong |
| (4) can make promises (4) can make promises (4) can motivate himself/herself (4) can focus on something (4) can do it even if something goes wrong |
| (5) can make promises (5) can make promises (5) can motivate himself/herself (5) can focus on something (5) can do it even if something goes wrong |

Published by ECSDEV, Via dei Fiori, 34, 00172, Rome, Italy
http://ecsdev.org
4. Results and Discussion

4.1 Consideration of Content Validity and Substantive Validity by Creating Development Regions and Items

We divided the composition of this scale into [Motor development] and [Social development] broadly. [Motor development] summarized in handling movement" and "locomotive movement", then the former was named "fine motor development" and latter was called " gross motor development ". Regarding ' Social development ', adaptation to 'Group activity', acquirement of "Rules in playing and rules of society ", "Morality", "Self-motivation" and "Self-control" are essential for young children, so we added them to the sub-scales.

In addition to the social adjustment, the coping skills for nature and their physical environment are necessary in early childhood, so we added "Attention " to the sub-scales too. Finally, "Chores " which is considered to be important for child development, especially in East Asian society, was added. Altogether, the questionnaires of this study consisting of 11 sub-scales shown in Table 1, which contain12 to 17 questions each totaled 149 items.

The age-specific total score, standard deviation and α coefficient obtained in the preliminary survey are also shown in Table 2. According to the results of this preliminary survey, the score of each region was high, due to the aging. Therefore, it is considered that our scale accurately capture a child’s developmental changes. Moreover, as the coefficient was high on all scales, it can be said that our developmental scale provides high reliability (Aoyagi et al.,2007, Yamagiwa et al.,2007).

| Table 2. The Score of each Subscale( Mean, SD ) and α |
|---------------------------------|
| Subscales                      | Number of items | 3yr  | 4yr  | 5yr  | α     |
| 1. Fine motor skills           | 13              | 34.5(6.3) | 40.3(5.7) | 44.7(4.9) | 0.890 |
| 2. Gross motor skills          | 14              | 34.7(4.8) | 39.1(5.6) | 44.7(5.9) | 0.851 |
| 3. Language comprehension      | 17              | 42.7(7.2) | 49.9(7.4) | 55.9(7.2) | 0.899 |
| 4. Language expression         | 15              | 36.6(7.0) | 43.9(7.1) | 49.9(6.9) | 0.919 |
| 5. Group activity              | 12              | 34.1(5.7) | 38.2(4.8) | 40.8(5.4) | 0.901 |
| 6. Play and Social development | 13              | 36.3(5.4) | 42.5(5.0) | 44.8(4.8) | 0.876 |
| 7. Morality                    | 13              | 32.6(6.4) | 36.5(5.4) | 39.9(5.9) | 0.899 |
| 8. Self-control                | 14              | 35.2(5.4) | 38.1(5.7) | 41.7(5.8) | 0.836 |
| 9. Self-motivation             | 13              | 32.5(6.0) | 37.5(5.8) | 40.7(5.5) | 0.852 |
| 10. Attention                  | 13              | 38.2(5.7) | 41.3(5.2) | 43.3(4.9) | 0.854 |
| 11. Help                       | 12              | 32.2(6.9) | 36.3(6.3) | 39.3(5.8) | 0.885 |

Translation from Yamagiwa et al. (2007)

4.2 Results of Examination of Validity

4.2.1 About Contents Aspect

We considered some questions about sociability and morality, not just about intelligence. In addition, its content validity considererred sufficiency, because in selecting the question items, we regarded the cultural perspective and discussed with great care and depth the usefulness and the attentional points on use. Also, in all
question items, the average value increases as the age. It can be said that “relevance” as a measure of development is sufficient. From the view point of the gender difference, girls tended to score higher with many question items. It has been pointed out that girls develop more quickly than boys, so it may guarantee the validity of this scale.

4.2.2 About Substantive Aspect

It is worth noting that the answers to the question items were given by the infant himself but by a parent who lives with an infant or nursing teacher. It is considered that it is possible for a child's condition to be evaluated more accurately if the observer becomes a respondent. In other words, the answer format by observation is practical and highly relevant.

4.2.3 About the Structural Aspect

The examination of the structural aspect is a verification on whether each subscale is consistently measuring psychological characteristics structurally and consistently, and it also confirms the reliability. We used a semi-seminary method to estimate reliability coefficients in order to verify the consistency of each scale as a whole. As shown in Table 3, the minimum value is 0.733, which is considered to be sufficiently consistent. Table 2 also shows the I-T correlation as to whether the score of each question item and the total score of the subscales are consistent. Only one item of "Self-control" showed an insufficient value, and it was only one out of 102 items. The consistency of the subscale showed that the variance of each question item was small, and the consistency was sufficient because of the high Cronbach's alpha coefficient.

Table3. Subscale Reliability Indicator

| Subscales                        | Number of items | Reliability coefficient of Spearman-Brown | Item-Total Correlation | α   |
|----------------------------------|-----------------|------------------------------------------|------------------------|-----|
| 1. Fine motor skills             | 11              | 0.846                                    | 0.402                  | 0.711 | 0.594 | 0.880 |
| 2. Gross motor skills            | 9               | 0.740                                    | 0.359                  | 0.616 | 0.483 | 0.788 |
| 3. Language comprehension        | 6               | 0.797                                    | 0.405                  | 0.756 | 0.623 | 0.837 |
| 4. Language expression           | 11              | 0.880                                    | 0.486                  | 0.767 | 0.669 | 0.904 |
| 5. Group activity                | 10              | 0.856                                    | 0.555                  | 0.729 | 0.659 | 0.899 |
| 6. Play and Social development   | 7               | 0.792                                    | 0.536                  | 0.633 | 0.567 | 0.876 |
| 7. Morality                      | 11              | 0.827                                    | 0.463                  | 0.683 | 0.573 | 0.880 |
| 8. Self-control                  | 9               | 0.733                                    | 0.273                  | 0.589 | 0.431 | 0.790 |
| 9. Self-motivation               | 10              | 0.806                                    | 0.381                  | 0.593 | 0.487 | 0.828 |
| 10. Attention                    | 9               | 0.797                                    | 0.387                  | 0.668 | 0.528 | 0.847 |
| 11. Help                         | 9               | 0.827                                    | 0.459                  | 0.699 | 0.579 | 0.869 |

Translation from Aoyagi et al. (2013)

4.2.4 About External Aspects

Because nursery teachers are experts of childcare, they are considered to evaluate the child's development objectively as a general child's developmental condition. External validity of this scale was verified by examining the relation with the numerical value evaluated by the childcare provider on each subscale. As shown in Table 4, the average score of evaluation scores by nurseries is higher according to age. However, the number of parents' data was more than 10 times that of the childcare center. Therefore,
we randomly extracted the data of parents and adjusted it to the same degree of data as the nursery teacher and created comparable samples. Therefore, we examined the average value of each subscale of parents and childcare givers. As a result, only for "Morality", the grades by the nursery teacher were significantly higher than that of the guardians (t (135) = 3.38, p < .01), but there was no significant difference in the other subscales. Therefore, since the guardians’ rating is the same level as the objective teachers’ assessments and there is little distortion in evaluation due to the close relationship with their child, it is suggested that our measure fulfills an external validity.

Table 4. The Score by Teachers (Mean and SD)

| Subscales                      | 3yr M | 3yr SD | 4yr M | 4yr SD | 5yr M | 5yr SD |
|--------------------------------|------|-------|------|-------|------|-------|
| 1. Fine motor skills           | 19.5 | 4.2   | 25.5 | 2.2   | 27.7 | 2.7   |
| 2. Gross motor skills          | 16.3 | 3.7   | 20.3 | 2.2   | 23.4 | 1.8   |
| 3. Language comprehension      | 10.5 | 2.3   | 14.2 | 2.1   | 16   | 1.4   |
| 4. Language expression         | 20.3 | 5.5   | 26.5 | 3.6   | 30.1 | 2.3   |
| 5. Group activity              | 21.8 | 4.5   | 26.5 | 2.2   | 28.8 | 1.8   |
| 6. Play and Social development | 15.8 | 3.9   | 19.2 | 1.6   | 20.3 | 1.2   |
| 7. Morality                    | 20.4 | 5.0   | 28.6 | 2.8   | 30.4 | 2.7   |
| 8. Self-control                | 18.1 | 4.2   | 21.7 | 3.7   | 24.1 | 3.7   |
| 9. Self-motivation             | 19.6 | 4.6   | 24.5 | 3.5   | 26.6 | 2.4   |
| 10. Attention                  | 17.9 | 4.0   | 21.9 | 2.8   | 23.5 | 3.7   |
| 11. Help                       | 20.1 | 4.0   | 23.9 | 2.2   | 24.5 | 2.5   |

Translation from Aoyagi et al. (2013)

Conclusion

Through the examinations we carried out, our developmental scale shows sufficient validity. In other words, this scale indicates there is little need to regard the cultural effect, at least in East Asian countries, and may be useful when capturing not only the developmental position of children in their group but their characteristics. To conclude, it is our hope that this scale will be helpful for parents and caregivers who support children’s development. However, it needs to be acknowledged that our research is based on crossing data, and therefore to collect essential validity by collecting longitudinal data will be necessary in the future.

References

Hajime Aoyagi, Yuichiro Yamagiwa, Junko Nakamura, Xiangshan Gao, Takayuki Umezaki, Ikuro Yoshino, Miie Sone, Rie Ishikawa, Mai Kominato (2007). Higashiai jiken Kosodatesshiyono Hattatsusakensano Kihatsuto sonono: Kihon kousoten no Shakudokosei (Development of Developmental Test for Child-rearing in East Asia Area (Part 1) Basic Concept and Scale Construction). Proceedings of the 18th Conference of the Japanese Society of Developmental Psychology, 414

Hajime Aoyagi, Yuichiro Yamagiwa, Aiobao Zhou, Jung Hwan Hyunn, Junko Nakamura (2013). Hattatsusinrinossan kenshiyono Sannai kura Sengu Konan shi no Kosodate shi no oto no (Parenting advice from 3 years old to preschool by development psychologists). Tokyo: Taken publishing Co.
Hambleton, R. K., Merenda, P. F., & Spielberger, C. D. (Eds.) (2005). *Adapting educational and psychological tests for cross-cultural assessment*. New Jersey: Lawrence Erlbaum Associates, Inc.

Rie Ishikawa, Aiobao Zhou, Taikou Kitano, Mie Sone, Mai Kominato, Noriko Sakai, Hajime Aoyagi (2005). Yojino Hattatsu to Hoikushi, Hahaoyano Hattatsu kinikansuru Hikakubunkatekikenkyu (3) Oyano Sutoresuhanno, Nicchukanhikaku (A Comparative Cultural Study on Early Childhood Development and Childcare Professional Development of Mothers (3), Comparison of Japan, China, and Korea). *Proceedings of the 16th Conference of the Japanese Society of Developmental Psychology*, 764.

Mai Kominato, Aiobao Zhou, Taikou Kitano, Mie Sone, Rie Ishikawa, Noriko Sakai, Hajime Aoyagi (2005). Yojino Hattatsu to Hoikushi, Hahaoyano Hattatsukitainikansuru Hikakubunkatekikenkyu (2) Hoikushino Hattatsukitaito Hahaoyano Hattatsu kaitai oyobi sono Nicchukanhikaku (A Comparative Cultural Study on Early Childhood Development and Childcare Professional Development of Mothers (2), Expectation of Nursery Education Development and Mothers' Developmental Expectations and Its Japan, China and Korea Comparison). *Proceedings of the 16th Conference of the Japanese Society of Developmental Psychology*, 763.

Messick, S. (1989) Validity. In R. L. Linn (Ed.), *Educational measurement*. 3rd ed. New York: McMillan. pp.13-103.

Mie Sone, Aiobao Zhou, Taikou Kitano, Mai Kominato, Rie Ishikawa, Noriko Sakai, Hajime Aoyagi (2005). Yojino Hattatsu to Hoikushi, Hahaoyano Hattatsu kinikansuru Hikakubunkatekikenkyu (1) Hahaoyaga Nintisuru Yojino Hattatsu nitsuiteno Nicchukanhikaku (A Comparative Cultural Study on Early Childhood Development and Childcare Professional Development of Mothers (1), Comparison of Japan, China and Korea on the Development of Infants Recognized by Mothers). *Proceedings of the 16th Conference of the Japanese Society of Developmental Psychology*, 762.

Tanaka Educational Research Institute ed. (1979). *TKshiki Yojihattatsu Kensa* (TK type infant development test). Tokyo: Taken publishing Co.

Yuichiro Yamagiwa, Hajime Aoyagi, Junko Nakamura, Xiangshan Gao, Takayuki Umezaki, Ikuro Yoshino, Mie Sone, Rie Ishikawa, Mai Kominato (2007). Higashiajaken Kosodate shienyo Hattatsu kensano Kaihatsu sono 2 Nihondeno Yobichosa (Development of Developmental Test for Child-rearing in East Asia Area (Part 2) Preliminary investigation in Japan). *Proceedings of the 18th Conference of the Japanese Society of Developmental Psychology*, 415

Yuichiro Yamagiwa, Hajime Aoyagi, Junko Nakamura, Xiangshan Gao, Takayuki Umezaki, Mie Sone, Mai Kominato, Marie Sato, Ikuro Yoshino (2008). Higashiajaken Kosodate shienyo Hattatsu kensano Kaihatsu sono3 Shinraiseino Kento (Development of Developmental Test for Child-rearing in East Asia Area (Part 3) Consideration of reliability). *Proceedings of the 19th Conference of the Japanese Society of Developmental Psychology*, 428