CrossMark

Detection and Treatment of Fake Math-Dislikes among Japanese Junior High School Students

Akitoshi Uchida[1] · Kazuo Mori[2,3]

Abstract International surveys repeatedly showed that Japanese children were good at mathematics but disliking it. We hypothesized there were a considerable number of "fake math-dislikes" among Japanese students who claimed they disliked mathematics explicitly but accepted it implicitly. To examine this hypothesis, we administered questionnaires and paper-based implicit association tests (IAT) to 204 Japanese junior high school students (13–14 years old) and found 38 fake math-dislikes (Study 1). We hypothesized further that those fake math-dislikes would become real math-dislikes eventually and that informing of their implicit attitude toward mathematics might work preventing this undesirable transition. Then, in Study 2, we randomly assigned them into experimental and control groups and informed only the experimental students of their positive implicit attitude toward mathematics we revealed with the IAT. One year later, we found 15 of the 16 experimental students improved their math achievement scores while only 9 of the 17 control students did. The simple practice of informing of their implicit attitude worked effectively for improving their math achievement. As hypothesized, it prevented the fake math-dislike students from turning into real math-dislikes.

Keywords Academic achievement · Implicit and explicit assessments · Math-dislikes · Paper-based implicit association test · Sex differences

✉ Kazuo Mori
    kazuo.mori@t.matsu.ac.jp

1   Sairyo Junior High School, Nagano 380-0913, Japan

2   Tokyo University of Agriculture and Technology, Koganei 184-8588, Japan

3   Department of School Education, Matsumoto University, 2095-1, Niimura, Matsumoto 390-1295, Japan

🅐 Springer

Introduction

Japanese Children Were Good at Math But Answered They Dislike Math

The International Association for the Evaluation of Educational Achievement (IEA) is an independent, international cooperative of national research institutions and governmental research agencies. It has conducted a series of international assessments of educational achievement in mathematics and science, the Trends in International Mathematics and Science Study (TIMSS) since 1995. Fourth- and eighth-grade Japanese children have participated in each cycle of the TIMSS assessments since the start year (TIMSS1995). According to the report from the National Institute for Educational Policy Research (NIER) of Japan on the results of TIMSS2011, both grades of Japanese children were among the top group along with the children in Singapore, South Korea, and Hong Kong (National Institute for Educational Policy Research, 2012). Even the score of the Japanese child in the 25th-percentile rank was higher than the TIMSS average score. As for the fourth - graders, their scores were improving from the former scores in TIMSS1995, 2003, and 2007. Meanwhile, the scores of the eighth - graders remained the same level as those of TIMSS2003 and 2007 but declined from those of TIMSS1995 and 1999.

Japanese children are good at mathematics but do not like mathematics. The TIMSS2011 included a student questionnaire asking about attitudes toward learning mathematics and science. As for mathematics, the participant children were asked to answer the question, "Do you like mathematics?" by choosing one of the four levels: "I like a lot," "I like a little," "I dislike a little," and "I dislike a lot." The percentage of Japanese fourth - graders who chose I like a lot was 31.1%, which was far below the international average of 58.7%. The percentage of math-dislike fourth - graders was 34.1%, which was much higher than that of the international average of 18.5%. The results were the same for Japanese eighth - graders of whom only 12.7% chose I like a lot whereas the international average was 32.2%. The percentage of math-dislike eighth - graders was as large as 60.8%, which was almost double the international average of 33.7%.

Meanwhile, the Curriculum Research Center, a subdivision of the NIER of Japan, has conducted a nationwide survey, the National Assessment of Academic Ability, every year since 2007. More than a million children from elementary and junior high schools all over Japan participate in this survey each year. It also includes a student questionnaire asking about their like/dislike of mathematics. The results of the 2015 survey showed that 38.8 and 29.1% of the elementary and junior high school children, respectively, answered that they liked mathematics (National Institute for Educational Policy Research, 2015). The percentages of the math-like children have been slightly increasing from 35.6 and 25.3% since the 2007 survey. Nevertheless, the percentages of elementary and junior high school children who disliked mathematics were still as high as 33.2 and 43.6%, respectively.

The Answers to Questionnaires Might Be Distorted

It has been known that the answers to a questionnaire may be biased for various reasons. The junior high school student participants were mature enough to be aware of

the purposes of the survey and might have distorted their answers consciously. The students seemed to have been concerned about the social desirability (Fisher, 1993) more when they answered the national survey than the international one. Although the ages of samples were slightly different between the two surveys, the national survey showed somewhat better results than the international one did, especially those of the junior high school students. The percentage of the math-like students in the national survey (25.3%) was two times more than that of TIMSS2011 (12.7%).

To remove the social desirability biases, Greenwald and Banaji (1995) argued the necessity of the procedures that could assess the implicit aspects of social attitudes. Various procedures have been developed to assess the implicit attitudes that are free from responders' intentional biases (Bar-Anan & Nosek, 2014). Among them, the Implicit Association Test (IAT: Greenwald, McGhee, & Schwartz, 1998) has been most widely used in the psychological research. With the IAT, the researchers measure the differences in the response times of the target concepts in the two cognitive evaluation tasks. Under the condition in which the target concepts were arranged congruently with positive and negative evaluation words, the response times would be shorter than the opposite condition in which they were arranged incongruently. Accordingly, by measuring the response time differences, the researchers can assess whether the participant associated the target concepts positively or negatively in the mind.

It would be suitable to use the IAT for the assessment of the like and dislike of school subjects without the social desirability biases. However, the IAT is a computer-based procedure requiring a personal computer for each student. This disadvantage makes it difficult to use the IAT in the school settings. Mori, Uchida, and Imada (2008) converted the IAT into a paper-and-pencil version (the "FUMIE test") so that it could be much easier to administer without using individual computers. It might sacrifice the precisions of the original IAT, but Mori et al. (2008) examined the validity/reliability of the FUMIE test and confirmed its validity/reliability. The FUMIE test comprised of a series of classification tasks of words with either positive or negative valence. The words were printed in lines on a sheet of paper. Participants performed the task by marking either "O" (denoting "good") or "X" ("bad) on each word in 20 s each line. The target words ("math" or "science" in the present study) were scattered randomly among the evaluation words. The experimenter instructed participants to mark the target either with O or X in turns for 20 s. The rationale of the FUMIE test is that if a participant has a positive implicit attitude toward the target word, the positive tasks (= marking O) are easier and hence are completed more in the limited time than are the negative tasks (= marking X). Accordingly, the difference in the performances on the positive and negative tasks serves as an index of the implicit attitude toward the target.

The FUMIE test has a further advantage. It can be administered along with the conventional questionnaires for measuring their explicit attitudes so as to assess both the implicit and explicit aspects. Mori and Mori (2007) reported that Japanese junior high school girls answered similarly as boys in questionnaires on like and dislike of mathematics but performed differently in the FUMIE test toward mathematics. In the present study, we aimed to investigate further the discrepancies in the explicit/implicit aspects of math- like/dislike among Japanese junior high school boys and girls.

We hypothesized that there would be considerable proportions of students who explicitly disliked math but implicitly accepted math ("fake math-dislikes") as we found in the previous study (Mori & Mori, 2007). We hypothesized further that the

fake math-dislike students would turn into "real math-dislikes" eventually without an appropriate intervention. The literature on self-concept and school attainment has shown that students' self-beliefs of their own ability, or attitude, are fundamental to their future achievement (e.g., Huang, 2011). It might be possible to prevent them from becoming real math-dislikes if we inform them of their positive implicit attitudes to encourage them not to give up studying mathematics.

We report here two successive studies for investigating the following two research questions. Study 1: How would Japanese junior high school students show the discrepancies of the explicit/implicit aspects of math like/dislike? Study 2: Whether the intervention of informing of the positive implicit attitudes would be effective to prevent fake math-dislikes from becoming real dislikes.

Study 1: Detection of Fake Math-Dislikes in Junior High School Students

We administered a battery of questionnaires and a paper-based IAT (FUMIE test) to assess the attitudes toward mathematics of 308 participants. For comparison, we also administered an equivalent test battery for the school subject of science to another set of 102 participants.

Method

Participants Four hundred ten Japanese junior high school students participated in Study 1. One hundred one seventh - graders (48 boys and 53 girls), 107 eighth - graders (50 boys and 57 girls), and 100 ninth - graders (51 boys and 49 girls) were assigned to the assessment of attitudes toward mathematics. For the purpose of baseline comparison, another set of 102 (17 boys and 16 girls in seventh- grade, 18 boys and 18 girls in eighth - grade, and 16 boys and 17 girls in ninth - grade) took part in the science assessment. Their school was a municipal school in a suburb of Nagano City, located about 200 km northwest of Tokyo. The socioeconomic status of the students' families was within a narrow middle-class range. All students were Japanese natives. The implicit tests could reveal hidden aspects of participants. Therefore, we obtained informed consent and ethics approval from the principal and teachers of the school. We also obtained the informed consent of the participants using the procedure approved by the school.

Explicit Assessment: Questionnaires Questionnaires for self-ratings on like and dislike of two school subjects, mathematics and science, were assembled, including the question "Do you like/dislike mathematics/science?" along with filler questions. Participants rated the questions twice (for "like" and "dislike" once each) on a five-point-scale (2 = yes, 1 = moderately yes, 0 = neutral, −1 = moderately no, −2 = no).

Implicit Assessment: Paper-Based Implicit Association Test (FUMIE Test) We prepared two versions of the FUMIE test (Mori et al., 2008) with math and science as target words. We administered the math versions to 308 participants and the science versions to a different set of 102 participants.

Procedure We prepared the administration manual for the questionnaires and the FUMIE test and asked the teachers to administer the assessment battery to their

mathematics class, using about 10 min. The procedural orders were counterbalanced among the participants.

Results

Explicit Assessment: Liking Scores We calculated the Liking scores by adding up the five-point-scale ratings (2, 1, 0, −1, −2) on the like questions and the reversed ones of the dislike questions. The average Liking scores of mathematics were higher for boys (.58) than for girls (−.06). A two-way ANOVA (2 sex × 3 grade) revealed that the main effect of sex was significant ($F_{(1, 302)} = 25.54, p = .0000, \eta^2 = .08$) while the grade effect was not ($F_{(2, 302)} = 2.32, p = .10$). The interaction was not significant either ($F_{(2, 302)} = .72, p = .49$).

Meanwhile, for science, the average Liking scores showed a different pattern. The scores were higher for boys (1.80) than for girls (.84), and boys maintained their high scores while the girls' scores decreased along with their grades. A two-way ANOVA (2 sex × 3 grade) revealed a significant interaction ($F_{(2, 96)} = 5.79, p = .0042, \eta^2 = .08$). The main effects were also significant ($F_{(1, 96)} = 14.54, p = .0002, \eta^2 = .11; F_{(2, 96)} = 12.63, p = .0007, \eta^2 = .11$). As the interaction was significant, we analyzed the differences among the grades separately for boys and girls. There were no significant differences among the grades in boys, but the girls' data showed significant differences among the grades ($F_{(2, 96)} = 13.14, p = .00001$).

Implicit Assessment: Implicit Association Quotients We counted separately the total number of words marked in 60 s (3 × 20 s) for the positive and negative tasks (W_P and W_N, respectively) and converted them into the Implicit Association Quotients (IAQ_{100}) using the following formula: $IAQ_{100} = 100 × (W_P − W_N) / (W_P + W_N)$. The IAQ_{100} is an index representing the difference in the numbers of words marked under the two conditions per 100 words. A positive/negative IAQ_{100} means a positive/negative implicit attitude toward the target.

The average Math-IAQ_{100} was positive, 3.81 ($SD = 7.00$). It was positive throughout the three grades: 2.93, 4.91, and 4.27, for seventh-, eighth-, and ninth - graders, respectively. Boys' average scores (5.00) were higher than girls' (2.69). A two-way ANOVA (2 sex × 3 grade) revealed that the main effect of sex was nearly significant ($F_{(1, 279)} = 3.57$, $p = .0599, \eta^2 = .01$) and the grade effect was significant ($F_{(2, 279)} = 3.49, p = .0318$, $\eta^2 = .03$). The interaction was not significant ($F_{(2, 302)} = .90, p = .41$). The multiple comparison analyses by the Ryan procedure revealed that the average Math-$IAQs_{100}$ rose significantly from seventh- to eighth - grade ($t_{(279)} = 2.62, p = .0093$) and stayed at a similar level (the difference between seventh- and ninth - graders, $t_{(279)} = 1.71, p = .0889$; the difference between eighth- and ninth - graders, $t_{(279)} = .86, p = .39$).

As for the Science-IAQ_{100}, they were also positive throughout the three grades: 6.27, 6.28, and 6.74, for seventh-, eighth-, and ninth- graders, respectively. Boys' average scores (7.76) were higher than girls' (4.48). A two-way ANOVA (2 sex × 3 grade) revealed that the main effect of sex was significant ($F_{(1, 82)} = 8.72, p = .0041, \eta^2 = .09$), but the grade effect was not significant ($F_{(2, 82)} = .09, p = .92$). The interaction was not significant ($F_{(2, 82)} = 1.96, p = .15$) either.

Fake Dislikes: Cross Tabulation of Liking and IAQ_{100} We classified students into the following three groups based on the liking scores: 1 or higher = Like, 0 = Neutral,

and −1 or lower = Dislike. As for mathematics, there were 87 boys and 57 girls in the Like group, 27 and 37 in Neutral, and 33 and 65 in Dislike. There were different distributions of students for the liking of science. There were 45 boys and 30 girls in the Like group, 6 and 8 in Neutral, and 0 and 13 in Dislike.

Then, we classified them according to their IAQ_{100} for mathematics and science. As for Math-IAQ_{100}, we excluded 23 outliers (12 boys and 11 girls) whose scores were either larger or smaller than the 95% reliability range ($−9.92 \leqq$ Math-$IAQ_{100} \leqq 17.54$). We also discarded 24 ambiguous data (11 boys and 13 girls) in the near-zero range ($−1.00 <$ Math-$IAQ_{100} < 1.00$). Finally, we obtained 106 boys and 104 girls in the Positive group and 20 and 31 in the Negative group. We conducted the same procedure for the Science-IAQ_{100} and got 42 boys and 31 girls in the Positive group and 4 and 11 in the Negative group.

These two classification results were cross-tabulated as shown in Table 1. Most of the students fell in the left-bottom and right-top cells with consistent results on the two assessments, having explicitly rated dislike and showed negative valences in IAQ or having rated like and positive IAQ scores. Meanwhile, there were some students who fell in the opposite pattern of cells with conflicting results, having rated dislike despite their positive IAQ scores or having rated like with negative IAQs. Those conflicting cases were scarce in general. However, there was a remarkable proportion of students in the left-top cell for mathematics (shown in bold font in Table 1). As compared with the corresponding ratio of science, 4.9%, it was as large as 20.1% of the students or one out of every five. It was strongly suspected that most of those students had lied, consciously or in a somewhat unconscious way, to the like/dislike questions. Therefore, we call them fake math-dislikes in the present study.

Grade Differences of the Proportions of Fake Math-Dislikes The proportion of fake math-dislikes in the seventh - graders was 9.9%, while those in the eighth- and ninth - graders were 28.0 and 22.0%, respectively. The differences of the

Table 1 Cross-tabulation of students based on the liking and IAQ_{100} scores

Mathematics		Questionnaires		
IAQ_{100}	Dislike	Neutral	Like	Total
Positive boys	**24**	16	66	106
Positive girls	**38**	24	42	104
Negative boys	5	4	11	20
Negative girls	14	7	10	31
Total	81	51	129	261
Science	Dislike	Neutral	Like	Total
Positive boys	0	4	38	42
Positive girls	5	7	19	31
Negative boys	0	1	3	4
Negative girls	5	0	6	11
Total	10	12	66	88

The fake math-dislike students are shown in bold font

proportions between the seventh - graders and the upper graders (eighth and ninth) were statistically significant ($z = 3.32$, $p = .0009$, $z = 2.34$, $p = .0191$). These results implied the fake math-dislikes would develop especially from seventh- to eighth - grade. There was a small sex difference in the proportions of fake math-dislikes (16.1% in boys and 23.9% in girls), but it did not reach the statistical significance ($z = 1.70$, $p = .0883$).

Study 2: Treatment of the Fake Math-Dislikes and Its Effect

We found as large as 20% of students were suspected to be fake math-dislikes in Study 1. Although they were in the stage of fake math-dislikes at that time, it should be concerned that they would become real math-dislikes eventually. It would be desirable if we could stop this unfortunate process with some educational interventions. One of the plausible treatments to prevent those fake math-dislike students from becoming real dislikes would be an intervention by informing of their positive implicit attitudes toward mathematics. Then, we aimed to examine the effects of this intervention on their future performance in mathematics. For this purpose, we administered the same assessment procedure to collect the experimental sample of fake math-dislikes in seventh - grade students. We chose seventh - graders for Study 2 because we had found the leap of the proportion of the fake math-dislikes between the seventh- to eighth - graders in Study 1. We hypothesized that the intervention of informing of the positive implicit attitudes would prevent seventh-grade fake math-dislikes from becoming real dislikes in the eighth - grade. We would test this hypothesis by assessing their achievement scores in mathematics at their eighth - grade year.

Method

Preliminary Study: Detection of Fake Math-Dislike Students We assessed 204 seventh - graders of a junior high school in Nagano City with the same question-naires for like and dislike of mathematics and the paper-based IAT with math as the target word. We found 38 (25 boys and 13 girls) fake math-dislike students. We also chose 24 (eight boys and 16 girls) real math-dislike students for the experiment. We obtained their achievement scores in mathematics assessed at the most recent term examination and 1 year later. We used the informed consent procedure to obtain the ethics approval from the school in the same way as Study 1.

Experimental Design We used a two-way between-subject design (fake vs. real dislike × with vs. without intervention). The dependent variable was the achievement score 1 year after the experimental intervention.

Randomized assignment to experimental groups. First, the 38 fake math-dislike students were matched into 19 pairs based on their sex and their achievement scores. Then, each of the pairs was randomly assigned to the two conditions. Because there were odd numbers of fake math-dislike boys and girls, one combination in each sex had three students. Finally, we had 20 students (13 boys and seven girls) in the experimental condition and 18 (12 boys and six girls) in the control condition. We randomly assigned 24 real math-dislike students to the two conditions (four boys and eight girls each to experimental and control conditions).

Academic achievement scores. The academic achievement of the students was regularly assessed in the term examinations. The test scores were converted into standardized scores (Z-scores) using the following equation:

$$Z = 50 + 10(X-m)/ \; SD; \; where \; X = raw \; score, m = mean, SD = standard \; deviation.$$

In Japan, most junior high schools use the Z-scores in the official school assessments of students (Saitoh & Newfields, 2010). We obtained the Z-scores in mathematics of the first term examination conducted in July of the experiment year and those of the next year. Therefore, all the achievement scores were the legitimate ones.

Intervention by informing of positive implicit attitudes toward math. Two months after the questionnaire and paper-based IAT administration, the experimenters sent a letter of thanks to the students individually for their participation in Study 1. For the students in the experimental condition, the following information was provided:

The questionnaire analysis showed your low Liking in mathematics. However, the newly invented psychological test revealed your positive attitude toward school life in your mind. *It also found out your forward-looking attitude toward mathematics in your mind.* Please study hard and enjoy your school life.

Control students received the same information except for the italicized sentence.

Results

Changes in the Achievement Scores after 1 Year Five fake math-dislike students (four in experimental and one in control condition) who were absent from the year-later term examination were removed from the following data analyses. In the fake math-dislike groups, the average Z-scores in mathematics of experimental students improved from 47.6 to 49.7 1 year after the intervention. Meanwhile, those of control students declined from 46.2 to 45.9 1 year later. These results showed the effectiveness of the intervention. In the real math-dislike groups, both experimental and control students did not improve their scores, from 45.83 to 44.56 and from 46.28 to 43.21, respectively. These data appeared to support the hypothesis that informing their implicit attitude would prevent the fake math-dislike students from becoming real math-dislikes.

However, the achievement scores of the students varied widely from 27.2 to 62.2. The Z-scores were inappropriate for the parametric statistical analyses because of large individual differences. Therefore, we counted the number of students who improved their scores among each group. As for the fake math-dislike groups, 15 of the 16 experimental students showed improvement, while only eight of the 17 control students improved their scores 1 year later (Fisher's exact test showed a statistical significance with $p = .0066$, *Phi* $= .508$). As for the real math-dislike groups, six and four of the 12 students showed improvement. A $2 \times 4 \; X^2$ test revealed that the differences in the ratios of the number of improved students in four groups were statistically significant ($X^2_{(3)} = 12.534$, $p < .01$, $\varphi_c = .468$). The probability of obtaining the ratio of 15/16

or larger by chance is only 0.00026. Therefore, the intervention worked but only for the fake math-dislike students in the experimental condition.

General Discussion

Why Do They Pretend to Be Math-Dislikes?

It is natural to assume that those who are not good at mathematics dislike it. It is also true that those who are good at math like it. However, it is uncertain whether the like or dislike of mathematics is a result of the poor achievement or a cause of it. The truth must be that they are interrelated in a causal chain. A student poor at math comes to dislike it and, then, will not study much, which will result in poor performance in the examination. It is the vicious cycle of the math-dislike student.

However, as the vicious cycle described above shows, disliking math is not a direct cause of poor achievement in math. Rather, it is a cause of or an excuse for not studying. Students with low achievement scores in mathematics will be told by their teachers and parents to study more diligently. By studying more, they should get better results in the next examination even though they dislike the subject. However, if they do not like to study, they need an excuse for not studying. Disliking math is a suitable excuse for not studying math. The excuse of disliking math is also helpful for maintaining self-esteem. Even if their math achievement scores are low and endanger their self-esteem, they can attribute their poor performance not to their competence but to their lack of studying because of disliking mathematics.

Although we dubbed "fake" math - dislike for those students who dislike mathematics but with positive implicit attitudes toward mathematics, they might not have a faking intention. Some of them might have some unclear, mixed feeling about mathematics. They experienced difficulties studying mathematics and might be uncertain of their liking or disliking, by themselves. On the other hand, some may be aware of their situation and consciously fake their own minds, a sort of self-deception or trying to deceive others such as teachers, parents, or classmates. In either case, they are in an incubation period prior to proceeding to deeper stages of math - dislike. Therefore, it is crucially important to prevent them at this stage from becoming real math-dislikes.

Did the Intervention Really Work?

It was a remarkable finding that informing students that "*It also found out your forward-looking attitude toward mathematics in your mind*" worked to prevent the fake math-dislike students from becoming real math-dislikes. There has been a large literature reporting that a mere belief might have a considerable effect in educational practices (e.g., "Pygmalion effect" by Rosenthal & Jacobson, 1968). The present study may be reporting another "placebo" effect in education. However, it was not a mere belief. The information the students believed was based on their own implicit attitude. Whether they were aware of it or not, it rooted on firm ground in their mind. It was noteworthy that the same information did not work for the real math-dislike students who not only explicitly disliked mathematics but also had a negative implicit attitude

toward mathematics. They could not believe it because the information seemed not true for them. One of the authors of this study is a schoolteacher himself. We would like to trust the power of belief for improving students in school. Even a small belief, if it is a true one, does make a difference in education.

Future Research

The characteristic of Japanese students learning mathematics reported in this article was not limited to Japanese students. There were similar tendencies found among the East Asian countries in the TIMSS2011. There may be found fake math-dislike students in those countries as well. It would be worth conducting similar research in East Asian countries such as Korea, Hong Kong, and Taiwan. Meanwhile, there was the opposite tendency among students in the Western nations as Sheldrake, Mujtaba, and Reiss (2015) wrote: "Students often, but not universally, over-evaluate their abilities. This has been seen at various ages, including at primary school and at university in Canada, Europe and North America. (p.464)" There might be fake math - likes among Western children. It would also be desirable to conduct research in Western countries as well.

Advantages and Limitations of Paper-Based IAT in Education

Several researchers have proposed a variety of paper-and-pencil versions of IAT (e.g., Lemm, Lane, Sattler, Khan & Nosek, 2008; Lowery, Hardin, & Sinclair, 2001; Mori et al., 2008; Teachman, Gapinski, Brownell, Rawlins & Jeyaram, 2003). The FUMIE test is one of them and has been utilized in educational research mostly in Japan (Kurita & Kusumi, 2009; Mori & Mori, 2007; Sakai & Koike, 2011). As shown in the present study, it is particularly useful in educational settings. Schoolteachers can administer it in their class along with the conventional questionnaires. The test is in an Excel format so the target words can be freely replaced with the most suitable ones for the research objectives. The data analysis procedure is simple and can be done by schoolteachers as well. The combinational usage with questionnaires can detect the discrepancies between the two measures as shown in the present study. The evaluation words were chosen based on the familiarity assessment with junior high school students. Therefore, it can be administered to a wide age range of participants, from junior high school students to elderly citizens.

   There are several limitations as well. Most of the limitations are in common with those of the original IAT, such as low reliability originated from the reaction time measurement. People may experience an occasional lapse time without any particular reasons. The reaction time measurement is vulnerable to it. The IAT measures each of the response times of the participant and uses only those made within a certain time limit so as to exclude an accidentally lapsed response. However, the FUMIE test does not have such devices so the assessed data may include those occasional time lapses. Mori et al. (2008) reported that the reliability of the FUMIE test was slightly lower than that of the IAT, as the reliabilities of the FUMIE test for the three targets, "marriage," "pregnancy," and "romantic love," were 0.56, 0.61, and 0.71, respectively. For this reason, we excluded the outliers whose scores dropped outside of the 95% reliability range.

   With this limitation, however, the convenience of the FUMIE test would be advantageous as a research tool for collecting massive implicit data without using

sophisticated equipment. As Vargas, Sekaquaptewa, and von Hippel (2007) claimed, "low-tech" measures of implicit attitudes have the potential to add predictive power beyond explicit measures. They minimize social desirability concerns, without requiring special equipment. Above all, they are easy to handle and applicable to large groups of respondents at the same time, even in schools.

Author Contributions   Both authors contributed to the study design, the data collection, the data analysis, and the interpretation. A. Uchida drafted the first manuscript in Japanese, and K. Mori wrote the English manuscript. Both authors approved the final version of the manuscript for submission.

References

Bar-Anan, Y., & Nosek, B. A. (2014). A comparative investigation of seven indirect attitude measures. *Behavior Research Methods, 46*, 668–688. doi:10.3758/s13428-013-0410-6.

Fisher, R. J. (1993). Social desirability bias and the validity of indirect questioning. *Journal of Consumer Research, 20*, 303–315. doi:10.1086/209351.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*, 4–27. doi:10.1037//0033-295X.102.1.4.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. K. L. (1998). Measuring individual difference in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74*, 1464–1480. doi:10.1037/0022-3514.74.6.1464.

Huang, C. (2011). Self-concept and academic achievement: A meta-analysis of longitudinal relations. *Journal of School Psychology, 49*(5), 505–528. doi:10.1016/j.jsp.2011.07.001.

Kurita, T., & Kusumi, T. (2009). Implicit and explicit attitudes toward people with disabilities and effects of the internal and external sources of motivation in moderating prejudice. *Psychologia: An International Journal of Psychological Sciences, 52*, 253–260. Kyoto, Japan: Psychologia Society.

Lemm, K. M., Lane, K. A., Sattler, D. N., Khan, S. R., & Nosek, B. A. (2008). Assessing implicit cognitions with a paper-format implicit association test. In M. A. Morrison & T. G. Morrison (Eds.), *The psychology of modern prejudice* (pp. 121–146). Hauppauge, NY: Nova Science Publishers.

Lowery, B. S., Hardin, C. D., & Sinclair, S. (2001). Social influence effects on automatic racial prejudice. *Journal of Personality and Social Psychology, 81*, 842–855. doi:10.1037/0022-3514.81.5.842.

Mori, K. & Mori, H. (2007). Shudan-shiki Senzai Renso Tesuto o Mochiita Chugakusei no Kyoka no Kouodo Sokutei. [Evaluation of scholastic subjects by junior high school students using a newly developed group performance implicit association test.] *Annual Letters of Clinical Psychology in Shinshu* (Faculty of Education, Shinshu University, Nagano, Japan), *6*, 1–4. (in Japanese)

Mori, K., Uchida, A., & Imada, R. (2008). A Paper-format group performance test for measuring the implicit association of target concepts. *Behavior Research Methods, 40*(2), 546–555. doi:10.3758/BRM.40.2.546

National Institute for Educational Policy Research (2012). *Kokusai Sugaku/Rika Kyoiku Doukou Chosa no 2011-nen Chosa (TIMSS2011) Kokusai Chosa Kekka Houkoku (Gaiyo)*[A brief report on the results of the trends in international mathematics and science study in 2011 (TIMSS2011)]. Tokyo, Japan: The National Institute for Educational Policy Research.

National Institute for Educational Policy Research (2015). *Heisei 26-nendo Zenkoku Gakuryoku/Gakushu Jokyo Chosa Houkokusho: Shitsumonshi Chosa* [The report on the results of the national assessment of academic ability 2014: Questionnaire survey]. Tokyo, Japan: The National Institute for Educational Policy Research.

Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom*. Austin, TX: Holt, Rinehart & Winston.

Saitoh, N., & Newfields, T. (2010). Insights in language testing: An interview with Shozo Kuwata—a pioneer of standardized rank scoring in Japan. *SHIKEN: JALT Testing & Evaluation SIG Newsletter, 14*, 2–5. Tokyo, Japan: The Japan Association For Language Teaching.

Sakai, H., & Koike, H. (2011). Implicitly and explicitly measured attitudes towards foreigners: A dual-process model perspective. *JABAET Journal, 14/15*, 39–58. Tokyo, Japan: The Japan-Britain Association of English Teachers.

Sheldrake, R., Mujtaba, T., & Reiss, M. J. (2015). Students' intentions to study non-compulsory mathematics: The importance of how good you think you are. *British Educational Research Journal, 41*(3), 462–488. doi:10.1002/berj.3150.

Teachman, B. A., Gapinski, K. D., Brownell, K. D., Rawlins, M., & Jeyaram, S. (2003). Demonstrations of implicit anti-fat bias: The impact of providing causal information and evoking empathy. *Health Psychology, 22*, 68–78. doi:10.1037/0278-6133.22.1.68.

Vargas, P. T., Sekaquaptewa, D., & von Hippel, W. (2007). Armed only with paper and pencil: "low-tech" measures of implicit attitudes. In B. Wittenbrink & N. Schwarz (Eds.), *Implicit measures of attitudes* (pp. 103–124). New York, NY: Guilford Press.