JOINT RESEARCH CENTER FOR PANEL STUDIES
SELECTED DISCUSSION PAPER SERIES

SDP2012-006
March, 2013

Japan Child Panel Survey: Reliability and Validity of the Academic Ability Test

Chizuru Shikishima*
Michio Naoi**
Jun Yamashita***
Hideo Akabayashi****

【Abstract】
This study aimed to shed light on how family background is related to children's cognitive ability. In order to carry this out, the Japan Child Panel Survey (JCPS) selected parents who have elementary and junior high school children from among the Japan Household Panel Survey (JHPS) subjects. Subsequently, concomitantly with JHPS follow-ups, the children and their parents were surveyed. In this paper, analysis was conducted on the data pertaining to 461 children from 312 households that were collected during the first survey conducted in 2010. Through this analysis, the reliability and validity of the JCPS academic ability test was verified. The high internal consistency of the item groups belonging to the three factors, arithmetic/mathematics, Japanese language, and reasoning, showed that the JCPS academic ability test has excellent reliability. A comparison with existing academic ability test data by prefecture showed that each item of the JCPS academic ability test has construct validity. The correlation between test scores for arithmetic/mathematics and Japanese language and the grades in both school subjects as reported by the parents showed that the construct measured by the JCPS academic ability test has concurrent validity. The scores for the JCPS academic ability test were validated as an effective index of academic ability.

* Keio Advanced Research Centers, Keio University
** Faculty of Marine Technology, Tokyo University of Marine Science and Technology
*** Keio Advanced Research Centers, Keio University
**** Faculty of Economics, Keio University

Joint Research Center for Panel Studies
Keio University

Japan Child Panel Survey: Reliability and Validity of the Academic Ability Test

Originally published in Japanese in June 2011 as Chapter 2 in Y. Higuchi, T. Miyauchi, C. R. McKenzie, and Keio University Joint Research Center for Panel Studies (eds.), *Dynamism of Education, Health, and Poverty*, Keio University Press, pp.23–48. Translated into English in March 2013.

Chizuru Shikishima (Keio Advanced Research Centers, Keio University)
Michio Naoi (Faculty of Marine Technology, Tokyo University of Marine Science and Technology)
Jun Yamashita (Keio Advanced Research Centers, Keio University)
Hideo Akabayashi (Faculty of Economics, Keio University)

Abstract

This study aimed to shed light on how family background is related to children's cognitive ability. In order to carry this out, the Japan Child Panel Survey (JCPS) selected parents who have elementary and junior high school children from among the Japan Household Panel Survey (JHPS) subjects. Subsequently, concomitantly with JHPS follow-ups, the children and their parents were surveyed. In this paper, analysis was conducted on the data pertaining to 461 children from 312 households that were collected during the first survey conducted in 2010. Through this analysis, the reliability and validity of the JCPS academic ability test was verified. The high internal consistency of the item groups belonging to the three factors, arithmetic/mathematics, Japanese language, and reasoning, showed that the JCPS academic ability test has excellent reliability. A comparison with existing academic ability test data by prefecture showed that each item of the JCPS academic ability test has construct validity. The correlation between test scores for arithmetic/mathematics and Japanese language and the grades in both school subjects as reported by the parents showed that the construct measured by the JCPS academic ability test has concurrent validity. The scores for the JCPS academic ability test were validated as an effective index of academic ability.

Section 1 Introduction[1]

There is increasing interest in the relationship between children's academic ability and family background. Discussions among researchers in the fields of sociology and economics regarding this relationship tend to take the position that children's academic ability is determined by the family's socioeconomic background (e.g., Mimizuka, 2007; Ojima & Below, 2010; Kariya, 2011), and such discussions highlight the issue of the education gap and its reproducibility.

However, in Japan, much of the above discussion has until now relied upon the results of cross-sectional surveys conducted in schools or the results of social survey analyses that target only adults. In other words, the enormous volume of data relating to children's academic ability and the enormous information group relating to family background are not being properly linked together. There is an urgent need to (1) track the same households within the scope of a balanced sample that represents the national population, (2) gather detailed information on both children's academic ability and the families' socioeconomic backgrounds, and thereby (3) provide empirical findings regarding the relationship between the two.

Against such a backdrop, an attempt is being made using a household budget survey targeting adults; that is, a test that actually measures the academic ability of the children from these families is to be incorporated into this family budget. This is the Japan Child Panel Survey (hereunder abbreviated as JCPS). If this design is utilized, not only can it make a static account of the variables measured between the two, but it will also make it possible to shed light on the dynamic correlation between family background and children's academic ability.

The goals of JCPS are as follows: (1) Track the academic ability of the same children and their living conditions and thereby elucidate the causal relationship between family socioeconomic background and children's level of academic achievement. (2) Enable comparisons to be made with US studies by designing a survey similar to the National Longitudinal Survey of Youth (NLSY). (3) Make use of the abundant household information of the Japan Household Panel Survey (JHPS) in

[1] A part of this paper was reported in the Keio University Joint Research Center for Panel Studies' workshop in November 2010. We would like to thank Hiroki Kawai (Faculty of Economics, Keio University) for his valuable comments. Upon the JCPS implementation, we would like to express heartfelt gratitude to Ryosuke Nakamura (Faculty of Economics, Graduate School of Economics, Keio University), Kazuki Kamimura (Faculty of Economics, Graduate School of Economics, Keio University), Sachi Aizawa (Faculty of Economics, Keio University), Takashi Hanazaki (Faculty of Economics, Graduate School of Economics, Keio University), and Hiroko Araki (Faculty of Economics, Graduate School of Economics, Keio University) for all their hard work. Additionally, we would like to thank Yuri Sato (Faculty of Economics, Graduate School of Economics, Keio University) for her help in developing logical reasoning questions. We would finally like to thank Shinji Yamagata (Keio Advanced Research Centers) for advising on the data analysis utilized in the structural equation modeling.

order to provide hitherto unnoticed material that reveals the relationship between family background and children's academic ability.

Hoping to accomplish these goals, this paper focuses on the academic ability test developed originally as a preliminary step before running JCPS with the purpose of measuring children's academic ability. The reliability and validity of the test will then be examined. The test's reliability refers to the degree of stability and consistency in which the test measures the data.[2] The test's validity, on the other hand, is an evaluation of the degree to which the test accurately reflects the construct, which is what the test aims to measure.[3]

The academic ability test must be created to function as an index of academic ability, which is one variable of analysis. To accomplish this, the test must have a high degree of precision and accurately capture the subject to be measured. In other words, the test needs to come with high reliability and validity. In order to confirm the above, this paper shows how the following three examinations were conducted.

The first is an examination of the test items' reliability. Thus, first, the structure of each test item group is checked to determine what type of interrelationship exists among items. The JCPS academic ability test consists of questions on arithmetic/mathematics, Japanese language, and reasoning directed at each school grade from the first year of elementary school through the third year of junior high school. The item groups of arithmetic/mathematics and Japanese language should be considered as sufficiently satisfying content validity, because they have been selected from the questions prepared in the government course guidelines for each of the school subjects. The reasoning item group is also considered to achieve content validity, because logic researchers developed them as logical reasoning questions. However, there is room for more investigation of the structure between constructs that each of the school subjects measures. Even though there is a correlation between the item groups of arithmetic/mathematics, Japanese language, and reasoning, in order to show that different constructs are being measured, it is necessary to show the goodness-of-fit of the three factor model; this model assumes that these item groups receive the respective factor

[2] Reliability coefficients are often used as indices expressing reliability. Reliability coefficients are defined by the ratio of σ^2_T/σ^2_X when test score variance (σ^2_X) is expressed in the sum of true variance (σ^2_T) and error variance (σ^2_E). Estimation methods include test-retest reliability, equivalent-forms reliability, parallel-forms reliability, and internal consistency.

[3] Validity is broadly divided into content validity, criterion-related validity, and construct validity. Content validity refers to the intuitive and experiential judgment by the test producers and other experts regarding how well the measurement scope that the test's content is concerned with is being represented. Criterion-related validity is evaluated through correlation with objective and external criteria, and correlation with other variables that are measured at the same time and involve the same test subjects is particularly referred to as concurrent validity. Construct validity is an essential concept in research and theoretical framework. It is an all-inclusive concept that evaluates the degree to which the construct is being measured.

loading from the constructs expressed by three latent factors—arithmetic/mathematics, Japanese language, and reasoning. In this paper, the test items' factor structure is verified by confirmatory factor analysis making use of structural equation modeling, the model-fitting technique that can assess a model's goodness-of-fit.

The reliability of the item groups based on factor structure is then evaluated from the perspective of the test's internal consistency. It may be said a test has internal consistency when an individual examinee gives the same types of answers to the same types of test questions. It is possible to examine internal consistency of responses using indices that numerically convert Cronbach's alpha.

Second, there is an examination of the construct validity at the academic ability test items level, which is based on a comparison of correct answer rate with existing data. Specifically, two types of methods are used to compare the correct answer rate of the various response items of the JCPS academic ability test to the preliminary survey conducted in advance of the JCPS or the academic ability test by prefecture, from which many of the arithmetic/mathematics and Japanese language questions from elementary school grade 4 to junior high school grade 3 were sourced. One of the methods is a test of the mean difference, that is, an examination of the sample in terms of difference in performance level. Assuming that the large-scale survey conducted at a prefectural level and the academic ability of the JCPS sample both represent the national standard, there should be no difference in terms of correct response rate in the same test items issued for both samples. Another method is a calculation of the correlation coefficient between the correct response rates in the multiple-item groups measured between the two surveys, that is, an examination of the consistency in covariation between the items. If the correlation coefficient is close to 1, this indicates that there is a high degree of consistency in covariation between the items in the two surveys and that the feasibility of assessing the construct validity of test items is high.

Third, there is an examination of the validity of the construct measured in the academic ability test. It is possible to examine the concurrent validity by working out the correlation coefficient between the scores for the JCPS academic ability test and the individual students' school grades From there, the existence of convergent validity can be confirmed if there is a high degree of correlation between the academic ability test scores for arithmetic/mathematics and the school grades for arithmetic/mathematics, as well as the academic ability test scores for Japanese language and the school grades for Japanese language.[4] Furthermore, it is also possible to determine whether

[4] Convergent validity refers to the extent to which scores on two different psychological tests that theoretically should be related are in fact related. Discriminant validity refers to the extent to which scores on two different psychological tests that theoretically should not be related are not in fact related. They may each be considered a type of construct validity evaluation methods.

discriminant validity exists if these correlations are higher than the correlation between the academic ability test scores for arithmetic/mathematics and the grades for Japanese language, or the academic ability test scores for Japanese language and the grades for arithmetic/mathematics.

This paper will examine the reliability and the validity of the JCPS academic ability test from these three angles.

Section 2 Methods

1 Participants

The Japan Household Panel Survey (JHPS) is conducted on approximately 4000 men and women as well as their spouses. These individuals were selected by stratified two-stage random sampling from a sample population consisting of adult men and women aged over 20 who reside throughout Japan. Follow-ups are then conducted on the same households. By doing this, the JHPS aims to shed light on the dynamic changes in economic behavior in Japan (Higuchi et al., 2011). The survey has been continually carried out every year since 2009, and the second survey (called JHPS 2010) was conducted between January and March 2010.

Those targeted in the first JCPS survey, namely JCPS 2010, which was conducted as a part of JHPS 2010, are the parents and children who responded to a request to participate in the survey. In JHPS 2010, those invited to take part in the survey were the JHPS subjects with children in compulsory schooling—either elementary school or junior high school. The survey designates an individual child of elementary school grade 1 to junior high school grade 3 as one unit, and the survey form is comprised of two parts. The first part, answered by the children, is an academic ability test relevant to the children's school grade and is also a questionnaire survey regarding study habits. The other part, answered by the parents who are the JHPS subjects, is a questionnaire survey regarding parenting behaviors and their children's study habits. The parents completed one copy per child.

Of all the subjects participating in JHPS 2010, 644 cases involved subjects with more than one elementary or junior high school child of an age eligible for JCPS, 312 of whom were effective participants for JCPS 2010. Regarding the numbers of participating children among the JHPS subjects, in 180 cases, there was 1 child; in 116 cases, 2 children; in 15 cases, 3 children; and in 1 case, 4 children. Among the children of subjects who participated in JHPS 2010, 959 were elementary or junior high school children of an age eligible for JCPS subjects. Of the 959, 461 were

effective participants for JCPS 2010. The breakdown by school grade is shown in Table 1.[5]

Table 1 Number of Valid Responses to the Academic Ability Test in JCPS 2010

Grade	Elementary School						Junior High School			Total
	1	2	3	4	5	6	1	2	3	
Respondents	62	44	63	47	62	38	57	46	42	461

After collecting the JHPS 2010 survey forms, surveyors visited the households who had consented to participate in the survey to hand the parents and children the JCPS 2010 survey forms.

The participants were given clear instructions on how to complete the academic ability test, which was in the children's survey form. These instructions were written in the survey request form sent to the subjects, as well as in the face sheet of the children's survey form. The instructions asked that the child answer the questions of the academic ability test by him/herself within a time period of 20 minutes, to be measured by the parent or the child him/herself if at least in grades 5 or 6. The instructions also required that the child him/herself should immediately seal the completed survey form in four places using the seal enclosed in the envelope and then hand it to his/her parent. The survey forms for the children and the survey forms for the parents were collected by post using self-addressed envelopes.

2 Test Items

JCPS is modeled on the US NLSY-Child Assessment, which is conducted on children once every two years. The NLSY-Child Assessment was also referenced for the preparation of the JCPS academic ability test.

For academic ability items, the same three subjects of arithmetic/mathematics, Japanese language, and reasoning were used from elementary school grade 1 to junior high school grade 3. This follows the example of the NLSY-Child Assessment, which since 1986 has conducted academic ability tests on 5- to 14-year-olds using the Peabody Individual Achievement Test (PIAT) to measure math, reading recognition, and comprehension as well as intelligence tests that measure short-term memory (digit span) from WISC-R. Unlike the NLSY, which uses interviewing, the JCPS is restricted due to the use of the mailing method; despite this, it is forced to rely on self-administration by families. Overall, the survey process was streamlined due to relative ease of scoring and

[5] This includes 4 children whose participation in JCPS 2010 cannot be confirmed. Their answers may therefore be based on an extraction error. (For example, the child's brother may have answered questions that should originally have been answered by a JCPS subject.)

subsequent feasibility of implementation.

The arithmetic/mathematics questions consisted of calculations and questions expressed in words concerning numbers and the manipulation of figures. The Japanese language questions consisted of vocabulary, in addition to the reading and writing of kanji characters. The intelligence test consisted of basic questions concerning logical reasoning. For each school grade, the volume of questions was set so as to allow answers to be completed within 20 minutes.

For arithmetic/mathematics and Japanese language, different sets of questions were prepared for each school grade in line with the government course guidelines. For elementary grades 1 to 3, a company that prepares learning drills was asked to prepare original questions. The questions used for elementary school grade 4 to junior high school grade 3 were selected from the academic ability tests by prefecture, which were independently carried out in prefectures. The researchers were able to view the questions for the academic ability tests by prefecture and their respective correct response rates by asking self-governing bodies to release the information. More specifically, we were able to access the questions set by Akita Prefecture in 2005–2008 for students in elementary school grade 4 to junior high grade 3, the questions set by Niigata Prefecture in 2004 and 2006 for students in elementary school grade 4 to junior high school grade 2, the questions set by Gifu Prefecture in 2003–2008 for students in elementary school grade 5 to junior high school grade 2, and the questions set by Kagawa Prefecture in 2006-2009 for students in elementary school grade 3 to junior high school grade 3. Consequently, a pool of arithmetic/mathematics and Japanese language questions for each school grade was prepared, and questions considered appropriate were selected from this pool.

For the reasoning questions, original logical reasoning questions were developed with the cooperation of the logics researchers, and they were used as an alternative index for intelligence measurement. There are various views on the definition of intelligence, but it is widely agreed among intelligence researchers that reasoning ability is central to intelligence (Gottfredson, 1997; Snyderman & Rothman, 1988). Several empirical studies in psychology have demonstrated that general intelligence is deeply related in particular to a deductive logical reasoning ability reflected in a syllogism (Stanovich & West,1998; Shikishima et al., 2009). Syllogisms are a form of logical deductive argument, relating three terms that consist of two premises and a conclusion. Incorporating multiple syllogism questions into a self-administered questionnaire can reportedly estimate intelligence (Shikishima et al., 2011). It is known that syllogistic reasoning ability can be effectively measured when children reach age nine (Bara et al., 1995); therefore, it was decided in the JCPS to introduce syllogistic logical reasoning questions for students in elementary school grades 4 and above. Common multiple-choice questions were then prepared for students in

elementary school grade 4 up to junior high school grade 3; however, for students in elementary school grade 4 up to grade 6, figures and symbols were included in each question to assist students in answering them. While not syllogistic, four logical reasoning questions were set for students in elementary school grade 1 up to grade 3. The students were given two statements representing the relationship between three terms, and they were required to deductively infer their mutual relationship and select the correct answer from the options provided.

The test questions prepared for students in elementary school grade 1 up to junior high school grade 3 were implemented as a preliminary survey in December 2009 with the cooperation of one elementary school and one junior high school in Shizuoka Prefecture. There were 419 effective responses in the preliminary survey. The breakdown of the numbers for each school grade is shown in Table 2.

Table 2 Number of Valid Responses to the Preliminary Survey

Grade	Elementary School						Junior High School			Total
	1	2	3	4	5	6	1	2	3	
Respondents	50	64	67	53	34	61	26	32	32	419
Classes	2	2	2	2	1	2	1	1	1	14

The goal of the preliminary survey was to conduct the prepared test in a real school classroom and thereby verify the difficulty level of the questions, the suitability of the question setting format, and the suitability of the volume and duration. As for the data collected from each school grade, the correct answer rate for each item was calculated, and then items with conspicuously high correct answer rates and items with an extremely small dispersion were removed. The distribution of simple calculation scores, whereby each item was scored with 1 point for a correct answer and 0 points for an incorrect answer, was then investigated. For items with a ceiling effect, various measures were implemented in an attempt to improve the difficulty level. These measures included adding revisions to the questions or introducing new questions for which the answer rates publicized in the academic ability tests by prefecture were lower. The items that had a low internal consistency in each school grade and each subject were also removed. Based on the information obtained from every participant about the time required to answer the academic ability test questions, the volume of questions for each school grade was adjusted to a level that could be judged appropriate for answering within 20 minutes. Furthermore, based on the comments of the site teachers, revisions were made to the wording of the questions in order to make them appropriate for the relevant school grades. Having undergone these amendments, the question items for the JCPS academic test were decided upon.

3 Statistical Analyses

(1) Method for Examining Reliability

The reliability of the JCPS academic ability test was examined through the following four methods.

1. Tetrachoric Correlation

For the examination of the test's reliability, tetrachoric correlation (fourfold point correlation coefficient) was introduced. Tetrachoric correlation is a 2 × 2 correlation that uses the maximum likelihood method to compute the covariance between two discrete variables that both lie on the ordinal scale and accompany category numbers (Greene & Hensher, 2010). Tetrachoric correlation assumes that the binaries have a normal distribution of latent variables, and it adjusts the distance threshold between the two response categories. For example, in the case of a disease, it is a useful method to state the correlation between the variables that accompany the two categories of an infected group and a healthy group, or in the case of the academic ability test, the correlation between the variables that accompany the correct/incorrect data.

In this paper, the items that have been categorized into the same categories in every case for each school grade and therefore have no variance were removed (as everyone answered either correctly or incorrectly). Furthermore, when every case showed a response category equivalent to 2 items or more and the correlation between items was 1, then the first item was left and the other item(s) was/were removed. After having completed the above, all remaining items for each school grade were used to calculate the tetrachoric correlation matrix.

2. Categorical Factor Analysis

Categorical factor analysis is a sub-model of structural equation modeling (Bartholomew, 1980; Toyoda, 1998). One-factor categorical factor analysis, which analyzes binary data, is expressed in the following formula.

$$z_{ij} = \alpha_j f_i + e_{ij}$$

In this case, α_j represents the factor pattern (factor loading) for item $_j$. Next, f_i represents the common factor for participant $_i$. Then, e_{ij} represents the error factor of item $_j$ for participant $_i$ In normal factor analysis, z_{ij} represents the observed variable, but in categorical factor analysis, it represents the latent variable. When γ_j is deemed item $_j$'s threshold, and the actually measured binary variable of 0 or 1 is deemed u_{ij}, it is observed that $u_{ij} = 1$ if $z_{ij} > \gamma_j$, and $u_{ij} = 0$ if $z_{ij} < \gamma_j$. The estimated parameter using the maximum likelihood method is the factor pattern α_j and threshold γ_j (j = 1, …, n).

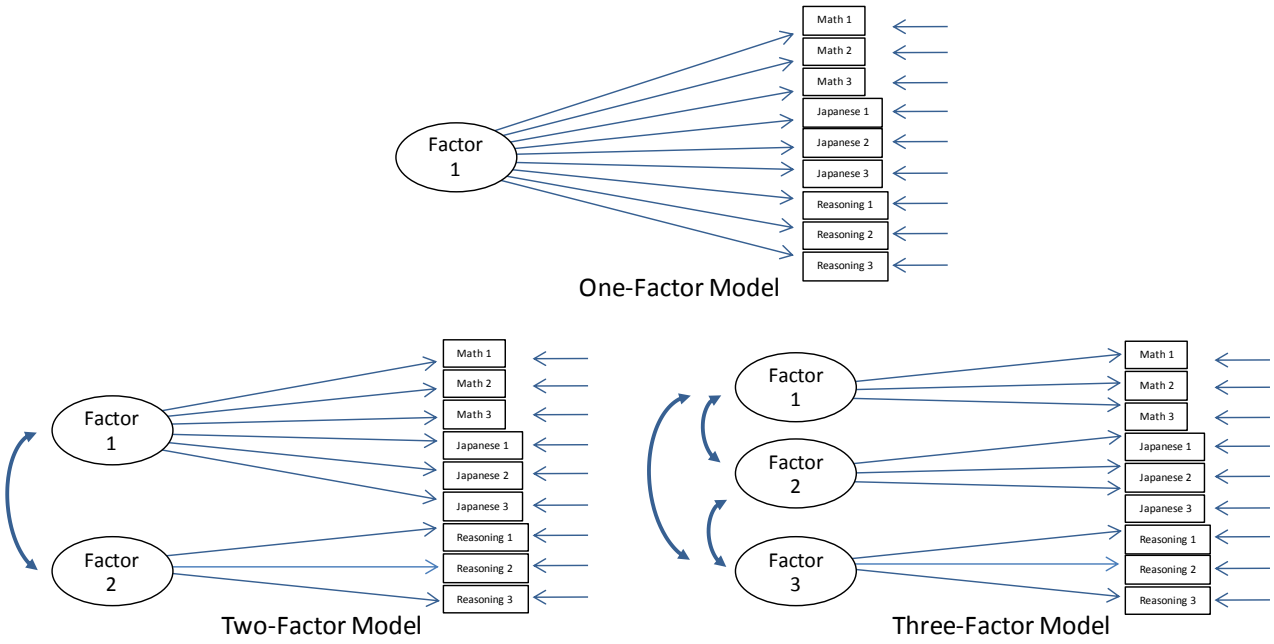
Categorical factor analysis of binary data based on structural equation modeling is known to correspond to the two-parameter item response theory model (Hambleton & Swaminathan, 1985).

In this paper, because the questions vary between school grades, estimates of factor pattern and threshold for each item concerning arithmetic/mathematics and Japanese language were worked out for each school grade and for each item. For reasoning, however, because common questions are set for elementary school grade 1 to grade 3, elementary school grade 4 to grade 6, and junior high school grade 1 to grade 3, estimates of factor pattern and threshold were worked out for each item by combining sets of data for three school grades.

3. Confirmatory Factor Analysis

Confirmatory factor analysis incorporates information concerning parameters that are known in advance, assumes a hypothetical structure, and then specifies a model from substantial scientific findings (Thompson, 2004). In other words, confirmatory factor analysis is a method that confirms the fitting of the model, having determined the factor structure from the beginning. It may be considered as an analysis in which a part of the parameter of the normal (exploratory) factor analysis model that calculates the factors in an explorative way was fixed as a theoretically predictable value.

Figure 1 Three Confirmatory Factor Analysis Models



In this paper, in order to examine the structure of each item group, three factor models were compared with regard to the item groups of arithmetic/mathematics, Japanese language, and reasoning (Figure 1). The three models are as follows: (i) One-factor model: explains all items in one

dimension; (ii) Two-factor model: combines the items of arithmetic/mathematics and Japanese language in one dimension and explains the reasoning items through another dimension that has correlation with them; and (iii) Three-factor model: explains the respective items of arithmetic/mathematics, Japanese language, and reasoning in three three-dimensions that correlate with each other.

The comparison of the models is based on model fit. A comparison based on RMSEA (root mean square error of approximation) and CFI (comparative fit index) was used together with a chi-squared test for evaluation of the models. RMSEA is an index that expresses the divergence between the models' distribution and true distribution per 1 degree of freedom. A model is judged to have a good fit if the degree of freedom is at 0.05 or below, and it is judged to have a poor fit if it is at 0.1 or above (Toyota, 1998). Compared to independent models that assume no correlation between the observed variables, CFI evaluates the degree of improvement in the model's goodness-of-fit. The values range from 0.0 to 1.0 with the model's goodness-of-fit judged better the closer it is to 1.0.

4. Cronbach's Alpha Coefficient

For examining the reliability of tests, Cronbach's alpha coefficient (Cronbach, 1951) is used the most frequently, particularly as an index of the reliability coefficient that estimates the internal consistency of item groups in a scale. Cronbach's alpha coefficient computes according to the following formula when N=number of items and $\bar{p}$=the mean correlation coefficient between items.

$$\alpha = \frac{N\bar{p}}{1+\bar{p}(N-1)}$$

If the alpha coefficient is close to 1, then the stability of the answers in the test is considered to have a high degree of consistency, and it is then possible to evaluate that the test has high reliability. A coefficient of around 0.8 is normally considered to represent sufficient consistency; even a coefficient of around 0.7 is considered to have a certain degree of consistency. It is known that, generally, the larger the number of items, the higher the alpha coefficient (Nunnaly, 1978).

In this paper, the tetrachoric correlation matrix formed from the item groups of each school grade was used to calculate the mean of the correlation coefficient, and the Cronbach's alpha coefficient was then worked out for each factor. The alpha coefficient for arithmetic/mathematics and Japanese language was calculated for each school grade. However, for reasoning, since common questions were set to elementary school grade 1 to grade 3, elementary school grade 4 to grade 6, and junior high school grade 1 to grade 3, the alpha coefficient was calculated using sets of data for three school grades.

(2) Method for Examining Validity

In order to examine the validity of the test items of the JCPS academic ability test, a comparative analysis of the mean values in the correct answer rate was carried out for each corresponding item between the JCPS academic ability test and the preliminary survey, as well as then between the JCPS and academic ability tests by prefecture.

Because the JCPS was supplemented with items to reflect the results of the preliminary survey, some of the questions are not included in the preliminary survey (five items of Japanese language for elementary school grade 1, one item of arithmetic for elementary school grade 2, two items of arithmetic for elementary school grade 3, one item of mathematics for junior high school grade 1, two items of mathematics and four items of Japanese language for junior high school grade 2, four items of mathematics and four items of Japanese language for junior high school grade 3, and one item of reasoning for elementary school grade 4 to junior high school grade 3; see Table 9). For common items other than these items, each of the correct answer rates were set as cases, and paired *t*-tests were conducted between the two surveys.

The same statistical test was also conducted between the JCPS and the academic ability tests by prefecture. In both the JCPS and the prefectural academic ability tests, the same arithmetic/mathematics and Japanese language questions were used for elementary school grade 4 to junior high school grade 3 (excluding one item of arithmetic for elementary school grade 6, four items of mathematics and two items of Japanese language for junior high school grade 2, and four items of mathematics and two items of Japanese language for junior high school grade 3; see Table 9). The prefectures from which the JCPS academic ability test questions were sourced, which were Akita Prefecture, Niigata Prefecture, Gifu Prefecture, and Kagawa Prefecture, cited the correct answer rates of the items for academic ability test by prefecture that correspond with the JCPS and then compared the mean values of correct answer rates between the two surveys. Two cases were removed from the analysis. In the first case, the period for implementing academic ability tests by prefecture for each grade differed from that of the JCPS by more than one year (19 items); in the second case, the questions were judged to be incomparable after having undergone revision (six items) (see Table 9).

After this, with the respective correct answer rates for the corresponding items mentioned above set as case, the correlation coefficient between the JCPS and the preliminary survey and between the JCPS and the academic ability tests by prefecture was worked out using Pearson's product-moment correlation coefficient. Regarding the academic ability tests by prefecture, comparison of the respective mean values and the correlation coefficients were calculated for all of the four prefectures.

In order to examine the validity of the construct measured in the test items of

arithmetic/mathematics and Japanese language, the children's school results in arithmetic/mathematics and Japanese language at the time they answered the questions were used as another index measuring the same construct. The JCPS inquired into the children's results by asking parents to answer on a five-point scale from "top grades" to "bottom grades." The degree of correlation between the factor scores for arithmetic/mathematics and Japanese language in each school grade and the school results in the two school subjects as reported by the parents was worked out using Pearson's product-moment correlation coefficient.

The statistical software Mplus 6.1, which is optimized for structural equation modeling, was used for the tetrachoric correlation coefficient, the calculation of each item's threshold value and factor pattern, and confirmatory categorical factor analysis. For other analyses, SPSS 18.0 was used.

Section 3 Results

1 The Factor Construction and Internal Consistency of the Test Items

Categorical factor analysis was conducted with the structural equation modeling's confirmatory factor analysis framework. Table 3 shows, with respect to the nine school grades of elementary school grade 1 to junior high school grade 3, the results of each grade's test items fitted with the one-factor model, two-factor model, and three-factor model.

The RMSEA values for elementary school grade 2, elementary school grade 6, junior high school grade 2, and junior high school grade 3 were lowest in the three-factor model. This revealed that the three-factor model has a better goodness-of-fit than the two-factor and one-factor models (Table 3). The one-factor and two-factor models had similar RMSEA values for elementary school grades 1, 3, 4, and 5, and the chi-squared test showed that the goodness-of-fit of the two-factor model was not significantly worse than that of the three-factor model. However, the three-factor model had slightly better CFI values for all three school grades. For elementary school grades 1, 2, 3, and 5 and junior high school grade 3, both the RMSEA and the CFI were formidable at below 0.03 and above 0.97, respectively. This showed that the model has an extremely impressive goodness-of-fit.

Table 3　Confirmatory Factor Analysis Model-Fitting

		Elementary School						Junior High School		
		1	2	3	4	5	6	1	2	3
Three-Factor Model	RMSEA	**0.02**	**0.01**	**0.01**	**0.04**	**0.02**	**0.07**	**0.06**	**0.06**	**0.03**
	CFI	**0.98**	**0.97**	**0.99**	**0.87**	**0.98**	**0.88**	**0.72**	**0.72**	**0.99**
Two-Factor Model	RMSEA	**0.02**	0.02	**0.01**	**0.04**	**0.02**	0.08	0.08	0.07	0.09
	CFI	**0.97**	0.90	**0.98**	**0.86**	**0.97**	0.87	0.50	0.62	0.86
	$\Delta\chi 2$	**2.13**	9.51	**4.79**	**3.76**	**3.14**	7.00	52.87	22.67	122.09
	Δdf	**2**	2	**2**	**2**	**2**	2	2	2	2
	p	***NS***	***	*****	***NS***	***NS***	**	***	***	***
One-Factor Model	RMSEA	0.04	0.03	0.01	0.05	0.04	0.10	0.08	0.07	0.09
	CFI	0.91	0.89	0.96	0.79	0.89	0.78	0.50	0.60	0.86
	$\Delta\chi 2$	32.12	11.89	8.58	18.79	24.08	56.15	55.23	28.18	122.47
	Δdf	3	3	3	3	3	3	3	3	3
	p	***	***	**	***	***	***	***	***	***

Note. ***, ** and * indicate that the coefficients are statistically significant at standards of 0.1%, 1%, and 5%, respectively. Bold font indicates the model with the best fit.

The factor pattern and threshold values of each estimated item are shown in Table 4 and Table 5. The cases in which the factor pattern was small but negative were observed for each one item of Japanese language in junior high school grades 2 and 3, but the load to the items was generally high and even.

Table 4　Factor Pattern and Threshold Values of the Items of Arithmetic/Mathematics and Japanese Language

Grade	Elemenatary School															Junior High School								
	1			2			3			4			5			6			1			2			3		
	Item	Factor Pattern	Threshold	Item	Factor Pattern	Threshold	Item	Factor Pattern	Threshold	Item	Factor Pattern	Threshold	Item	Factor Pattern	Threshold	Item	Factor Pattern	Threshold	Item	Factor Pattern	Threshold	Item	Factor Pattern	Threshold	Item	Factor Pattern	Threshold
Arithmetic/ Mathematics	1(1)	0.78	-1.40	1(1)	0.81	-1.69	1	0.95	-1.86	1(1)	0.88	-0.80	1(1)	0.74	-2.14	1(1)	0.77	-1.12	1(1)	0.87	-0.69	1(1)	0.69	-0.71	1	0.83	-0.30
	1(2)	0.69	-1.21	1(2)	0.70	-0.91	2	0.14	-1.22	1(2)	0.60	-0.73	1(2)	0.19	-1.21	1(2)	0.76	-1.12	1(2)	0.63	-0.93	1(2)	0.78	-1.23	2(1)	0.68	-0.18
	2(1)	0.88	-1.85	2(1)	-	-	3(1)	0.15	-1.22	1(3)	0.65	-1.72	1(3)	0.78	-0.60	1(3)	0.81	-0.56	1(3)	0.95	-0.93	1(3)	0.71	-0.58	2(2)	0.91	0.12
	2(2)	0.64	-1.30	2(2)	-	-	3(2)	0.28	-1.14	1(4)	0.64	-1.37	1(4)	0.80	-0.55	1(4)	0.73	0.81	1(4)	0.83	-0.58	1(4)	0.79	-0.33	3(1)	0.91	-0.57
	3(1)	0.60	-1.40	3(1)	0.45	-1.21	4	0.67	-1.14	2(1)	0.42	-0.53	2(1)	0.24	-0.70	2	0.80	-0.34	2	0.58	-0.63	2(1)	0.58	-0.11	3(2)	0.81	-0.06
	3(2)	0.52	-0.29	3(2)	0.61	-0.83	5(1)	0.70	-1.53	2(2)	0.59	0.24	2(2)	0.78	-0.81	3(1)	0.92	-0.20	3(1)	0.74	-0.63	2(2)	0.72	0.06	3(3)	0.81	-1.18
	4(1)	0.85	-1.21	4	0.53	-1.10	5(2)	0.30	0.10	2(3)	0.58	-1.25	3	0.68	-0.04	3(2)	0.73	-0.56	3(2)	0.66	-0.29	3(1)	0.46	-0.51	4(1)	0.85	-0.12
	4(2)	0.67	-0.81	5	0.68	-1.21	6_1	0.42	-0.39	2(4)	0.89	-0.95	4(1)	0.82	-0.65	4(1)	0.76	0.41	3(3)	0.42	0.38	3(2)	0.81	-0.71	4(2)	0.98	-0.37
	5(1)	0.75	-2.14	6	0.62	-1.34	6_2	-	-	3(1)	0.45	-0.53	4(2)	0.81	-0.99	4(2)	0.90	0.13	4(1)	0.42	-0.20	4(1)	0.86	-0.86	5(1)	1.00	-0.43
	5(2)	0.68	-1.52	7(1)	0.92	-2.00	7	0.62	-1.86	3(2)	0.93	-1.14				4(3)	0.43	-0.63	4(2)	0.40	0.38	4(2)	0.41	-0.06	5(2)	0.85	0.00
	5(3)	0.79	-1.40	7(2)	0.92	-2.00	8	0.40	-1.41																		
	5(4)	0.60	-1.06	7(3)	0.91	-1.69	9(1)	0.95	-1.86																		
	6(1)	0.49	-0.75	7(4)	0.92	-2.00	9(2)	0.95	-1.86																		
	6(2)	0.49	-0.60	8(1)	0.16	-1.34	9(3)	0.36	-1.41																		
				8(2)	0.40	-1.21	9(4)	0.45	-1.14																		
				9_1	0.03	-1.21	9(5)	0.67	-1.31																		
				9_2	0.43	-0.41	9(6)	0.86	-1.22																		
							10(1)	0.88	-1.67																		
							10(2)	0.52	-0.77																		
Japanese Language	7(1)	0.91	-1.30	10(1)	0.51	-0.67	11(1)	0.69	-1.86	4(1)	0.73	-2.03	5(1)	0.64	-1.66	5(1)	0.84	-0.72	5(1)	0.74	-1.08	5(1)	0.39	-0.64	6(1)	0.71	-1.07
	7(2)	0.96	-1.21	10(2)	0.25	-1.10	11(2)	0.83	-1.53	4(2)	0.73	-2.03	5(2)	0.84	-0.75	5(2)	0.78	-1.00	5(2)	0.46	-1.36	5(2)	0.27	-1.71	6(2)	0.77	-0.88
	7(3)	0.91	-1.52	10(3)	-	-	11(3)	0.90	-2.15	4(3)	0.95	-1.04	5(3)	0.52	0.16	5(3)	0.84	-0.56	5(3)	0.90	-0.93	5(3)	0.63	-0.51	6(3)	0.33	-1.98
	8(1)	-	-	11(1)	0.57	-1.00	11(4)	0.75	-1.07	5(1)	0.55	-1.25	6(1)	0.69	0.16	5(4)	0.45	-0.27	5(4)	0.65	-0.16	6(1)	0.73	-1.03	7(1)	0.36	-0.64
	8(2)	-	-	11(2)	0.58	-0.83	12(1)	0.25	-0.82	5(2)	0.79	-0.95	6(2)	0.54	-0.70	6(1)	0.77	-0.34	6(1)	0.76	-0.87	6(2)	0.68	-0.86	7(2)	0.71	-0.79
	8(3)	0.85	-1.85	12(1)	0.69	-1.00	12(2)	0.63	-0.57	5(3)	0.72	-0.08	6(3)	0.87	-1.52	6(2)	0.63	-0.41	6(2)	0.79	-1.25	6(3)	0.69	0.06	7(3)	0.82	-1.31
	8(4)	0.95	-1.66	12(2)	0.77	-1.69	12(3)	0.73	-1.00	5(4)	0.49	-0.19	6(4)	0.90	-1.30	6(3)	0.89	-0.56	6(3)	0.94	-1.47	6(4)	0.39	-1.71	7(4)	0.63	-1.47
	8(5)	0.46	-0.99	12(3)	0.92	-1.69	12(4)	0.49	-0.35	5(5)	0.90	-0.35	6(5)	0.89	-1.40	6(4)	0.88	0.20	6(4)	0.71	-0.93	6(5)	0.63	0.33	7(5)	0.72	0.37
	9_1	0.93	-1.66	12(4)	0.86	-0.91	12(5)	0.21	-1.00	6	0.71	-1.25	7(1)	0.69	-2.14	6(5)	0.85	-0.20	6(5)	0.64	-0.02	7	-0.10	-1.03	8(1)	-0.02	-1.07
	9_2	-	-	12(5)	0.62	-1.49	12(6)	0.34	-0.06	7	0.69	-0.87	7(2)	0.43	-0.93	7(1)	0.24	-0.81	7	0.62	-1.16	8(1)	0.63	1.23	8(2)	0.75	-0.79
	9_3	-	-	13(1)			13(1)	0.46	-0.18	8	0.53	-0.35	8(1)	0.64	-1.40	7(2)	0.76	-0.34	8(1)	0.63	-1.16	8(2)	0.52	-0.28	8(3)	0.87	-0.57
	10(1)	0.90	-1.66	13(2)	0.18	0.23	13(2)	0.47	0.06	9(1)	0.37	-0.80	8(2)	0.18	-1.40	8	0.73	-0.72	8(2)	0.53	-1.25	9(1)	0.10	-0.94	8(4)	0.24	0.97
	10(2)	0.91	-1.85	13(3)			14(1)	-	-	9(2)	0.72	-2.03	8(3)	0.33	-0.29	9	0.50	-0.72	9	0.18	-0.07	9(2)	0.13	-0.58	9(1)	0.63	0.88
	10(3)	0.96	-1.66	14(1)			14(2)	-	-													9(3)	0.93	0.11	9(2)	0.49	-0.79
	11(1)	0.84	-1.85	14(2)	0.83	-0.11	15(1)	-	-													9(4)	0.28	1.23	10	0.28	-1.18
	11(2)	-	-	14(3)			15(2)	0.94	-1.86																		
	11(3)_1	0.82	-1.66				16(1)	0.53	-1.00																		
	11(3)_2	0.72	-1.40				16(2)	0.33	-0.94																		
	11(4)_1	0.42	0.12																								
	11(4)_2	0.88	-1.85																								

Table 5 Factor Pattern and Threshold Values of the Items of Reasoning

Elementary School						Junior High School		
Grades 1 to 3			Grades 4 to 6			Grades 1 to 3		
Item	Factor Pattern	Threshold	Item	Factor Pattern	Threshold	Item	Factor Pattern	Threshold
(1)	0.70	-0.76	(1)	0.78	-0.03	(1)	0.88	-0.29
(2)	0.87	-0.66	(2)	0.80	-0.57	(2)	0.78	-1.19
(3)	0.87	-0.72	(3)	0.85	0.06	(3)	0.44	-0.03
(4)	0.80	-0.39	(4)	0.66	-0.09	(4)	0.36	-0.17

Table 6 shows the correlation coefficients between the factors worked out for each school grade. The range in correlation coefficients is large, but the values are generally high in most school grades. In particular, after removing junior high school grade 1, which exceptionally did not show a significant correlation, and junior high school grade 2, for which the correlation was significant but low, the correlation between the two constructs of "arithmetic/mathematics" and "Japanese language" was, at over 0.76, extremely high. The correlations between "arithmetic/mathematics" and "reasoning" and between "Japanese language" and "reasoning" were, at below 0.52, lower than the correlation between "arithmetic/mathematics" and "Japanese language."

Table 6 Correlation between the Factors using the Three-Factor Model (Pearson's *r*)

	Elementary School						Junior High School		
	1	2	3	4	5	6	1	2	3
Arithmetic/Mathematics – Japanese Language	0.99 ***	0.76 ***	0.88 ***	0.98 ***	0.88 ***	0.89 ***	0.19	0.36 **	0.97 ***
Arithmetic/Mathematics – Reasoning	0.34 ***	0.48 ***	0.27 **	0.27 *	0.20	-0.03	0.08	0.24	0.50 ***
Japanese Language – Reasoning	0.33 ***	0.37 **	0.20	0.25 *	0.20	0.15	0.15	0.42 ***	0.52 ***

Note. ***, ** and * indicate that the coefficients are statistically significant at standards of 0.1%, 1%, and 5%, respectively.

Cronbach's alpha coefficient showed sufficiently high values in arithmetic/mathematics. The lowest value was 0.83, which was in elementary school grades 4 and 5, and the highest value was 0.95, which was in junior high school grade 3. The average for all school grades was 0.88. These high values confirmed the high degree of internal consistency (Chart 3-7). There was also a sufficiently high degree of internal consistency for Japanese language. The lowest value for Japanese language was 0.79, which was in junior high school grade 2 (0.81 when one item for which the factor pattern showed negative is removed), and the highest value was 0.96, which was in elementary school grade 1. The mean value for all school grades was 0.87. As for reasoning,

although the value from junior high school grade 1 to 3 was, at 0.69, low, from elementary school grade 1 to 3 it was, at 0.88, high, and the mean value was 0.80.

Table 7 Internal Consistency of the Subject Items in Each School Grade (Cronbach's Alpha)

		Arithmetic/ Mathematics		Japanese Language		Reasoning	
		α	Number of Items	α	Number of Items	α	Number of Items
Elementary School	1	0.89	14	0.96	15		
	2	0.90	15	0.82	11	0.88	4
	3	0.89	19	0.85	15		
	4	0.83	10	0.88	13		
	5	0.83	9	0.87	13	0.85	4
	6	0.90	10	0.89	13		
Junior High School	1	0.87	10	0.89	13		
	2	0.87	10	0.79 (0.81)	15 (14)	0.69	4
	3	0.95	10	0.85 (0.87)	15 (14)		
Mean		0.88		0.87 (0.88)		0.80	

Note. Parenthesis = When items showing negative factor pattern are removed.

2 Comparison of the Correct Answer Rates in the Preliminary Survey and the Academic Ability Test by Prefecture

Regarding the correct answer rates in the JCPS 2012 and the preliminary survey, in elementary school grade 1 to elementary school grade 6, the JCPS had significantly higher rates for all three school subjects of arithmetic ($t = 6.37$, $df = 75$, $p < 0.001$), Japanese language ($t = 8.43$, $df = 87$, $p < 0.001$), and reasoning ($t = 6.94$, $df = 20$, $p < 0.001$) (Table 8; Table 9). In junior high school grades 1 to 3, the JCPS had significantly higher rates for all three subjects of arithmetic ($t = 2.43$, $df = 24$, $p < 0.05$), Japanese language ($t = 2.50$, $df = 34$, $p < 0.05$), and reasoning ($t = 2.40$, $df = 8$, $p < 0.05$).

However, regarding the correct answer rates in the JCPS and the academic ability test by prefecture, in elementary school grades 4 to 6, the JCPS had a significantly lower correct answer rate for arithmetic ($t = -2.27$, $df = 22$, $p < 0.05$), but there was no significant difference in the correct answer rate for Japanese language ($t = -0.84$, $df = 34$, $p = 0.41$). There was no significant difference in the correct answer rates for arithmetic from junior high school grades 1 to 3 ($t = 0.95$, $df = 14$, $p = 0.36$), but the JCPS had significantly higher correct answer rates for Japanese language ($t = 3.06$, $df = 30$, $p < 0.01$).

Even when including all the elementary and junior high school grades and comparing the correct answer rates for corresponding items among every school subject, the JCPS has significantly higher correct answer rates than the preliminary survey ($t = 11.79$, $df = 253$, $p < 0.001$). However, there was

no such significance between the JCPS and the academic ability test by prefecture ($t = 0.95$, $df = 103$, $p = 0.34$).

Table 8　Comparison of Correct Answer Rates in the JCPS / Preliminary Survey (Elementary School Grade 1 to Elementary School Grade 3)

Subject	Grade 1			Grade 2			Grade 3		
	JCPS		Preliminary Survey	JCPS		Preliminary Survey	JCPS		Preliminary Survey
	Item	Rate	Rate	Item	Rate	Rate	Item	Rate	Rate
Arithmetic/ Mathematics	1(1)	0.93	0.94	1(1)	0.95	0.91	1	0.97	0.93
	1(2)	0.90	0.81	1(2)	0.82	0.56	2	0.89	0.73
	2(1)	0.98	0.94	2(1)	1.00	0.92	3(1)	0.89	0.85
	2(2)	0.92	0.88	2(2)	1.00	0.98	3(2)	0.87	0.88
	3(1)	0.93	0.92	3(1)	0.89	0.78	4	0.87	0.78
	3(2)	0.62	0.49	3(2)	0.82	0.48	5(1)	0.94	0.90
	4(1)	0.90	0.49	4	0.86	NA	5(2)	0.46	NA
	4(2)	0.80	0.80	5	0.89	0.95	6_1	0.65	0.63
	5(1)	0.98	0.94	6	0.91	0.77	6_2	0.60	0.49
	5(2)	0.95	0.84	7(1)	0.98	0.95	7	0.97	0.93
	5(3)	0.93	0.86	7(2)	0.98	0.95	8	0.92	0.87
	5(4)	0.87	0.68	7(3)	0.95	0.92	9(1)	0.97	NA
	6(1)	0.79	0.76	7(4)	0.98	0.92	9(2)	0.97	0.99
	6(2)	0.74	0.64	8(1)	0.91	0.81	9(3)	0.92	0.88
				8(2)	0.89	0.69	9(4)	0.87	0.82
				9_1	0.89	0.81	9(5)	0.90	0.90
				9_2	0.66	0.44	9(6)	0.89	0.91
							10(1)	0.95	0.96
							10(2)	0.78	0.75
Japanese Language	7(1)	0.92	0.82	10(1)	0.75	0.11	11(1)	0.97	0.96
	7(2)	0.90	0.86	10(2)	0.86	0.91	11(2)	0.94	0.82
	7(3)	0.95	0.82	10(3)	1.00	0.97	11(3)	0.98	0.96
	8(1)	1.00	0.96	11(1)	0.84	0.09	11(4)	0.86	0.73
	8(2)	1.00	0.94	11(2)	0.80	0.66	12(1)	0.79	0.71
	8(3)	0.98	0.82	12(1)	0.84	0.73	12(2)	0.71	0.58
	8(4)	0.97	NA	12(2)	0.95	0.88	12(3)	0.84	0.82
	8(5)	0.85	0.58	12(3)	0.95	0.78	12(4)	0.63	0.80
	9_1	0.97	0.92	12(4)	0.82	0.02	12(5)	0.84	0.83
	9_2	0.97	0.92	12(5)	0.93	0.85	12(6)	0.52	0.32
	9_3	0.97	0.86	13(1)	0.84	0.31	13(1)	0.57	0.52
	10(1)	0.97	0.94	13(2)	0.43	0.13	13(2)	0.48	0.16
	10(2)	0.98	0.92	13(3)	0.82	0.22	14(1)	1.00	0.99
	10(3)	0.97	0.86	14(1)	0.68	0.34	14(2)	1.00	0.97
	11(1)	0.98	0.98	14(2)	0.55	0.17	15(1)	1.00	0.97
	11(2)	1.00	0.98	14(3)	0.75	0.28	15(2)	0.97	0.96
	11(3)_1	0.97	NA				16(1)	0.84	0.73
	11(3)_2	0.93	NA				16(2)	0.83	0.50
	11(4)_1	0.46	NA						
	11(4)_2	0.98	NA						
Reasoning	12(1)	0.67	0.56	15(1)	0.82	0.63	17(1)	0.86	0.64
	12(2)	0.62	0.56	15(2)	0.77	0.64	17(2)	0.86	0.67
	12(3)	0.66	0.54	15(3)	0.84	0.61	17(3)	0.83	0.76
	12(4)	0.49	0.22	15(4)	0.82	0.50	17(4)	0.70	0.62

Table 9　Comparison of the Correct Answer Rates in the JCPS / Preliminary Survey / Academic Ability Test by Prefecture Correct Answer Rates Comparison (Elementary School Grade 4 to Junior High School Grade 3)

Elementary School Grade 4

Subject	JCPS Item	JCPS Rate	Preliminary Survey Rate	Academic Ability Test by Prefecture Rate	School Grade	Implementation Period		Prefecture
Arithmetic/ Mathematics	1(1)	0.79	0.68	0.87	E4	2006 /7		Akita
	1(2)	0.77	0.51	0.80	E4	2006 /7		Akita
	1(3)	0.96	0.76	0.96	E4	2006 /7		Akita
	1(4)	0.91	0.87	0.97	E4	2007 /7		Akita
	2(1)	0.70	0.17	0.54	E4	2006 /1		Niigata
	2(2)	0.40	0.13	0.46	E4	2004 /1		Niigata
	2(3)	0.89	0.64	0.85	E4	2006 /1		Niigata
	2(4)	0.83	0.45	0.79	E4	2006 /1		Niigata
	3(1)	0.70	0.55	0.78	E4	2004 /1		Niigata
	3(2)	0.87	0.51	0.95	E4	2006 /1		Niigata
Japanese Language	4(1)	0.98	0.89	0.97	E4	2006 /7		Akita
	4(2)	0.98	0.93	0.98	E4	2004 /1		Niigata
	4(3)	0.85	0.45	0.87	E4	2005 /7		Akita
	5(1)	0.89	0.66	0.89	E4	2007 /7		Akita
	5(2)	0.83	0.77	0.81	E4	2006 /7		Akita
	5(3)	0.53	0.02	0.75	E4	2008 /12		Akita
	5(4)	0.57	0.25	0.72	E4	2008 /12		Akita
	5(5)	0.64	0.25	0.61	E5	2008 /1		Gifu
	6	0.89	0.66	0.88	E4	2008 /12		Akita
	7	0.81	0.19	0.78	E4	2005 /7		Akita
	8	0.64	0.34	0.69	E4	2007 /7		Akita
	9(1)	0.79	0.72	0.87	E5	2006 /4-5		Kagawa
	9(2)	0.98	0.91	0.90	E5	2006 /4-5		Kagawa
Reasoning	10(1)	0.40	0.38	NA				
	10(2)	0.62	0.59	NA				
	10(3)	0.40	NA	NA				
	10(4)	0.45	0.34	NA				

Elementary School Grade 5

Subject	JCPS Item	JCPS Rate	Preliminary Survey Rate	Academic Ability Test by Prefecture Rate	School Grade	Implementation Period		Prefecture
Arithmetic/ Mathematics	1(1)	0.98	0.15	0.95	E5	2008 /1		Gifu
	1(2)	0.89	0.77	0.96	E4	2006 /7	*	Akita
	1(3)	0.73	0.35	0.88	E5	2006 /7		Akita
	1(4)	0.71	0.68	0.69	E5	2003 /2		Gifu
	2(1)	0.76	0.74	0.79	E4	2006 /1	*	Niigata
	2(2)	0.79	0.47	0.74	E5	2004 /1		Niigata
	3	0.52	0.06	0.54	E5	2007 /7		Akita
	4(1)	0.74	0.32	0.63	E5	2008 /12	**	Akita
	4(2)	0.84	0.79	0.87	E5	2004 /1		Niigata
Japanese Language	5(1)	0.95	0.85	0.87	E4	2005 /7	*	Akita
	5(2)	0.77	0.53	0.63	E5	2009 /4-5		Kagawa
	5(3)	0.44	0.12	0.26	E5	2009 /4-5		Kagawa
	6(1)	0.44	0.15	0.38	E5	2008 /1		Gifu
	6(2)	0.76	0.91	0.62	E5	2005 /1		Gifu
	6(3)	0.94	0.82	0.92	E5	2004 /1		Gifu
	6(4)	0.90	0.65	0.86	E5	2006 /1		Niigata
	6(5)	0.92	0.38	0.81	E4	2006 /7	*	Akita
	7(1)	0.98	0.85	0.94	E5	2004 /1		Niigata
	7(2)	0.82	0.62	0.75	E5	2004 /1		Niigata
	8(1)	0.92	0.97	0.87	E5	2006 /4-5		Kagawa
	8(2)	0.92	0.85	0.90	E5	2006 /4-5		Kagawa
	8(3)	0.61	0.35	0.72	E6	2008 /4-5		Kagawa
Reasoning	9(1)	0.55	0.50	NA				
	9(2)	0.77	0.71	NA				
	9(3)	0.52	NA	NA				
	9(4)	0.55	0.50	NA				

Elementary School Grade 6

Subject	JCPS Item	JCPS Rate	Preliminary Survey Rate	Academic Ability Test by Prefecture Rate	School Grade	Implementation Period		Prefecture
Arithmetic/ Mathematics	1(1)	0.87	0.85	0.95	E5	2008 /1	*	Gifu
	1(2)	0.87	0.82	0.83	E6	2003 /2		Gifu
	1(3)	0.71	0.90	0.90	E6	2003 /2		Gifu
	1(4)	0.21	0.11	0.38	E6	2008 /4-5		Kagawa
	2	0.63	0.66	0.74	E6	2006 /1		Niigata
	3(1)	0.58	0.61	0.54	E6	2007 /7	*	Akita
	3(2)	0.71	0.72	0.74	E6	2006 /4-5		Kagawa
	4(1)	0.34	0.36	NA				
	4(2)	0.45	0.28	0.60	E6	2007 /7		Akita
	4(3)	0.74	0.77	0.80	J1	2006 /4-5		Kagawa
Japanese Language	5(1)	0.76	0.56	0.63	E5	2009 /4-5	*	Kagawa
	5(2)	0.84	0.92	0.90	J1	2006 /1		Gifu
	5(3)	0.71	0.41	0.84	E6	2008 /12		Akita
	5(4)	0.61	0.38	0.64	E6	2008 /12		Akita
	6(1)	0.63	0.08	0.66	E6	2006 /4-5		Kagawa
	6(2)	0.66	0.79	0.62	E5	2005 /1	*	Gifu
	6(3)	0.71	0.69	0.77	E6	2005 /1		Gifu
	6(4)	0.42	0.25	0.56	E6	2008 /1		Gifu
	6(5)	0.58	0.41	0.79	E6	2008 /12		Akita
	7(1)	0.79	0.51	0.72	E6	2008 /4-5		Kagawa
	7(2)	0.63	0.64	0.69	E6	2008 /4-5		Kagawa
	8	0.76	0.85	0.84	E6	2008 /4-5		Kagawa
	9	0.76	0.39	0.82	E6	2006 /1		Gifu
Reasoning	10(1)	0.58	0.51	NA				
	10(2)	0.74	0.66	NA				
	10(3)	0.50	NA	NA				
	10(4)	0.63	0.46	NA				

Junior High School Grade 1

Subject	JCPS Item	JCPS Rate	Preliminary Survey Rate	Academic Ability Test by Prefecture Rate	School Grade	Implementation Period		Prefecture
Arithmetic/ Mathematics	1(1)	0.75	NA	0.64	J1	2007 /1		Gifu
	1(2)	0.83	0.85	0.91	E6	2004 /1	*	Gifu
	1(3)	0.83	0.77	0.71	J1	2007 /1		Gifu
	1(4)	0.72	0.65	0.60	J1	2007 /1		Gifu
	2	0.74	0.77	0.74	E6	2006 /1	*	Niigata
	3(1)	0.74	0.58	0.76	J1	2003 /2		Gifu
	3(2)	0.61	0.58	0.49	J1	2004 /1		Gifu
	3(3)	0.35	0.42	0.40	J1	2008 /12		Akita
	4(1)	0.58	0.46	0.73	J1	2007 /7	**	Akita
	4(2)	0.35	0.23	0.73	J2	2008 /4-5	**	Kagawa
Japanese Language	5(1)	0.86	0.54	0.77	J2	2006 /1		Niigata
	5(2)	0.91	0.89	0.90	J1	2006 /1		Gifu
	5(3)	0.83	0.85	0.60	J2	2008 /1		Gifu
	5(4)	0.56	0.19	0.49	J1	2007 /7		Akita
	6(1)	0.81	0.69	0.77	E6	2008 /12	*	Akita
	6(2)	0.90	0.69	0.87	J1	2005 /1		Gifu
	6(3)	0.93	0.77	0.59	J1	2003 /2		Gifu
	6(4)	0.83	0.54	0.73	J1	2008 /1		Gifu
	6(5)	0.51	0.35	0.36	J1	2009 /4-5		Kagawa
	7	0.88	0.81	0.84	E6	2008 /4-5	*	Kagawa
	8(1)	0.88	0.77	0.79	J1	2003 /2		Gifu
	8(2)	0.90	0.73	0.69	J1	2005 /7		Akita
	9	0.53	0.58	0.31	J1	2009 /4-5		Kagawa
Reasoning	10(1)	0.61	0.42	NA				
	10(2)	0.84	0.77	NA				
	10(3)	0.58	NA	NA				
	10(4)	0.61	0.27	NA				

Junior High School Grade 2

Subject	JCPS Item	JCPS Rate	Preliminary Survey Rate	Academic Ability Test by Prefecture Rate	School Grade	Implementation Period		Prefecture
Arithmetic/ Mathematics	1(1)	0.76	0.72	0.60	J1	2007 /1	*	Gifu
	1(2)	0.89	0.84	0.88	J2	2003 /2		Gifu
	1(3)	0.72	0.72	0.61	J2	2004 /1		Gifu
	1(4)	0.63	0.53	0.50	J2	2003 /2		Gifu
	2(1)	0.54	NA	0.48	J3	2006 /7		Akita
	2(2)	0.48	0.25	0.40	J1	2008 /12	*	Akita
	3(1)	0.70	NA	NA	J3	2009	***	Tokyo
	3(2)	0.76	0.50	0.78	J3	2007 /7		Akita
	4(1)	0.80	0.66	0.75	J2	2007 /1	**	Gifu
	4(2)	0.52	0.41	0.56	J1	2005 /7	*	Akita
Japanese Language	5(1)	0.74	0.88	0.45	J2	2003 /2		Gifu
	5(2)	0.96	0.91	0.90	J2	2003 /2		Gifu
	5(3)	0.70	0.66	0.66	J2	2004 /1		Gifu
	6(1)	0.85	0.75	0.66	J2	2004 /1		Gifu
	6(2)	0.80	0.75	0.76	J2	2005 /1		Gifu
	6(3)	0.48	0.53	0.33	J2	2006 /1		Gifu
	6(4)	0.96	0.88	0.89	J2	2006 /1		Niigata
	6(5)	0.37	NA	0.52	J3	2007 /7		Akita
	7	0.85	0.72	0.82	J2	2006 /1		Gifu
	8(1)	0.11	NA	NA				
	8(2)	0.61	NA	NA				
	9(1)	0.83	0.84	0.79	J2	2006 /1		Niigata
	9(2)	0.72	0.91	0.79	J1	2003 /2	*	Gifu
	9(3)	0.46	NA	0.46	J2	2003 /2		Gifu
	9(4)	0.11	0.03	0.25	J2	2007 /1		Gifu
Reasoning	10(1)	0.65	0.59	NA				
	10(2)	0.91	0.72	NA				
	10(3)	0.39	NA	NA				
	10(4)	0.54	0.56	NA				

Junior High School Grade 3

Subject	JCPS Item	JCPS Rate	Preliminary Survey Rate	Academic Ability Test by Prefecture Rate	School Grade	Implementation Period		Prefecture
Arithmetic/ Mathematics	1	0.62	NA	NA	J3	2009	***	Tokyo
	2(1)	0.57	NA	NA	J3	2008	***	Akita
	2(2)	0.45	NA	NA	J3	2005	***	Ehime
	3(1)	0.71	0.63	0.73	J3	2007 /7		Akita
	3(2)	0.52	0.75	0.67	H1	2006 /4-5		Kagawa
	3(3)	0.88	0.81	0.94	H1	2006 /4-5		Kagawa
	4(1)	0.55	0.56	0.69	H1	2006 /4-5		Kagawa
	4(2)	0.64	NA	NA	J3	2008	***	Akita
	5(1)	0.67	0.66	0.68	J3	2005 /7	**	Akita
	5(2)	0.50	0.56	0.61	J3	2006 /7	**	Akita
Japanese Language	6(1)	0.86	0.78	0.66	J2	2004 /1	*	Gifu
	6(2)	0.81	0.88	0.85	J3	2005 /7		Akita
	6(3)	0.98	0.94	0.82	J3	2005 /7		Akita
	7(1)	0.74	0.84	0.33	J2	2006 /1	*	Gifu
	7(2)	0.79	0.84	0.83	J3	2005 /7		Akita
	7(3)	0.91	0.94	0.91	J3	2006 /7		Akita
	7(4)	0.93	0.91	0.90	J3	2007 /7		Akita
	7(5)	0.36	0.45	0.52	J3	2007 /7		Akita
	8(1)	0.86	0.97	0.85	J3	2006 /7		Akita
	8(2)	0.79	0.63	0.58	J3	2006 /7		Akita
	8(3)	0.71	NA	0.46	J2	2003 /2	*	Gifu
	8(4)	0.17	NA	0.25	J2	2007 /1	*	Gifu
	9(1)	0.19	NA	NA				
	9(2)	0.79	NA	NA				
	10	0.88	0.88	0.91	J3	2006 /7		Akita
Reasoning	11(1)	0.57	0.50	NA				
	11(2)	0.91	0.78	NA				
	11(3)	0.55	NA	NA				
	11(4)	0.52	0.63	NA				

Note. *: Cases where the dates of the academic ability test by prefecture and the JCPS implemented

(in April to May of the school year above the school year in question) are more than 12 months apart.
**: Cases where there are revisions made to the questions set in the academic ability test by prefecture.
***: Items were extracted from prefectural/metropolitan high school entrance exam.

What follows is the correlation of the correct answer rates for corresponding items. In elementary school grades 1 to 3 and in junior high school grades 1 to 3, there was around 0.8 of correlation between the JCPS and the preliminary survey; however, in elementary school grades 4 to 6, arithmetic showed a lower value at 0.55 (Table 10). In the preliminary survey, it was revealed through comments given by teachers of the participating schools that the elementary school grade 5 students had not been taught calculating fractions yet at the time of the survey. When calculating the correlation coefficient after removing the two items in question (elementary school grade 5 arithmetic 1 (1), 3 in Table 9), arithmetic in elementary school grades 4 to 6 rose to 0.68 (27 items). When including every school grade, the correlation was 0.75 for arithmetic/mathematics (101 items), 0.74 for Japanese language (123 items), and 0.76 for reasoning (30 items).

Regarding the comparison of the JCPS and the academic ability test by prefecture, when mathematics for junior high school grades 1 to 3 (15 items) was removed, the correlation coefficients for arithmetic/mathematics and Japanese language all exceeded 0.80 and, at 0.89, they were highest in arithmetic for elementary school grades 4 to 6 (23 items). When including all school grades, the correlation was 0.85 for arithmetic/mathematics (38 items) and 0.82 for Japanese language (66 items) (Table 10).

Table 10 Correlation of the Correct Answer Rates in the JCPS, the Preliminary Survey, and the Academic Ability Test by Prefecture (Pearson's *r*)

School Grade		Subject	Preliminary Survey		Academic Ability Test	
			r	Number of Items	*r*	Number of Items
Elementary School	1 to 3	Arithmetic	0.82 ***	47	NA	
		Japanese Language	0.77 ***	49	NA	
		Reasoning	0.77 ***	12	NA	
	4 to 6	Arithmetic	0.55 ***	29	0.89 ***	23
		Japanese Language	0.76 ***	39	0.84 ***	35
		Reasoning	0.93 ***	9	NA	
Junior High School	1 to 3	Mathematics	0.80 ***	25	0.79 ***	15
		Japanese Language	0.81 ***	35	0.83 ***	31
		Reasoning	0.68 **	9	NA	
All		Arithmetic/Mathematics	0.75 ***	101	0.85 ***	38
		Japanese Language	0.74 ***	123	0.82 ***	66
		Reasoning	0.76 ***	30	NA	

Note. *** and ** indicate that the coefficients are statistically significant at standards of 0.1% and 1%, respectively.

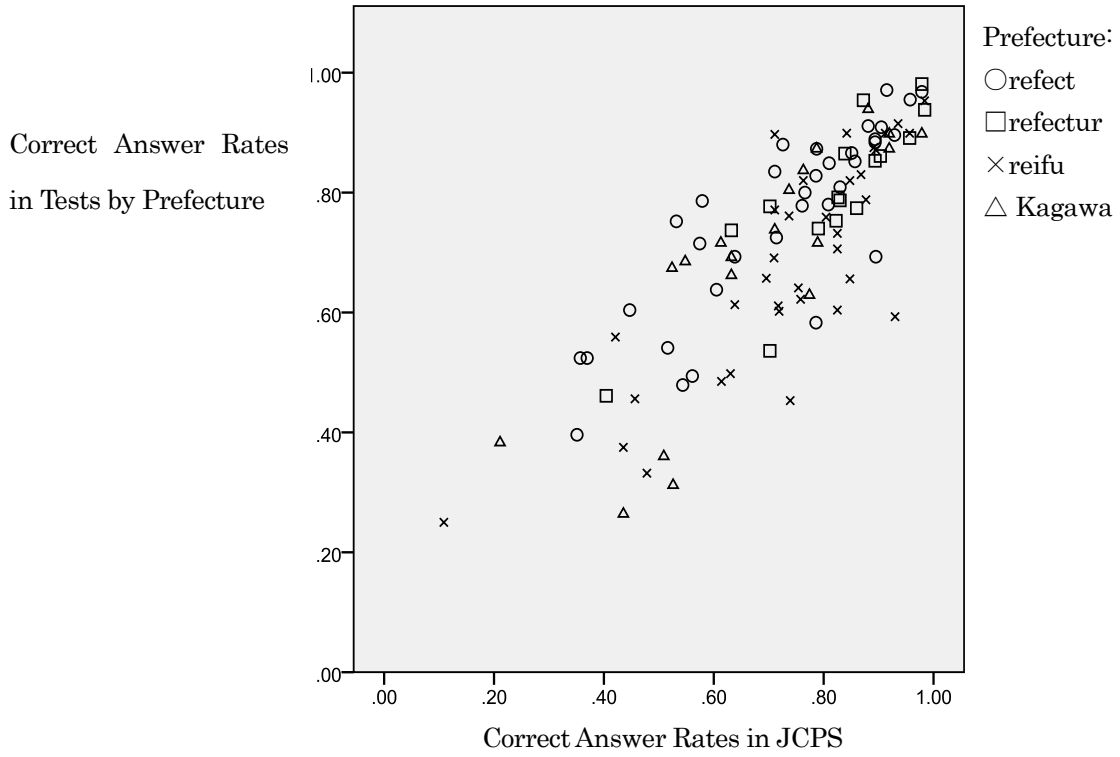
When the information regarding the academic ability test by prefecture is categorized according to prefecture and analyzed, it is revealed that the academic ability test by prefecture has significantly higher correct answer rates in Akita Prefecture yet significantly lower correct answer rates in Gifu Prefecture, but there is no significant difference with Kagawa Prefecture and Niigata Prefecture (Table 11). Regarding corresponding items among elementary school grade 4 to junior high school grade 3, Figure 2 depicts the distribution of the correct answer rates of the JCPS and the correct answer rates for each prefecture in the academic ability test by prefecture. The correlation coefficient for these 104 items was 0.82 (Table 10). This figure rises even more when calculating by prefecture (Table 11; 0.82 – 0.88).

Table 11　Comparison by Prefecture of the Average Correct Answer Rates in the Academic Ability Test by Prefecture and the JCPS (Elementary School Grade 4 to Junior High School Grade 3)

	Correct Answer Rate Average	Standard Deviation	Number of Items	*t Value*	*df*	Pearson's *r*
JCPS	0.72	0.18	34	-2.08 **	33	0.85 ***
Prefectural Test: Akita	0.76	0.16				
JCPS	0.81	0.15	16	1.03	15	0.88 ***
Prefectural Test: Niigata	0.79	0.14				
JCPS	0.74	0.18	35	3.02 ***	34	0.82 ***
Prefectural Test: Gifu	0.68	0.18				
JCPS	0.68	0.19	19	-0.13	18	0.84 ***
Prefectural Test: Kagawa	0.68	0.21				

Note. *** and ** indicate that the coefficients are statistically significant at standards of 0.1% and 1%, respectively.

Figure 2　Distribution by Prefecture of the Correct Answer Rates for Each Item in the JCPS and the Academic Ability Test by Prefecture (Elementary School Grade 4 to Junior High School Grade 3)



3 Test Scores and School Results

The correlation between the arithmetic/mathematics factor scores and the arithmetic/mathematics results was low in elementary school grade 1 and elementary school grade 3 at 0.39 and 0.47, respectively, but the correlation was higher in other school grades, ranging from 0.57 to 0.67, and the mean value for all school grades was 0.57 (Table 12). As for the Japanese language factor scores and the Japanese language results, they were low in elementary school grade 3 and junior high school grade 1 at 0.39 and 0.14, respectively, but the correlation ranged from 0.43 to 0.62 in other school grades, and the mean value for all school grades was 0.46.

In all the school grades with the exception of elementary school grade 1, arithmetic/mathematics factor scores tended to correlate more strongly with mathematics results than with Japanese language results. However, it was not possible to identify a clear tendency for Japanese language factor scores to correlate more strongly with mathematics results than with Japanese language results.

Table 12　Correlation between School Results and the Factor Scores for Two School Subjects (Pearson's *r*)

	School Results											
	Elementary School											
	Grade 1		Grade 2		Grade 3		Grade 4		Grade 5		Grade 6	
	Arithmetic	Japanese	Arithmetic	Japanese	Arithmetic	Japanese	Arithmetic	Japanese	Arithmetic	Japanese	Arithmetic	Japanese
Arithmetic	.39 ***	.44 ***	.59 ***	.47 ***	.47 ***	.41 ***	.57 ***	.54 ***	.58 ***	.44 ***	.61 ***	.57 ***
Japanese Language	.39 ***	.44 ***	.55 ***	.43 ***	.41 ***	.39 ***	.61 ***	.57 ***	.56 ***	.49 ***	.58 ***	.62 ***

	School Results							
	Junior High School						Average for All Grades	
	Grade 1		Grade 2		Grade 3			
	Mathematics	Japanese	Mathematics	Japanese	Mathematics	Japanese	Mathematics	Japanese
Mathematics	.61 ***	.46 ***	.67 ***	.14	.64 ***	.52 ***	.57 ***	.44 ***
Japanese Language	.00	.14	.52 ***	.50 **	.61 ***	.53 ***	.47 ***	.46 ***

Note. *** and ** indicate that the coefficients are statistically significant at standards of 0.1% and 1%, respectively.

Section 4 Discussion

The confirmatory categorical factor analysis conducted on each item of arithmetic/mathematics, Japanese language, and reasoning for every school grade showed that the three-factor structure has good fit. This suggests that the constructs measured for each school subject correlate together but are not the same. However, the two-factor model also has an admirable goodness-of-fit in a number of school grades; additionally, the validity of the model that considers arithmetic/mathematics and Japanese language together in one-dimension also stands up. Although the reliability coefficients that evaluate based on the internal consistency of the items of arithmetic/mathematics, Japanese language, and reasoning were slightly lower in reasoning, which has a low number of items, they were generally extremely high. Consequently, it could be confirmed that the test items have high reliability.

A comparison of the average values of the correct answer rates for the JCPS academic ability test and those of the existing correct answer rates data showed, when junior high school mathematics is excluded, that the JCPS has higher correct answer rates than the preliminary survey. However, the preliminary survey data for both elementary school and junior high school were collected from 1 to 2 classes in one school in Shizuoka Prefecture, and it is not considered that the test items' construct validity is challenged by the results of comparing the correct answer rates between two specific samples. Indeed, in view of the fact that the JCPS was implemented around 5 months after the

preliminary survey, these results are entirely appropriate, and they cannot be considered proof of the test items' weak construct validity.

What is important is the comparison with the large-scale academic ability test by prefecture. It is remarkable that an examination of all the items in all the school grades did not reveal any difference in the mean values for the correct answer rates. While the task of examining the representativeness of the JCPS sample is left to Yamashita et al. (2011), based on the premise that not only the sample properties of the JCPS but also the academic level of the sampled children are representative of Japan as a whole, then the fact that there was no difference in mean values in the correct answer rates for the four prefectures may suggest the strength of the test items' construct validity. Moreover, when comparing prefectures, although the correct answer rates were significantly higher than those to the JCPS in Akita Prefecture and lower than those to the JCPS in Gifu Prefecture, it is particularly noteworthy that there was no difference between the survey results in Kagawa Prefecture and Niigata Prefecture. The national academic ability survey, which targeted elementary school grade 6 and junior high school grade 3, has disclosed the scores by prefecture for the correct answer rates in arithmetic/mathematics and Japanese language. (Regarding "2010 National Academic Ability / Learning Situation Survey / Report / Collected Results," please refer to the National Institute for Educational Policy Research website http://www.nier.go.jp/10chousakekkahoukoku/.) The survey reports that Akita Prefecture has a steady, high level of academic ability in arithmetic/mathematics and Japanese language. The fact that Akita's correct answer rates were higher than those to the JCPS may provide additional evidence that corroborates the strong construct validity of the JCPS academic ability test items. Furthermore, the correlation coefficient of all the items corresponding between the preliminary survey and the JCPS is 0.75, and that with the academic ability test by prefecture is 0.82. Both these figures are high, affirming the strength of the test items' construct validity.

Regarding the correlation between the children's school results as reported by the parents and the JCPS academic ability test scores, in terms of the mean values for every school grade, arithmetic/mathematics was 0.57 and Japanese language was 0.46. Such a moderate degree of correlation supports the concurrent validity of the academic ability test and, simultaneously, shows that the test scores accompany elements that the school results reported by parents cannot explain. Furthermore, the arithmetic/mathematics tests had a higher concurrent validity with the corresponding school subject results; to add to this, although discriminant validity was observed in the arithmetic/mathematics tests, it was not observed in the Japanese language tests. These findings show that arithmetic/mathematics in the JCPS academic ability test has a higher predictive capability regarding the results of this school subject compared to Japanese language. This may be because

arithmetic/mathematics test questions are more inclusive than Japanese language test questions, which are limited to vocabulary and kanji character reading and do not include reading comprehension and composition; hence, they reflect more broadly the elements learnt as part of that specific school subject. Alternatively, this could be interpreted as being suggestive of how the degree of correlation between the children's measured academic ability and the school's performance assessment varies according to school subject.

The reliability and the validity of the JCPS academic ability test have been confirmed. Even though the JCPS is a self-administered survey that utilized the postal method and was thus implemented in test conditions that do not necessarily guarantee exact measurement, the results of the examination indicate that it is highly precise and that it does measure the concepts that need to be measured. The results appear to back up the effectiveness of the JCPS academic ability test answers.

Admittedly a disadvantage, 38 to 63 students in each school grade were targeted in this paper; therefore, this examination was based on a small sample. In particular, the sample size was too small to obtain stable results from categorical factor analysis. It is therefore unclear whether the range observed in the degrees of correlation between the factors calculated for each year (i.e., arithmetic/mathematics, Japanese language, and reasoning) is generated from school grade characteristics, from differences in test questions, or from estimation error potentially due to insufficient sample size. There is a need to conduct further examinations using a larger sample size.

Although such an issue remains, the JCPS academic test scores can serve as an effective academic ability index variable. By conducting analysis linked to the abundant data set of the JHPS, they promise to offer major advances toward clarifying the correlation between academic ability in Japan and Japan's socioeconomic background variables.

References

Bara, B. G., M. Bucciarelli and P. N. Johnson-Laird (1995) "The development of syllogistic reasoning," *American Journal of Psychology*, vol.108, pp.157–193.

Bartholomew D. J. (1980) "Factor analysis for categorical data," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol.42, pp.293–321.

Cronbach, L. J. (1951) "Coefficient alpha and the internal structure of tests," *Psychometrika*, vol.16, pp.297-334.

Gottfredson, L. S. (1997) "Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography," *Intelligence*, vol.24, pp.13-23.

Greene, W. H. and D. A. Hensher (2010) *Modeling Ordered Choices: A Primer*, Cambridge

University Press.

Hambleton, R. K. and H. Swaminathan (1985) *Item Response Theory: Principles and Applications*, Kluwer-Nijhoff Publishing.

Higuchi, Y., T. Miyauchi, C. R. McKenzie, and Keio University Joint Research Center for Panel Studies (eds.), *Dynamism of Education, Health, and Poverty*, Keio University Press (in Japanese).

Kariya, T. (2011) "Japanese solutions to the equity and efficiency dilemma? Secondary schools, inequity and the arrival of 'universal' higher education," *Oxford Review of Education*, vol. 37, pp. 241–266.

Mimizuka, H. (2007) "Dare ga gakuryoku wo kakutoku surunoka (Who acquires the academic ability?)," in H. Mimizuka and K. Makino (eds.), *Gakuryoku to Toranzisshon no Kiki* (Crisis in Academic Ability and Transmission), Kaneko Shobo, pp.3–24 (in Japanese).

Nunnaly, J. (1978) *Psychometric theory*. New York: McGraw-Hill.

Ojima, F. and S. von Below (2010) "Family background, school system and academic achievement in Germany and in Japan," in J. Dronkers (ed.), *Quality and Inequality of Education: Cross-National Perspectives*, Springer, pp.275–297.

Shikishima, C., K. Hiraishi, S. Yamagata, Y. Sugimoto, R. Takemura, K. Ozaki, M. Okada, T. Toda and J. Ando (2009) "Is *g* an entity? A Japanese twin study using syllogisms and intelligence tests," *Intelligence*, vol.37, pp.256–267.

Shikishima, C., S. Yamagata, K. Hiraishi, Y. Sugimoto, K. Murayama and J. Ando (2011) "A simple syllogism-solving test: Empirical findings and implications for g research," *Intelligence*, vol.39, pp.89–99.

Snyderman, M, and S. Rothman (1988) *The IQ Controversy, the Media and Public Policy*, Transaction Books.

Stanovich, K. E. & R. F. West (1998) "Individual differences in rational thought," *Journal of Experimental Psychology*, vol.127, pp.161−188.

Thompson, B. (2004) *Exploratory and Confirmatory Factor Analysis: Understanding Concepts and Applications*, American Psychological Association.

Toyoda, H. (1998) *Kyobunsankozobunseki: Structural Equation Modeling (Nyumonhen)*, Asakura Publishing (in Japanese).

Yamashita, J., R. Nakamura, H. Akabayashi, M. Naoi, and C. Shikishima (2011) "The Characteristics of Households who Responded to Japan Child Panel Survey 2010," in Y. Higuchi, T. Miyauchi T., C. R. McKenzie and Keio University Joint Research Center for Panel Studies (eds.), *Dynamism of Education, Health, Poverty*, Keio University Press, pp.49–67 (in Japanese).