In this paper, we estimated the traits of AIDS knowledge items found in Japanese junior high school tests using Item Response Theory. Discrimination and difficulty, which are parameters that regulate the item traits, were estimated. With this method, we were able to describe the characteristics of AIDS knowledge items. As a result, items like “one has more risk of getting infected by AIDS when fatigued”, “the body fluid that contains the HIV virus the most is blood”, and “AIDS can be transmitted through sexual intercourse” were found low in difficulty for students, meaning that many students do have an understanding of those items. On the other hand, items that relate to the routes of infection other than humans, like “mosquitoes can carry HIV viruses”, “the disease cannot be transmitted from pets and other animals”, and “there are effective vaccines for AIDS” were found to be high in difficulty. However, “the HIV virus can be transmitted by shaking hands”, “AIDS can be transmitted when the blood of an infected patient touches a wound of a healthy individual”, and others regarding the routes of infection among human beings were also low in difficulty, proving a high understanding by the students.

Keywords: Item Response Theory, AIDS knowledge, difficulty, discrimination, multi-group analysis

1. Significance and Purpose of the Research

AIDS cases continue to increase in the 21st century, which makes a pressing problem globally.

It is known that the risk of AIDS infection can be obviated with proper knowledge and appropriate behavior as is generally well put in the phrase “Education is the best vaccine.” Hence UN institutions and governments of developing countries are greatly concerned about AIDS education. Teaching about AIDS at school brings up many delicate issues because it closely relates to the culture, religion, and customs of the concerned society, and because it is associated with sex education. Teachers have to face challenging questions of what content to teach and in what order it should be taught. In addition, if teachers expect students to understand the etiologic and pathologic background of this disease, they have to teach advanced knowledge including the immune system, which can be difficult to understand.

For instance, in Japan where a highly sophisticated school education system has been established, multi-media-based teaching materials are provided for AIDS education. There are a few studies examining the learning effect of teaching materials and programs through tests (surveys) (Koiso, 1999).

Among these challenging issues, we determined to explore measurement of understanding of basic AIDS knowledge.

Previous studies, however, only explored the outcomes of specific examinees on how much AIDS knowledge was understood (or how much their understanding changed by teaching materials and learning methods) (Iwata, 1998; Watanabe, 1998). No studies have examined the difficulty of survey items, nor have such studies properly discussed the generality of the tests.

Kokudo (2001) estimated some AIDS knowledge items by Item Response Theory (IRT) to examine the items.

There are two problems with using IRT to examine
the appropriateness of test items:

First, “correct response rate (passage rate),” which is used as an index of difficulty of items, depends on a specific test group of examinees. That is, even with the same test items, the correct response rate becomes high if they are given to a group of high achievers while it becomes low if they are given to a group of low achievers. Accordingly, there are limitations of using the correct response rate as an index to express item-specific difficulty. Therefore the item-specific traits cannot be examined by the rate. Thus the appropriateness of items cannot be explored in correspondence with the developmental stage of children.

Second, in all types of tests, the number of items in a test for the same group is limited. For instance, when hundreds of AIDS knowledge items are to be tested, only a certain volume of items (the number of items) can be administered at a time if the burden on examinees is considered. To ease the burden, test items should be divided (to limit the number of items in a test), several test patterns should be prepared, and each should be administered to several different groups of examinees. However, this method cannot compare obtained items because of the first problem stated above. In other words, even if many items are prepared under the conventional test theory, the items cannot be securely examined in a common scale.

For this reason, most test development under the classical test theory holds many contradictions. AIDS knowledge tests are no exception.

The purpose of the present study, focusing on AIDS knowledge understanding, is to develop a scale of AIDS knowledge understanding items as inclusive as possible by adding new items to Kokudo’s (2001) study. We conducted a large scale survey to junior high school students for developing a scale of AIDS knowledge understanding items using IRT. We also examined sustainable survey methods.

2. Method

2.1. Item Response Theory (IRT)

IRT is a test theory, in which statistical models express test item responses and the traits of examinees besides correct response rate and average scores. In the present study, to overcome the two problems presented above, the following methodological steps were taken:

In order to examine as many items as possible for the development of a scale of items, question items are divided in advance to prepare different question items for each examinee group. For the development of a scale of many items, the number of such items is too large to be administered to the same group of examinees. As a solution, the items have to be evenly divided among several groups of examinees. Each group of items should include some common items for each group of examinees. Several types of test form arrangement are prepared applying common items and examinees when massive data matrices are taken into account. With common item parameter values set to be the same as inter-group estimates by putting constraints, analysis can express these items in the same scale. Thus, continuous item analysis without unnecessary burdens on examinees will develop an item scale in the same scale. Also, continuous surveys including common items can estimate item parameters one step at a time. This brings about great merit on test preparation and management.

If IRT permits scale development and operation of many AIDS knowledge understanding items in the same scale, a test will be insusceptible to biased difficulty of pre/post test and distribution of groups of examinees. The present study attempts to develop a scale of AIDS knowledge understanding items using IRT.

We employ a two-parameter model for analysis using IRT. Here, we do not use a three-parameter model which requires more examinees than the two-parameter model for stable estimates. The three-parameter model including a guessing parameter may be appropriate for this kind of item analysis, which has two-check factors. However, for our continuous future surveys, the three-parameter model might prove more difficult in estimating item parameters. Accordingly, the present study employs the two-parameter model and studies the result.

2.2. Data analysis applied from Multi group analysis

This study prepares two survey forms. They consist of items common to both forms and items specific to each form. The advantage of these survey forms is that they enable a wide variety of question items for examinees. Their disadvantage is the unfeasibility of comparison between groups if each
survey form (in each group of examinees) is analyzed separately. Therefore, the present study applies multi group analysis, which uses the framework of covariance structure analysis to denote all items in a common scale through common items. Then, it analyzes two non-identical survey forms. This method is described by Allison (1980) and Toyoda (1998) as factor analysis of data including missing.

The method constructs two check factor analysis models, in which items for two groups denote observed variables. That is, the study constructs one one-factor factor analysis model having the items of survey form A as one observed variable and the other one-factor factor analysis model having the items of survey form B as the other observed variable. The mean and variance of the two factors are held constant to be 0 and 1, respectively. The items common to both forms have a constraint that the factor pattern is the same between the factor models. That is, even if the survey forms have different make-up items, they can be discussed in a common scale in respect to factor patterns and factor scores when an equivalent constraint is set in scores and factor patterns of the common items.

Covariant structure analysis software MPlus is used for the analysis. Using this kind of software, multi-group estimation values can be computed concurrently by setting an equivalent constraint in inter-factor factor patterns. The present study adopts two one-factor models: one using the items of survey form A and the other, the items of survey form B. However, it is computed with a constraint that the estimating factor patterns are equivalent in the items common to survey form A and B, whereas, the factor analysis scale used in the different survey forms can be denoted in a common scale of equivalent factor patterns.

2.3. Survey Items

Test items measuring AIDS knowledge count a vast number of items of etiology, pathology, the route of infection, symptoms, treatment, and prevention. Here, we add newly created numeral items primarily including items of “application and assessment of multi-media related practical research in AIDS education” (Ohsawa, et al., 1998). The survey answer had two-check factors of “think so” and “do not think so” in each item.

The survey comprised a total of 34 items. The survey form had two types, type A and type B. Thirty-one items were chosen for each form of A and B from a total of 34 items. The items common to both forms accounted for 28 items.

2.4. Subjects and Time of Survey

The survey had 2462 student examinees from ten junior high schools in K City, Saitama Prefecture, Japan. Form A had 1274 examinees having 31 items while form B had 1188 examinees having 31 items. Each format is detailed in Table 1. The survey took place in December of 2002.

2.5. Data Coding

The survey outcomes were subjected to coding of “1” if the answer was correct and “0” if the answer was incorrect or unanswered. The data were then analyzed.

3. Results and Discussion

3.1. Correct Response Rate

From the survey, the correct response rate was calculated for each question (Table 1). The correct response rate differed between groups even though both had the same items. For example, items 6 and 17 had inter-group differences of more than 10%.

Theoretically, there should not be any survey items whose correct response rate was less than 50% because the questions had only two-check factors. The result, however, was that there was a much lower correct response rate of less than 50% in item 6. It might imply an important problem that “AIDS knowledge is incorrectly understood.”

3.2. Factor Patterns

As described in the previous section, factor patterns of obtained data were concurrently estimated between the groups of examinees by multi group analysis (Table 1). Here, the correct response rate was not separately calculated in each group but the common factor patterns were calculated between the groups A and B. As a whole, the factor patterns yielded more than 0.2 in many items. The patterns of the following items showed less than the other items:

10. HIV is not transmitted to others before the
Okubo, T., Ohsawa, S. and Nakagawa, M.

### Table 1 Correct response rate, factor pattern, discrimination, and difficulty in the items by survey form

<table>
<thead>
<tr>
<th>No.</th>
<th>Correct Response Rate (A)</th>
<th>Correct Response Rate (B)</th>
<th>Factor Pattern</th>
<th>Discrimination</th>
<th>Difficulty</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.596</td>
<td>0.561</td>
<td>0.170</td>
<td>0.248</td>
<td>-0.112</td>
</tr>
<tr>
<td>2</td>
<td>0.600</td>
<td>0.596</td>
<td>0.250</td>
<td>0.402</td>
<td>-0.224</td>
</tr>
<tr>
<td>3</td>
<td>0.651</td>
<td>0.737</td>
<td>0.310</td>
<td>0.766</td>
<td>-2.125</td>
</tr>
<tr>
<td>4</td>
<td>0.600</td>
<td>0.641</td>
<td>0.250</td>
<td>0.356</td>
<td>-1.358</td>
</tr>
<tr>
<td>5</td>
<td>0.788</td>
<td>0.796</td>
<td>0.300</td>
<td>0.404</td>
<td>-0.942</td>
</tr>
<tr>
<td>6</td>
<td>0.767</td>
<td>0.758</td>
<td>0.300</td>
<td>0.404</td>
<td>-0.942</td>
</tr>
<tr>
<td>7</td>
<td>0.706</td>
<td>0.699</td>
<td>0.250</td>
<td>0.356</td>
<td>-1.123</td>
</tr>
<tr>
<td>8</td>
<td>0.796</td>
<td>0.794</td>
<td>0.250</td>
<td>0.356</td>
<td>-1.123</td>
</tr>
<tr>
<td>9</td>
<td>0.751</td>
<td>0.758</td>
<td>0.300</td>
<td>0.404</td>
<td>-0.942</td>
</tr>
<tr>
<td>10</td>
<td>0.761</td>
<td>0.768</td>
<td>0.300</td>
<td>0.404</td>
<td>-0.942</td>
</tr>
<tr>
<td>11</td>
<td>0.767</td>
<td>0.768</td>
<td>0.300</td>
<td>0.404</td>
<td>-0.942</td>
</tr>
<tr>
<td>12</td>
<td>0.772</td>
<td>0.774</td>
<td>0.300</td>
<td>0.404</td>
<td>-0.942</td>
</tr>
<tr>
<td>13</td>
<td>0.788</td>
<td>0.796</td>
<td>0.300</td>
<td>0.404</td>
<td>-0.942</td>
</tr>
<tr>
<td>14</td>
<td>0.796</td>
<td>0.796</td>
<td>0.300</td>
<td>0.404</td>
<td>-0.942</td>
</tr>
<tr>
<td>15</td>
<td>0.801</td>
<td>0.800</td>
<td>0.300</td>
<td>0.404</td>
<td>-0.942</td>
</tr>
<tr>
<td>16</td>
<td>0.801</td>
<td>0.800</td>
<td>0.300</td>
<td>0.404</td>
<td>-0.942</td>
</tr>
<tr>
<td>17</td>
<td>0.801</td>
<td>0.800</td>
<td>0.300</td>
<td>0.404</td>
<td>-0.942</td>
</tr>
<tr>
<td>18</td>
<td>0.801</td>
<td>0.800</td>
<td>0.300</td>
<td>0.404</td>
<td>-0.942</td>
</tr>
<tr>
<td>19</td>
<td>0.801</td>
<td>0.800</td>
<td>0.300</td>
<td>0.404</td>
<td>-0.942</td>
</tr>
<tr>
<td>20</td>
<td>0.801</td>
<td>0.800</td>
<td>0.300</td>
<td>0.404</td>
<td>-0.942</td>
</tr>
<tr>
<td>21</td>
<td>0.801</td>
<td>0.800</td>
<td>0.300</td>
<td>0.404</td>
<td>-0.942</td>
</tr>
<tr>
<td>22</td>
<td>0.801</td>
<td>0.800</td>
<td>0.300</td>
<td>0.404</td>
<td>-0.942</td>
</tr>
<tr>
<td>23</td>
<td>0.801</td>
<td>0.800</td>
<td>0.300</td>
<td>0.404</td>
<td>-0.942</td>
</tr>
<tr>
<td>24</td>
<td>0.801</td>
<td>0.800</td>
<td>0.300</td>
<td>0.404</td>
<td>-0.942</td>
</tr>
<tr>
<td>25</td>
<td>0.801</td>
<td>0.800</td>
<td>0.300</td>
<td>0.404</td>
<td>-0.942</td>
</tr>
<tr>
<td>26</td>
<td>0.801</td>
<td>0.800</td>
<td>0.300</td>
<td>0.404</td>
<td>-0.942</td>
</tr>
<tr>
<td>27</td>
<td>0.801</td>
<td>0.800</td>
<td>0.300</td>
<td>0.404</td>
<td>-0.942</td>
</tr>
<tr>
<td>28</td>
<td>0.801</td>
<td>0.800</td>
<td>0.300</td>
<td>0.404</td>
<td>-0.942</td>
</tr>
<tr>
<td>29</td>
<td>0.801</td>
<td>0.800</td>
<td>0.300</td>
<td>0.404</td>
<td>-0.942</td>
</tr>
<tr>
<td>30</td>
<td>0.801</td>
<td>0.800</td>
<td>0.300</td>
<td>0.404</td>
<td>-0.942</td>
</tr>
<tr>
<td>31</td>
<td>0.801</td>
<td>0.800</td>
<td>0.300</td>
<td>0.404</td>
<td>-0.942</td>
</tr>
<tr>
<td>32</td>
<td>0.801</td>
<td>0.800</td>
<td>0.300</td>
<td>0.404</td>
<td>-0.942</td>
</tr>
<tr>
<td>33</td>
<td>0.801</td>
<td>0.800</td>
<td>0.300</td>
<td>0.404</td>
<td>-0.942</td>
</tr>
</tbody>
</table>

- Indicates no conduction (missing).

(○) is the item whose content is correct and is regarded as the right answer when the respondent answered "think so".

(×) is the item whose content is incorrect and is regarded as the right answer when the respondent answered "do not think so".

onset of AIDS;
13. Symptoms sometimes are not manifested even if one is infected with HIV;
14. The incubation period of AIDS is long before the onset, about seven to ten years on average;
25. The onset of AIDS can be delayed with early treatment.

These results indicate slight difference in the traits of these items as test items from those of many other items.

When a scale is composed only from used test items, items of low factor patterns are normally excluded. Considering the present study being part of subsequent surveys on AIDS knowledge items, their exclusion was disregarded even if the factor patterns are found to be low. Results obtained from future surveys will require close review.

More than half of the items of the present survey are associated with the route of infection. The factor patterns were likely to be influenced by the items of the route of infection. For our future study, we need to increase items about incubation period, onset, treatment, hosts, and tests. In successive item surveys, each item will be denoted in the same scale by making the best use of IRT while the items are regarded as the items for "overall AIDS-related knowledge understanding".

### 3.3. Estimation Parameters by Item Response Theory (IRT)

Basic statistics were computed for correct response rate and factor patterns.

An IRT analysis was performed under the assumption that AIDS knowledge understanding items can be expressed later in a single dimension. The present study applied a two-parameter model and used software Mplus for its analysis.

In the study, difficulty and discrimination were not estimated from each survey form but were concurrently estimated by the method applying multi-group analysis, as previously mentioned. Therefore, only a single estimate of difficulty and discrimination was obtained regardless of the number of group examinees.

Table 1 shows the difficulty and discrimination of each item. Items of high discrimination are:
1. HIV is transmitted by kissing AIDS-infected people.
2. HIV is transmitted by sharing juice.
3. HIV is not transmitted by sneezes.
4. HIV is transmitted by coughs from AIDS-infected people.
5. HIV may be transmitted if one eats food cooked by AIDS-infected people.
7. HIV is transmitted by eating food cooked by AIDS-infected people.
8. AIDS cannot be treated even at an early stage.
9. AIDS infection can be tested at public health centers.
10. HIV is transmitted to others before the onset of AIDS.
11. The disease is not transmitted from pets and other animals.
12. Symptoms sometimes are not manifested even if one is infected with HIV.
13. The incubation period of AIDS is long before the onset, about seven to ten years on average.
14. The body fluid that contains HIV virus most is blood.
15. The onset of AIDS can be delayed by early treatment.
16. Babies and old people do not get infected with HIV.
17. AIDS is transmitted through sexual intercourse.
18. AIDS may be transmitted when the blood of an infected patient touches a wound of a healthy individual.
19. The disease is not transmitted by sharing food.
20. There are effective vaccines for AIDS.
21. AIDS cannot be treated even at an early stage.
22. There is no effective treatment for AIDS.
23. There are effective vaccines for AIDS.
24. HIV is not transmitted by sharing juice.
25. The onset of AIDS can be delayed by early treatment.

These results indicate that primal “AIDS knowledge understanding” drawn from the item groups probably measured the AIDS knowledge area of route of infection. These are mostly associated with the incubation period, which may be relatively low in priority as test items. These results indicate that primal “AIDS knowledge understanding” drawn from the item groups probably measured the AIDS knowledge area of route of infection. These are mostly associated with the incubation period, which may be relatively low in priority as test items.

Next, the item traits were examined according to difficulty as shown in Table 1.

4. Summary

(1) This survey developed a scale by IRT for AIDS knowledge understanding items used in a test for junior high school students. It could partly clarify the item property on AIDS knowledge understanding, mainly on the route of infection, through estimation of item parameters using an IRT-based two-parameter model. Item characteristics depended on examinees in the conventional method. By the estimation of our item parameters, the following two points should be considered:

1. Following this survey, continuous surveys on AIDS knowledge understanding items should allow understanding of the characteristics among the aspects of AIDS knowledge including the route of infection, incubation period, onset, treatment, hosts, and tests. This can be accomplished by developing a common scale for the items.
2. More examination may be possible for such issues as teaching order by knowing the characteristics of each knowledge item.

(2) In difficulty, knowledge of the routes of infection from human to human showed low difficulty while the same routes of infection from...
animal or special area to human showed high difficulty. This suggests that under the current status they are less recognized than those of human-to-human infection.

(3) Items relating to animal-mediated routes of infection and AIDS treatment tended to show high difficulty.

(4) Knowledge level about infection was relatively low in onset and treatment.

(5) As a whole, the item parameters were estimated low because they were analyzed by a two-parameter model.

(6) The present study made a scale of items, mainly the routes of infection, by item response theory. Accordingly, it is important for future study to estimate parameters by increasing in balance items of incubation period, onset, treatment, hosts, tests, using “the other groups of examinees” to discuss them in a common scale. The present study, using the IRT-based two parameters to question items with two-check factors, had low estimation in difficulty. This will need to be reviewed. Further, we should survey more items besides the items estimated in the present study and develop scales for more items for children and young people of AIDS-endemic areas of Asian and African countries. There are many tasks that lie ahead.

References
Koiso, T., et al., (1999) Research by educational experiment regarding the educational effect of the class about AIDS in the health instruction at junior high school. – Comparative analysis of the educational effect regarding knowledge obtained from the class which used multimedia (CD-ROM) and from the one which used VTR -. Japanese Journal of school health. 41 : 153-167 (in Japanese with English abstract).