Development of pre-tests for measuring learning effects of computer-manipulation skills and knowledge and grouping Sports Science students by them

NAGAMINE Ken (Assistant Professor, Department of Health and Sports Science)
AOYAGI Osamu (Professor, Department of Health and Sports Science)
IKEDA Takahiro (Professor, Fukuoka Prefectural University)
IKEDA Tomoko (Part-time lecturer, Fukuoka University)

Abstract

In order to develop a test to measure the computer-manipulation skills and knowledge in the computer-skill improvement practice classes for the Sports Science major, we administered 20 items each for Word, Excel, and PowerPoint operation, for a total of 60 items for Sports Science major students. From the perspective of 1) the number of substantial options, 2) internal validity, 3) the improvement in the percentage of correct answers, and 4) unstable measurement, 26 test items satisfying these criteria were chosen. Then, from the perspective of increasing the reliability of the composite scores, Quantification Theory Type Three (QTTT) was implemented, and a practical scoring table consisting of simple integers was developed based on the category weights of the first dimension. This evaluation table is simple to calculate and can be easily calculated in class, and it is expected that this test will be used to make homogenous groups before classes and confirm the effects of learning.

1. Introduction

Since students who do not have academic ability above a certain passing score in English, Japanese, and other subjects for entrance examinations are not allowed to enter the school, classes can naturally be held for a homogeneous group with a certain level of academic ability or even higher. However, the computer-skill improvement practice classes at universities are conducted based on these as an extension of the computer skills and knowledge education at junior and senior high schools. Therefore, those who graduated from high schools with extensive computer-manipulation

education have a high degree of skills, but those who did not receive such education do not have a high level of skills or knowledge, which naturally leads to a gap between them(Ishizaki, 2009 [1] ; Yasui, 2017 [2]). When conducting practice, especially with groups of individuals with large differences, those with higher skills and knowledge take less time to complete certain tasks and are able to complete tasks on their own in a shorter period of time. However, those with limited skills and knowledge need a lot of assistance and require time to complete their assignments (Kayaki et al., 2008 [3]). If the class progression is adjusted to those with higher abilities, those with lower abilities will not be able to keep up, and if the class progression is adjusted to those with lower abilities, those with higher abilities will have too much time on their hands.

For this reason, as Okamoto (2007) [4] pointed out, this makes simultaneous teaching inefficient. Therefore, in the case of the Sports Science Department where the class is divided into several groups, it is desirable, as much as possible, for each group to be a homogeneous group. For this purpose, it is necessary to develop a test to determine the level of computer-manipulation skills and knowledge of students prior to taking the course. In addition, when evaluating the learning effect, if the students are evaluated in terms of their level of computer-manipulation skills and knowledge at the end of the unit, those who have higher skills and knowledge before the classes will receive a high evaluation even if they do not attend the class. Even if the student's skills and knowledge are low before the class, and if there is significant development after the class because the course is effective, his/her efforts should also be recognized. Although the test must consist of the same test items in order for it to be used to determine the effectiveness of learning, there are other necessary requirements.

Now, what requirements should such a test have in order to rate computer-manipulation ability in advance, and at the same time, measure the results of the class by testing again after the unit? First, if the test is a multiple-choice test (objective test; Watanabe, 1972 [5]) in which one correct answer is chosen from several alternatives, then distracters (options other than the correct answer) must be effectively handled. If the answers are concentrated on a particular option, then the test is ineffective because there is no distracter effect, and it is true even though the number of options is

increased. Then, the content of the options should be revised to increase the effectiveness of the distracter.

Next, all test items must measure the same abilities, i.e., computer-manipulation skills and knowledge. For example, if a person can answer a test item correctly without computer-manipulation skills or knowledge, then that item is measuring a different ability from what the other items are measuring. This ability can be examined by computing the correlation coefficient between the total number of correct answers and each item and examining whether it shows a significant positive correlation coefficient. If the correct and incorrect answers are determined by different abilities than the other items, the correlation will not be significant and positive.

Furthermore, in addition to the above general test requirements for diagnosis, a test to measure learning effects must also demonstrate certain learning effects if they are to be achieved. In other words, if the post-learning correct response rate is not higher than the pre-learning rate, the test is not measuring learning effects (Takeya, 1991〔6〕, pp. 107-113).

In addition to the overall change in correct and incorrect responses described above, we need to take note of the response pattern. First, those who answered correctly both before and after the class already have the necessary computer-manipulation skills and knowledge and do not need to learn. Second, those who answered incorrectly both before and after the unit were those who did not have the computer-manipulation skills and knowledge before the unit and, unfortunately, did not acquire them after the unit. Third, those who answered incorrectly before learning but correctly after learning are those who acquired computer-manipulation skills and knowledge only after learning.

These are all-natural response patterns. However, those who answered correctly before the unit but incorrectly after the unit are considered to have an unnatural pattern. The correct response before learning is considered to be an unnatural response, most likely due to cheating or a lucky guess. In any case, if there are many such unnatural responses, it is not a desirable test item.

Therefore, in this study, the test items were selected based on the following

criteria: 1) the distracters are valid and are not biased toward a particular option (the number of substantial options); 2) it measures the same computer-manipulation skills and knowledge as other test items (internal validity); 3) the correct response rate increases before and after the class (increase in correct response rate); and 4) those who answered correctly before class rarely answered incorrectly on the same item after class (unstable responses).

Furthermore, when evaluating the overall score, all items should not be assigned the same score, i.e., the score should not be based on the number of correct answers, etc., but rather on the characteristics of the items. The criterion for item selection in 2) has been found to increase the reliability of the total score (Ikeda, 1973 [7]). A similar point of view could be used to allot points based on the Quantification Theory Type Three (QTTT; Kobayashi, 1981 [8] ; Komazawa, 1982 [9]), which is a technique to make category weights between items that show similar patterns as equal as possible, and the resulting sample scores are known to be more reliable.

However, we use QTTT to find similar category weights among items that show similar patterns, but, usually, these values include a decimal point and are not constrained by a maximum or minimum value. Since these are inconvenient to handle in the classroom, the scores are converted to integer scores with a maximum value (e.g., 50 points) and a minimum value (e.g., 0 points), and a score allotment table is developed in a tabular format.

2. Methods

2.1 Test items

While referring Kawai et al. (2009) [10] , Ono et al. (2011) [11] and Mon (2009) [12] , a total of 60 test items were developed, 20 for each of about seven options, concerning the operation and knowledge of Word, Excel, and PowerPoint, which are the main units of the computer-skill improvement practice classes for a Sports Science major. The questions were based on the content actually covered in class by the teachers in charge of each unit. The 60 test items are listed in Table 1, along with their abbreviated names. The letter “W“ at the beginning of the abbreviated name indicates Word, “E“ indicates

Excel, and “P“ indicates PowerPoint. The questions were submitted via radio buttons on a Google Form, and the collected answers were compiled in a spreadsheet (Aoyagi and Nagamine, 2021 [13]).

2.2 Subjects

The subjects were 278 students of the Sports Science Department who took the computer-skill improvement practice classes for the Sports Science major in 2021 and 2022.

2.3. Test implementation

The subjects gathered in one classroom under the direction of one teacher with a set time limit for answering the questions, while the other teachers watched them. Therefore, it was not possible to consult with each other while answering the questions, as the answers were not given at each home, etc. Since the test was not conducted using paper and writing instruments, individuals could not be identified by their handwriting.

Table 1-1. Test items and their abbreviations

Test items	Abbreviations
01. Which tab to select when setting up a page?	W1.SETUP-PAGE
02. Which tabs to select for header and footer settings?	W2.HEADER
03. Which tab to you select for page numbering?	W3.NUMBERING
04. Which tab to select to set the page background?	W4.BACKGROUND
05. Which is the shortcut key for printing text?	W5.PRINTING
06. When you "cut" or "copy" a sentence, where is it temporarily saved?	W6.CUT&COPY
07. Which is the shortcut key for copying sentences?	W7.COPYING
08. Which is the shortcut key for pasting text?	W8.PASTING
09. Which tab is selected to search a word in text?	W9.WORD-SEARCH
10. Which tab to select to search the text?	W10.TEXT-SEARCH
11. Which is the shortcut key to select all sentences?	W11.ALL-SENTENCES
12. What is the name of the indentation setting that appears above the horizontal ruler?	W12.INDENTATION
13. Which to use to insert special effect characters?	W13.SPECIAL-EFFECT
14. When deleting a table, which method is used to delete all the characters in the table only?	W14.DELETING-TABLE
15. Which group in the Layout tab of the Table tool do you select to rearrange numbers/words in the tables?	W15.REARRANGE
16. Which method is used to automatically adjust table column widths?	W16.WIDTHS
17. What tabs can shapes be added from?	W17.SHAPE-ADD
18. How do you drag to create a shape with a 1:1 aspect ratio?	W18.1:1 ASPECT
19. What settings to make figures or shapes move freely?	W19.MOVE-FREELY
20. Which tab to select to apply the figure effect?	W20.FIGURE-EFFECT
21. Which is the shortcut key to search in an Excel book?	E1.SEARCH-IN-BOOK
22. How to adjust the width and height automatically in an Excel sheet?	E2.WIDTH/HIGHT
23. What is the name of the place where the button to change the book view is located?	E3.CHANGE-BOOK-VIEW
24. Which tab to print from?	E4.PRINT
25. Which is the way to save an Excel file as PDF?	E5.PDF
26. Which of the following copying methods can be used to increase or decrease the number of copies in a sequence in a regular manner?	E6.SEQUENCE
27. Which is the shortcut key to copy a cell?	E7.COPY-CELL
28. Which is the shortcut key for pasting cells?	E8.PASTING-CELL
29. Is the range selected while what button to press and hold to select cells that are far apart?	E9.SELECT-AWAY
30. What to change on the Home tab when changing the formatting of the numbers in the cells?	E10.HOME-TAB

Table 1-2. Test items and their abbreviations

Test items	Abbreviations
31. Which symbols are absolute references to utilize in functions?	E11.ABSOLUTE-REFERENCE
32. Which is input method other than direct typing of absolute references?	E12.DIRECT-TYPING
33. Which function does calculate the "sum"?	E13.SUM
34. Which function does calculate the "mean"?	E14.MEAN
35. Which function does calculate the "minimum value"?	E15.MIN
36. Which function does calculate the "maximum value"?	E16.MAX
37. Which function does convert lowercases of all English sentences to "uppercase"?	E17.UPPERCASE
38. Which function does convert uppercases of all English sentences to "lowercase"?	E18.LOWERCASE
39. Which tab is a graph inserted from?	E19.GRAPH-INSERT
40. What is the main purpose of a line chart?	E20.LINE-CHART
41. Which method is used to insert new slides?	P1.NEW-SLIDE
42. Which is method of changing the layout of the slides?	P2.LAYOUT
43. Which tabs are set to hide slides?	P3.HIDE-SLIDE
44. Which method is used to erase slides?	P4.ERASE-SLIDE
45. Which method is used to make multiple selections on a series of slides?	P5.MULTI-SELECT
46. Which method is used for multiple selection of slides that are far apart?	P6.APART-MULTI-SELECT
47. Which shortcut key is used to print slides?	P7.PRINT-SLIDE
48. What is called the frame already pre-displayed in the slide?	P8.PRE-DISPLAY
49. Which tabs do change text to word art style?	P9.WORD-ART
50. Which tab adds a figure?	P10.FIGIURE
51. Which method makes the figures resizable?	P11.FIGURE-RESIZABLE
52. Which tab adds the table?	P12.ADD-TABLE
53. Which tab does switch to the note master view?	P13.NOTE-MASTER
54. Which shortcut key undoes the previous operation?	P14.UNDO
55. Which shortcut key repeats the previous operation?	P15.REPEAT
56. What markings are attached on the slides to which animation has been added?	P16.ATTACHED-MARK
57. Which tabs rehearses before presentation?	P17.REHEARSE
58. Which button turns the screen black during the presentation?	P18.BLACK-SCREEN
59. Which button is pressed at the end of the presentation?	P19.END-PRESENTATION
60. What is called a diagram that combines multiple shapes to represent interrelationships of information in a visually clear manner?	P20.INTERRELATIONSHIP

2.4. Statistical analysis

2.4.1 Selection of valid test items

A well-written test should first be examined in terms of whether it is a valid test irrespective of time changes before it is examined in terms of whether it measures learning effects, which change over time. If the test is a multiple-choice test, test items for which the correct answer is easily found will result in a lack of ability to discriminate between those who are more knowledgeable and those who are less so. If the non-correct answer option, i.e. distracter, is not valid, the answers will be concentrated on the correct answer option. In other words, for a test item to be good, the answers must not be biased toward a particular option. The amount of information (entropy) when only certain options are correctly answered and the other options are unanswered can be expressed by Equation 1, and the number of substantial options (k) at that time can be expressed by Equation 2 (the number of substantial options; Takeya, 1991 [6] , p. 45).

$$H = -\Sigma pi \log s(pi) \quad \cdots\cdots (1)$$

$$k = 2^{H} \quad \cdots\cdots (2)$$

In addition, a good test item should measure the same abilities as other test items, e.g., computer-manipulation skills and knowledge. If the test item is not one that those who excel in the abilities answer correctly and those who do not answer incorrectly, the test item is measuring abilities other than the abilities it is trying to measure. From the statistical perspective, this examination can be conducted using the correlation coefficient between the total number of correct answers (total score) and the item (internal validity).

In addition, considering the criteria for a good test in terms of whether it measures learning effects, assuming that there are learning effects in the class unless the percentage of correct answers increases before and after the class, the test does not measure the learning effects. In other words, as a test to measure learning effectiveness, the percentage of correct answers must improve compared to before and after the class.

From the statistical perspective, this can be examined in terms of whether there was a significant improvement in the percentage of correct answers before and after the class (Kanji, 1999 [14] ; improvement in the percentage of correct answers).

Furthermore, when we focus on the pattern of correct and incorrect answers, assuming that the class has a learning effect, it is quite possible that each individual (1) answered correctly before and after class on the same test item regardless of the class effect; (2) answered incorrectly before and after class on the same test item regardless of the class effect; (3) answered incorrectly before class on the same test item, but answered correctly after class because the class effect allows students to understand the test item. However, (4) the same test item was answered correctly before the class but incorrectly after the class is not possible unless the errors are taught in the class. Therefore, a test item that exhibits many such patterns is considered to be unstable in measurement and therefore inadequate as a test item. From the statistical perspective, this can be examined from the ratio of the frequency of pattern (4) in the cross table consisting of correct and incorrect answers before and after the class (unstable measurement).

2.4.2 Development of a practical test-allotment table

Test items selected as good test items should also be scored according to their item characteristics, and from the perspective of increasing the reliability of the scoring, a scoring method based on the sum of the product of the category weights (sample scores) obtained using QTTT can be considered. Here, a simple scoring table is developed based on the category weights of the items selected as having the highest contribution. In other words, Equation (3) is transformed so that the maximum total of the category weights is, for example, 50 points and the minimum is 0 points, using the following Equation (Ohyama et al., 2016 [15]).

$$\frac{(w_i - wmin/n)}{(wmax - wmin)} \times 50 \quad \cdots\cdots (3)$$

where “win” indicates the category weight of item i, “wmax” indicates the maximum value of the sum of the category weights, “wmin” indicates the minimum value of the sum of the category weights, and “n” indicates the number of items.

However, since there is a decimal point and the calculation is complicated, the decimal point is rounded off to an integer. If the minimum value is not zero or the maximum value is not 50 when rounding off, fine adjustments are made for values with a large rounding range.

2.5 Ethical Considerations

We also explained to subjects that their personal information would be carefully managed/stored and that when publishing the results of the analysis, the obtained data would be statistically processed and published in such a way that individuals could not be identified.

3. Results

3.1 Selection of valid test items

Table 2 shows the entropy (H), which indicates the variability of option selection for each test item, and the number of substantial options in descending order. The number of substantial options was five in the 3-item range, 38 in the 2-item range, and 17 in the 1-item range, with no unit-by-unit differences. If the number of substantial options is less than two, it means that the answers are concentrated on one option (correct answer) and that the item is too easy for the test. Therefore, we decided not to consider any test item with the number of substantial options less than 2.0 as a valid test item.

Table 2. H values and the numbers of substantial options

Itrems	H	Number of substantial options
E3.CHANGE-BOOK-VIEW	1.767	3.4
P20.INTERRELATIONSHIP	1.720	3.3
W19.MOVE-FREELY	1.698	3.2
P2.LAYOUT	1.576	3.0
W15.REARRANGE	1.567	3.0
P9.WORD-ART	1.537	2.9
W16.WIDTHS	1.491	2.8
E6.SEQUENCE	1.490	2.8
W2.HEADER	1.476	2.8
P6.APART-MULTI-SELECT	1.474	2.8
W20.FIGURE-EFFECT	1.468	2.8
P8.PRE-DISPLAY	1.448	2.7
W12.INDENTATION	1.446	2.7
P11.FIGURE-RESIZABLE	1.446	2.7
W1.SETUP-PAGE	1.356	2.6
E2.WIDTH/HIGHT	1.350	2.5
P15.REPEAT	1.349	2.5
P5.MULTI-SELECT	1.344	2.5
E10.HOME-TAB	1.342	2.5
P14.UNDO	1.342	2.5
W3.NUMBERING	1.328	2.5
W18.1:1 ASPECT	1.327	2.5
E1.SEARCH-IN-BOOK	1.312	2.5
P3.HIDE-SLIDE	1.312	2.5
E9.SELECT-AWAY	1.274	2.4
W9.WORD-SEARCH	1.252	2.4
P1.NEW-SLIDE	1.247	2.4
P13.NOTE-MASTER	1.234	2.4
E17.UPPERCASE	1.230	2.3
W14.DELETING-TABLE	1.197	2.3
W13.SPECIAL-EFFECT	1.190	2.3
E20.LINE-CHART	1.181	2.3
E11.ABSOLUTE-REFEREN	1.117	2.2
P16.ATTACHED-MARK	1.111	2.2
E12.DIRECT-TYPING	1.093	2.1
E18.LOWERCASE	1.090	2.1
E8.PASTING-CELL	1.087	2.1
P12.ADD-TABLE	1.083	2.1
W4.BACKGROUND	1.080	2.1
E4.PRINT	1.037	2.1
E5.PDF	1.037	2.1
P19.END-PRESENTATION	1.006	2.0
P4.ERASE-SLIDE	0.995	2.0
P7.PRINT-SLIDE	0.914	1.9
W5.PRINTING	0.900	1.9
E7.COPY-CELL	0.852	1.8
W8.PASTING	0.834	1.8
W11.ALL-SENTENCES	0.824	1.8
W6.CUT&COPY	0.822	1.8
P18.BLACK-SCREEN	0.754	1.7
W10.TEXT-SEARCH	0.726	1.7
P10.FIGIURE	0.691	1.6
W7.COPYING	0.625	1.5
E13.SUM	0.624	1.5
E19.GRAPH-INSERT	0.606	1.5
W17.SHAPE-ADD	0.511	1.4
P17.REHEARSE	0.456	1.4
E14.MEAN	0.378	1.3
E16.MAX	0.252	1.2
E15.MIN	0.240	1.2

Note. Items are sorted in descending order by H value.

In order to examine the internal validity, Table 3 shows the results of the correlation coefficients between the total score (number of correct answers) and each item in the order from the highest correlation coefficient to the lowest. E1.SEARCH-IN-BOOK, which showed the highest correlation, 0.731, was significant at the 1% level, indicating that those who answered correctly on this test item tended to answer

correctly with greater frequency for other items also, and those who answered incorrectly tended to answer correctly with less frequency also on other items very strongly. SEQUENCE and thereafter also show this tendency according to the value of their correlation coefficients. However, P12.ADD-TABLE and P9.WORD.ART, for which the correlation coefficients were not significant, did not show such a tendency, indicating that the total number of correct or incorrect responses was not related to the correct or incorrect responses for these two items and that the correctness of these two items was determined by computer-manipulation skills and knowledge different from the other items. Therefore, we decided not to consider them valid test items.

Table 4 shows the results of the tests of the pre- and post-unit correct response percentages for the test items, the improvement in the percentage of correct responses, and the difference between the percentages, in descending order of the value of “u.” Generally, a 10% increase in the percentage of correct responses is significant. However, the 25 items with an improvement percentage lower than it did not show significant development, so we can decide that they do not measure that learning effect if we assume that there was a learning effect in the class. Therefore, we decided not to consider them valid test items.

Table 5 shows the frequency of unstable answer patterns and their percentages in ascending order of percentage. Although it is difficult to determine what criteria should be used to select an unstable percentage, in this study, we used 30% as the criterion and did not consider anything higher than that with more than 30% unstable percentage to be a valid test item.

Table 3. The correlation coefficients between total scores and item scores

Item	Correlation coefficients
E1.SEARCH-IN-BOOK	0.731 **
E6.SEQUENCE	0.730 **
P3.HIDE-SLIDE	0.712 **
W20.FIGURE-EFFECT	0.700 **
P7.PRINT-SLIDE	0.693 **
P15.REPEAT	0.680 **
W3.NUMBERING	0.674 **
P14.UNDO	0.671 **
P5.MULTI-SELECT	0.663 **
E11.ABSOLUTE-REFERENCE	0.624 **
P13.NOTE-MASTER	0.602 **
E17.UPPERCASE	0.599 **
E8.PASTING-CELL	0.583 **
E12.DIRECT-TYPING	0.555 **
W13.SPECIAL-EFFECT	0.553 **
E18.LOWERCASE	0.551 **
P19.END-PRESENTATION	0.546 **
W18.1:1 ASPECT	0.533 **
E7.COPY-CELL	0.526 **
W11.ALL-SENTENCES	0.525 **
P16.ATTACHED-MARK	0.516 **
E13.SUM	0.503 **
W8.PASTING	0.503 **
P6.APART-MULTI-SELECT	0.498 **
W2.HEADER	0.489 **
W1.SETUP-PAGE	0.485 **
W5.PRINTING	0.484 **
E10.HOME-TAB	0.482 **
W10.TEXT-SEARCH	0.475 **
W9.WORD-SEARCH	0.457 **

Item	Correlation coefficients
W19.MOVE-FREELY	0.453 **
W7.COPYING	0.438 **
E5.PDF	0.430 **
E9.SELECT-AWAY	0.428 **
P18.BLACK-SCREEN	0.428 **
E19.GRAPH-INSERT	0.420 **
W15.REARRANGE	0.398 **
E14.MEAN	0.392 **
E16.MAX	0.373 **
P1.NEW-SLIDE	0.368 **
E20.LINE-CHART	0.367 **
W17.SHAPE-ADD	0.359 **
P8.PRE-DISPLAY	0.358 **
E3.CHANGE-BOOK-VIEW	0.358 **
P17.REHEARSE	0.357 **
E15.MIN	0.354 **
E2.WIDTH/HIGHT	0.348 **
P2.LAYOUT	0.347 **
W6.CUT&COPY	0.338 **
W14.DELETING-TABLE	0.336 **
W12.INDENTATION	0.327 **
P11.FIGURE-RESIZABLE	0.323 **
P10.FIGIURE	0.318 **
P20.INTERRELATIONSHIP	0.299 **
E4.PRINT	0.284 **
W16.WIDTHS	0.282 **
W4.BACKGROUND	0.268 **
P4.ERASE-SLIDE	0.235 **
P12.ADD-TABLE	0.124
P9.WORD-ART	0.057

Note1. Items are sorted in descending order by correlation coefficient.

Note2. ** indicates significance at the 1% significance level

Table 4. The improvement of correct answer rates

Items	the percentage of correct answer in the pre-test	the percentage of correct answer in the pre-test	The significant improvement of correct-answer percentages	uo
P14.UNDO	0.407	0.907	0.500	9.122 **
P3.HIDE-SLIDE	0.395	0.895	0.500	9.049 **
P15.REPEAT	0.407	0.872	0.465	8.387 **
P19.END-PRESENTATION	0.570	0.942	0.372	7.501 **
P5.MULTI-SELECT	0.419	0.837	0.418	7.490 **
P7.PRINT-SLIDE	0.628	0.965	0.337	7.249 **
P13.NOTE-MASTER	0.488	0.860	0.372	6.873 **
P16.ATTACHED-MARK	0.547	0.837	0.290	5.440 **
P8.PRE-DISPLAY	0.244	0.547	0.303	5.367 **
P11.FIGURE-RESIZABLE	0.395	0.686	0.291	5.057 **
E12.DIRECT-TYPING	0.587	0.840	0.253	4.846 **
P18.BLACK-SCREEN	0.733	0.930	0.197	4.558 **
P6.APART-MULTI-SELECT	0.360	0.616	0.256	4.435 **
E6.SEQUENCE	0.280	0.527	0.247	4.360 **
W12.INDENTATION	0.300	0.533	0.233	4.093 **
W13.SPECIAL-EFFECT	0.553	0.747	0.194	3.522 **
P20.INTERRELATIONSHIP	0.221	0.407	0.186	3.471 **
E11.ABSOLUTE-REFERENCE	0.567	0.753	0.186	3.400 **
W20.FIGURE-EFFECT	0.353	0.547	0.194	3.377 **
P2.LAYOUT	0.302	0.488	0.186	3.295 **
W16.WIDTHS	0.400	0.580	0.180	3.118 **
W8.PASTING	0.733	0.873	0.140	3.048 **
W6.CUT&COPY	0.727	0.860	0.133	2.845 **
W7.COPYING	0.800	0.913	0.113	2.791 **
W10.TEXT-SEARCH	0.767	0.887	0.120	2.747 **
W11.ALL-SENTENCES	0.747	0.867	0.120	2.633 **
P1.NEW-SLIDE	0.628	0.767	0.139	2.621 **
E2.WIDTH/HIGHT	0.473	0.620	0.147	2.557 *
E17.UPPERCASE	0.560	0.700	0.140	2.511 *
W5.PRINTING	0.720	0.840	0.120	2.509 *

Items	the percentage of correct answer in the pre-test	the percentage of correct answer in the pre-test	The significant improvement of correct-answer percentages	uo
P17.REHEARSE	0.860	0.942	0.082	2.378 *
E5.PDF	0.667	0.780	0.113	2.188 *
W9.WORD-SEARCH	0.467	0.593	0.126	2.186 *
P4.ERASE-SLIDE	0.709	0.814	0.105	2.134 *
W19.MOVE-FREELY	0.333	0.447	0.114	2.024 *
E4.PRINT	0.600	0.707		1.947
W18.1:1 ASPECT	0.533	0.633		1.756
E9.SELECT-AWAY	0.493	0.593		1.738
E1.SEARCH-IN-BOOK	0.580	0.673		1.665
W17.SHAPE-ADD	0.853	0.913		1.617
E7.COPY-CELL	0.760	0.833		1.570
E13.SUM	0.807	0.873		1.559
W1.SETUP-PAGE	0.493	0.580		1.511
E8.PASTING-CELL	0.673	0.747		1.412
W14.DELETING-TABLE	0.567	0.640		1.292
E3.CHANGE-BOOK-VIEW	0.273	0.333		1.131
E14.MEAN	0.900	0.933		1.033
W15.REARRANGE	0.280	0.333		0.996
W4.BACKGROUND	0.587	0.627		0.709
E18.LOWERCASE	0.673	0.707		0.637
E10.HOME-TAB	0.553	0.587		0.595
W3.NUMBERING	0.567	0.600		0.580
E15.MIN	0.933	0.947		0.511
E16.MAX	0.933	0.933		0.000
P9.WORD-ART	0.302	0.302		0.000
P10.FIGIURE	0.837	0.837		0.000
E19.GRAPH-INSERT	0.860	0.820		-0.945
W2.HEADER	0.540	0.480		-1.039
E20.LINE-CHART	0.700	0.593		-1.938
P12.ADD-TABLE	0.698	0.500		-3.499

Note1. Items are sorted in descending order by uo value.

Note2. * and ** indicate significance at the 5% and 1% significance level, respectively.

Table 5. The number and percentage of unstable responses

Items	the number of unstable responses	the percentage of unstable responses (%)
P15.REPEAT	1	1.16
P19.END-PRESENTATION	2	2.33
P3.HIDE-SLIDE	3	3.49
P7.PRINT-SLIDE	3	3.49
P14.UNDO	3	3.49
E15.MIN	7	4.67
E16.MAX	8	5.33
P17.REHEARSE	5	5.81
P18.BLACK-SCREEN	5	5.81
P20.INTERRELATIONSHIP	5	5.81
W17.SHAPE-ADD	9	6.00
W6.CUT&COPY	10	6.67
W7.COPYING	10	6.67
E14.MEAN	10	6.67
P8.PRE-DISPLAY	6	6.98
P11.FIGURE-RESIZABLE	6	6.98
P16.ATTACHED-MARK	6	6.98
W10.TEXT-SEARCH	12	8.00
P13.NOTE-MASTER	8	9.30
W8.PASTING	14	9.33
E7.COPY-CELL	14	9.33
E12.DIRECT-TYPING	14	9.33
E13.SUM	14	9.33
W11.ALL-SENTENCES	15	10.00
W13.SPECIAL-EFFECT	15	10.00
W20.FIGURE-EFFECT	15	10.00
P2.LAYOUT	9	10.47
P5.MULTI-SELECT	9	10.47
P6.APART-MULTI-SELECT	9	10.47
W5.PRINTING	16	10.67

Items	the number of unstable responses	the percentage of unstable responses (%)
W12.INDENTATION	16	10.67
W16.WIDTHS	17	11.33
E6.SEQUENCE	17	11.33
E11.ABSOLUTE-REFERENCE	18	12.00
E5.PDF	19	12.67
E19.GRAPH-INSERT	19	12.67
P4.ERASE-SLIDE	11	12.79
P10.FIGIURE	11	12.79
W15.REARRANGE	21	14.00
E2.WIDTH/HIGHT	21	14.00
E3.CHANGE-BOOK-VIEW	21	14.00
E4.PRINT	22	14.67
E17.UPPERCASE	23	15.33
W9.WORD-SEARCH	24	16.00
W19.MOVE-FREELY	24	16.00
P1.NEW-SLIDE	14	16.28
P9.WORD-ART	14	16.28
W1.SETUP-PAGE	25	16.67
E8.PASTING-CELL	25	16.67
E1.SEARCH-IN-BOOK	26	17.33
W14.DELETING-TABLE	27	18.00
E18.LOWERCASE	27	18.00
W18.1:1 ASPECT	28	18.67
W4.BACKGROUND	29	19.33
E9.SELECT-AWAY	32	21.33
E10.HOME-TAB	33	22.00
E20.LINE-CHART	33	22.00
W3.NUMBERING	34	22.67
W2.HEADER	45	30.00
P12.ADD-TABLE	32	37.21

Note1. Items are sorted in ascending order by the percentage of unstable responses.

Note2. The number of unstable responses does not correspond to the percentage of them because of missing values.

Table 6 lists whether the four types of criteria we have examined so far have been met or not. The 26 test items that met all four criteria are W9.WORD-SEARCH, W12.INDENTATION, W13.SPECIAL-EFFECT, W16.WIDTHS, W19.MOVE-FREELY, W20.FIGURE-EFFECT, E2.WIDTH/HIGHT, E5.PDF, E6.SEQUENCE, E11.ABSOLUTE-REFERENCE, E12.DIRECT-TYPING, E17.UPPERCASE, P1.NEW-SLIDE, P2.LAYOUT, P3.HIDE-SLIDE, P4.ERASE-SLIDE, P5.MULTI-SELECT, P6.APART-MULTI-SELECT, P8.PRE-DISPLAY, P11.FIGURE-RESIZABLE, P13.NOTE-MASTER, P14.UNDO, P15.REPEAT, P16.ATTACHED-MARK, P19.END-

PRESENTATION, P20.INTERRELATIONSHIP. In the following sections, we will examine how to measure computer-manipulation skills and knowledge using these 26 items. The final selections include six items related to Word and six items related to Excel, but PowerPoint had a significant number, as many as 14 items.

3.2 Development of a practical test-allotment table

First, in order to determine the score allotment of these 26 items, QTTT was performed in terms of obtaining more reliable scale values. The obtained eigenvalues from the first to the fifth dimension were 0.3263, 0.0574, 0.0516, 0.0505, and 0.0442, and their percent contributions were 32.63%, 5.74%, 5.16%, 5.05%, and 4.42%. In other words, the contribution of the first dimension was significantly larger than that of the second and subsequent dimensions, and the first dimension explained a considerable amount of the variance. Therefore, we decided to develop a score allotment table using the category weights for the first dimension only. The obtained category weights are shown on the left side of Figure 1 in a tabular format with two columns made up of correct and incorrect. All category weights for correct answers were smaller than those for wrong answers.

Since the direction of correctness is arbitrary, the signs of correct and incorrect answers will be reversed from now on so that the allotment of scores for correct answers is higher than that for incorrect answers. The table at the center of Figure 1 was transformed so that the maximum total category weight is 50 points and the minimum is 0 points according to Equation 1. Here, the arbitrary maximum total score of 50 points is an example of assuming a range that can be handled by mental arithmetic. However, since there is a decimal point and the calculation is complicated, the decimal point was rounded off to an integer. Since the minimum value was not zero when rounding, fine adjustments were made for values with a large rounding range.

Table 6. The summary table of whether 4 criteria have been met or not

Items	The number of criteria met	the numbers of substantial options	Correlation with total scores	Improvement of correct answers	Unstable responses
W1.SETUP-PAGE	3	○	○		○
W2.HEADER	2	○	○		
W3.NUMBERING	3	○	○		○
W4.BACKGROUND	3	○	○		○
W5.PRINTING	3		○	○	○
W6.CUT&COPY	3		○	○	○
W7.COPYING	3		○	○	○
W8.PASTING	3		○	○	○
W9.WORD-SEARCH	4	○	○	○	○
W10.TEXT-SEARCH	3		○	○	○
W11.ALL-SENTENCES	3		○	○	○
W12.INDENTATION	4	○	○	○	○
W13.SPECIAL-EFFECT	4	○	○	○	○
W14.DELETING-TABLE	3	○	○		○
W15.REARRANGE	3	○	○		○
W16.WIDTHS	4	○	○	○	○
W17.SHAPE-ADD	2		○		○
W18.1:1 ASPECT	3	○	○		○
W19.MOVE-FREELY	4	○	○	○	○
W20.FIGURE-EFFECT	4	○	○	○	○
E1.SEARCH-IN-BOOK	3	○	○		○
E2.WIDTH/HIGHT	4	○	○	○	○
E3.CHANGE-BOOK-VIEW	3	○	○		○
E4.PRINT	3	○	○		○
E5.PDF	4	○	○	○	○
E6.SEQUENCE	4	○	○	○	○
E7.COPY-CELL	2		○		○
E8.PASTING-CELL	3	○	○		○
E9.SELECT-AWAY	3	○	○		○
E10.HOME-TAB	3	○	○		○

Items	The number of criteria met	the numbers of substantial options	Correlation with total scores	Improvement of correct answers	Unstable responses
E11.ABSOLUTE-REFERENCE	4	○	○	○	○
E12.DIRECT-TYPING	4	○	○	○	○
E13.SUM	2		○		○
E14.MEAN	2		○		○
E15.MIN	2		○		○
E16.MAX	2		○		○
E17.UPPERCASE	4	○	○	○	○
E18.LOWERCASE	3	○	○		○
E19.GRAPH-INSERT	2		○		○
E20.LINE-CHART	3	○	○		○
P1.NEW-SLIDE	4	○	○	○	○
P2.LAYOUT	4	○	○	○	○
P3.HIDE-SLIDE	4	○	○	○	○
P4.ERASE-SLIDE	4	○	○	○	○
P5.MULTI-SELECT	4	○	○	○	○
P6.APART-MULTI-SELECT	4	○	○	○	○
P7.PRINT-SLIDE	3		○	○	○
P8.PRE-DISPLAY	4	○	○	○	○
P9.WORD-ART	2	○			○
P10.FIGIURE	2		○		○
P11.FIGURE-RESIZABLE	4	○	○	○	○
P12.ADD-TABLE	1	○			
P13.NOTE-MASTER	4	○	○	○	○
P14.UNDO	4	○	○	○	○
P15.REPEAT	4	○	○	○	○
P16.ATTACHED-MARK	4	○	○	○	○
P17.REHEARSE	3		○	○	○
P18.BLACK-SCREEN	3		○	○	○
P19.END-PRESENTATION	4	○	○	○	○
P20.INTERRELATIONSHIP	4	○	○	○	○

Note1. ○ indicates that the criterion has been met.

Note2. Shading indicates that all four criteria are met.

	Category weights		The score table of a full mark with 50 points		The score table by integer after fine adjustment	
Items	Correct	Incorrect	Correct	Incorrect	Correct	Incorrect
W9.WORD-SEARCH	-0.742	0.975	1.836	0.183	2	0
W12.INDENTATION	-0.902	0.652	1.990	0.494	2	0
W13.SPECIAL-EFFECT	-0.726	1.622	1.821	-0.440	2	0
W16.WIDTHS	-0.765	0.668	1.858	0.479	2	0
W19.MOVE-FREELY	-0.901	0.748	1.989	0.402	2	0
W20.FIGURE-EFFECT	-1.156	1.346	2.234	-0.175	2	0
E2.WIDTH/HIGHT	-0.716	0.849	1.812	0.304	2	0
E5.PDF	-0.474	1.559	1.578	-0.379	1	0
E6.SEQUENCE	-1.500	1.355	2.566	-0.182	2	0
E11.ABSOLUTE-REFERENCE	-0.875	1.953	1.964	-0.759	2	-1
E12.DIRECT-TYPING	-0.840	2.118	1.930	-0.917	2	-1
E17.UPPERCASE	-0.741	1.721	1.835	-0.535	2	-1
P1.NEW-SLIDE	-0.551	0.862	1.652	0.292	2	0
P2.LAYOUT	-0.723	0.479	1.818	0.661	2	1
P3.HIDE-SLIDE	-1.103	1.608	2.183	-0.427	2	0
P4.ERASE-SLIDE	-0.162	0.437	1.278	0.701	1	0
P5.MULTI-SELECT	-1.074	1.513	2.156	-0.335	2	0
P6.APART-MULTI-SELECT	-0.965	0.774	2.051	0.377	2	0
P8.PRE-DISPLAY	-1.057	0.532	2.140	0.610	2	1
P11.FIGURE-RESIZABLE	-0.489	0.524	1.593	0.617	2	1
P13.NOTE-MASTER	-0.817	1.477	1.908	-0.301	2	0
P14.UNDO	-1.044	1.606	2.127	-0.424	2	0
P15.REPEAT	-1.014	1.587	2.098	-0.406	2	0
P16.ATTACHED-MARK	-0.694	1.256	1.790	-0.088	2	0
P19.END-PRESENTATION	-0.619	1.737	1.717	-0.550	2	-1
P20.INTERRELATIONSHIP	-0.992	0.338	2.077	0.796	2	1
Total	-21.642	30.294	50.000	0.000	50	0

Figure 1. The process of making the score list

The final allotment table is shown on the right side of the figure. The larger the difference between scores of correct and incorrect answers on the obtained score allotment table, the more the total score changes depending on whether the correct or incorrect answer is given, so the difference between correct and incorrect answers is important for discriminating computer manipulating skills and knowledge. The largest difference was three points, which was found in as many as three items for Excel, compared to one item for Word and one item for PowerPoint. Therefore, performance for Excel was important in discriminating overall computer-manipulation skills and knowledge.

4. Discussion

4.1 Reason why so many test items related to PowerPoint were selected

Many test items related to PowerPoint were selected as test items for evaluation. It is considered that this is because in many cases PowerPoint is not installed even when standard Office is loaded onto students' notebook PCs, and there are fewer opportunities to actually operate PowerPoint at home as compared to Word or Excel, so the computer-manipulation skills and knowledge may be greatly improved in the class.

4.2 Reasons why the large difference in allotment scores for Excel was found

The test item with the largest difference in the allotment of correct and incorrect scores (three points) was the most common for Excel compared to the other software. About this point, Ishizaki and Masumoto (2014)〔16〕 also pointed out rapid improvement of manipulating skills in Excel. As Ishizaki (2013)〔17〕 pointed out, this may be because, unlike Word and PowerPoint, Excel requires knowledge of mathematics when using functions, and there are more unexplored areas in computer-manipulation education up to the high school level than for Word and other programs.

For this reason, the correct responses to those test items are considered to be highly valued for computer-manipulation skills and knowledge.

4.3 Usage of this test

As described above, we were able to develop a practical 26-item test to determine computer-manipulation skills and knowledge. This allows grouping based on their performance on the test given before the start of the unit, resulting in a homogeneous group with no differences in computer-manipulation skills or knowledge. As Nagai et al.(2005)〔18〕 pointed out, this is expected to enable the implementation of homogeneous class content and stress-free teaching and learning for both teachers and students. In addition, by comparing pre- and post-unit tests, the degree of learning in the class can be determined. This makes it possible to obtain individualized learning effects in substantive computer- manipulation skills and knowledge through the classes.

References

（1）Ishizaki, R. (2009) Computer Skills and Computer Literacy Education at Entry Level for New Students in the Faculty of Humanities and Social Sciences, Fukuoka Prefectural University. Bulletin of the Faculty of Human Sociology, Fukuoka Prefectural University 18(1): 43-60. (in Japanese)

（2）Yasui, H. (2017) Computer-related Environment Survey of Freshmen 2016. Bulletin of Tokyo City University, Common Education 10: 97 - 115. (in Japanese)

（3）Kayaki, N., Ueda, C. and Wakabayashi, Y. (2008) The Influence of Awareness of Others and Confidence Level in Simultaneous Literacy Classes. Computers and Human Society, Proceedings of the National Conference of Japan Information Processing Society 70: 459-460. (in Japanese)

（4）Okamoto, R. (2007) A Case Study of Information Literacy Education: A Review of the First Semester Classes at Sagami Women's University. Research Report of Japan Information Processing Society 101(1): 1-8. (in Japanese)

（5）Watanabe, K. (1972) Multiple Choice Format. In N. Hidano (Ed.) Psychological Research Methods, Volume 7: Test I. Tokyo University Press, Inc.: Tokyo, pp. 42-52. (in Japanese)

（6）Takaya, M. (1991) New Test Theory: Structural Analysis of Educational Data. Waseda University Press, Inc.: Tokyo. (in Japanese)

（7）Ikeda, H. (1973) Psychological Research Methods Volume 6 Test2. Tokyo University Press, Inc. : Tokyo, pp. 237-241. (in Japanese)

（8）Kobayashi, R. (1981) Introduction to Quantification Theory. Nikka-Giren: Tokyo. (in Japanese)

（9）Komazawa, T. (1982) Statistics Library, Quantification Theory and Data Analysis. Asakura Shoten: Tokyo. (in Japanese)

（10）Kawai, H., Fukuyama, H., Iwase, H. and Handa, K. (2009) Survey on Computer Literacy of Freshmen at Tokyo Seitoku University. Bulletin of Tokyo Seitoku University, Faculty of Humanities and Faculty of Applied Psychology16: 59-71. (in Japanese)

（11）Ono, K., Matsumoto, Y., Nagakwa, K. and Yamada, S. (2011) Computer

Manipulating Education in Junior College. Bulletin of Nagoya Bunka Junior College 36: 1-7. (in Japanese)

（12） Mon, M. (2009) Computer Manipulating Education in Junior Colleges. Bulletin of Osaka Shoin Women's University, Human Sciences, 8: 263-273. (in Japanese)

(13) Aoyagi, O. and Nagamine, K. (2021) Statistical Analysis of Questionnaire Surveys in Physical Education Students' Theses with Google Forms and Excel. Toka-Shobo: Fukuoka. (in Japanese)

（14） Kanji, G. K. (1999) 100 Statistical tests. Sage Publications: London, p. 25.

（15） Ohyama, Y., Aoyagi, O., Yaita, A., Komure, I. and Kawazura, T. (2016) The Evaluation of Basketball Group Offensive Skills in College Physical Education Classes. Japanese Journal of Basketball Research 2: 1-14. (in Japanese with English abstract)

（16） Ishizaki, R. and Masumoto, K. (2014) Effectiveness of Computer Literacy Education in the Faculty of Humanities and Social Sciences, Fukuoka Prefectural University 2013 . Bulletin of the Faculty of Human Sociology, Fukuoka Prefectural University 23(1): 37-57. (in Japanese)

（17） Ishizaki, R. (2013) Educational Effectiveness of Statistical Analysis Practice in the Faculty of Human Sociology, Fukuoka Prefectural University . Bulletin of the Faculty of Human Sociology, Fukuoka Prefectural University 22(2): 117-132. (in Japanese)

（18） Nagai, M., Shimizu, D., Okuda, T. and Yamaguchi, E. (2005) A Study of Classification by Computer Manipulating Ability Based on Students' Questionnaire in Computer Literacy Class. Proceedings of the Japan Society for Educational Technology 29(Suppl.) 225-228. (in Japanese)

和文抄録

スポーツ科学部生のための情報処理事前スキルや知識による グルーピングと授業成果測定のためのテスト作成

長嶺健（スポーツ科学部助教）
青柳領（スポーツ科学部教授）
池田孝博（福岡県立大学教授）
池田知子（福岡大学非常勤講師）

スポーツ科学の専門科目であるスポーツ情報処理実習において、その情報処理スキルや知識を測定するテストを作成するために、Word、Excel、 PowerPoint の操作に関する各 20 項目、計 60 項目をスポーツ科学生に実施した。そして、①実質選択肢数、②内的妥当性、③正答率の向上、④不安定な測定という観点から、それらの基準を満たす 26 項目を選んだ。そして、評価の信頼性を上げるという観点から数量化理論Ⅲ類を実施し、その第 1 次元のカテゴリーウエイトに基づいた。簡便な整数からなる配点表を作成した。この評価表は計算が簡便で授業中でも容易に計算でき、このテストを利用して、事前のグルーピングや事後の学習効果の確認に活用することが期待できると考えられる。