

# Evaluation Of The Translated Version Of The FMCE

Michi Ishimoto

*School of Environmental Science and Technology, Kochi University of Technology, Tosayamada-cho Kami-shi Kochi, 782-8502, Japan*

**Abstract:** The Force and Motion Conceptual Evaluation (FMCE) is a test that was originally developed in English to assess American students' understanding of Newton's laws of motion. The present study evaluates the Japanese translation of the FMCE (abbreviated as "FMCEJ") using data consisting of the pretest results of 1095 Japanese students, most of whom were first-year students at a mid-level engineering school between 2003 and 2012 in Japan. The classical test theory indices of the FMCEJ indicate that its reliability and discrimination are adequate in assessing Japanese students' views about force and motion. A comparison of these students' understanding of the concepts of motion and force, as assessed with the FMCEJ, and that of American students, as assessed with the FMCE, indicates the validity of the FMCEJ.

**Keywords:** FMCE, Evaluation of Concept Inventory, Japanese translation

**PACS:** Physics education, 01.40.-d

## INTRODUCTION

The use of standard examinations, such as the Force Concept Inventory (FCI) [1] and the Force and Motion Conceptual Evaluation (FMCE) [2], to test physics concepts has led to the development of many effective instructional methods in the United States [3]. Japanese educators, however, do not have the means to measure the efficacy of their instructional methods.

The Concept Inventory is designed to probe not only students' understanding of Newtonian concepts but also concepts commonly held by students. The context and the writing style of the multiple-choice questions may influence students' responses to the test. The FCI and the FMCE are written in a style so as to measure accurately students' understanding of concepts. A technical translation must use consistent and accurate terminology and incorporate appropriate linguistic features. However, a translated version of the Concept Inventory must also be style sensitive because it needs to elicit not only correct concepts but also commonly held concepts.

The Japanese language, which belongs to the Altaic family of languages that includes the Turkic and Mongolic languages, is a very different language from English, which belongs to the Indo-European family of languages. Thus, the translation of English sentences into Japanese adds a layer of complication to the writing style.

Japan has developed its own educational system. As an educator working within this system, I have observed many students with the same commonly held

naïve concepts about force and motion as those previously mentioned by Hestenes *et al.* [1]. I assume that a properly translated concept inventory could produce similar outcomes as the original version. In this study, I evaluate my translated version of the FMCE, a test originally developed in English for American students, in terms of the validity of its administration to Japanese students.

The format of the FMCE [2] is suitable for assessing students' views of force and motion by cluster analysis. The structure enables the translated version to be validated by comparing Japanese data to American data in clusters. In this study, the FMCE, translated by the author into Japanese (abbreviated as "FMCEJ"), was administered under the same conditions to 1095 students. The FMCEJ was evaluated statistically based on the classical test theory [4]. The clusters were then compared to American data to validate the efficacy. For the purpose of comparison, I used the data reported by Smith and Wittmann [5] to represent American data.

## METHODS

### Surveyed Students

The surveyed students were mostly first-year students attending a basic mechanics class at the Kochi University of Technology (KUT) between 2003 and 2012. Each year, more than 500,000 university candidates in Japan take the National Center Test

(NCT), an academic aptitude test. On average, KUT students surveyed for the present study earned a score of 63 points (of 100 points) on the NCT in physics, which is comparable to the national average of 62.68 points reported in 2013 [6]. Thus, it is assumed that the data must be suitable for assessing the physics pre-concepts of Japanese students at a mid-level engineering school.

## General Statistics

The FMCE is designed to probe students' views of force and motion concepts using clusters with respect to subjects and representation. First, I used simple sum scores (SSS's) with all 47 FMCE questions for the statistical analysis based on the classical test theory. Second, I used the scoring systems used by Thornton *et al.* [7] and clustering suggested by Smith and Wittmann [5] to compare Japanese and American data.

### Single Number Scores

Thornton *et al.* [7] determined a single number score (SNS) for the FMCE to characterize students' understanding of Newtonian concepts. These authors omitted 7 questions (5, 6, 15, 33, 35, 37, and 39) that they considered to be less reliable in probing students' understanding of the concepts. They suggested that 9 of the 36 questions be divided into three sets of 3 questions, with a score of 2 points being obtained if all 3 questions in a set are answered correctly. The rationale behind this scoring system is that the FMCE designers believe that students only understand the concepts involved if they answer all three questions correctly. They advocated SNS's for the questions (excluding those questions on energy conservation), producing a total of 33 points.

The SNS put forth by Thornton *et al.* [7] excludes the 4 questions on energy conservation. Because the FMCEJ includes these 4 questions, I used the 33 points designated for the SNS by Thornton *et al.* [7] and 4 points for the 4 questions on energy conservation and referred to them as the SNS in this study. The inclusion of the 4 questions on energy conservation thus changes the possible total score on the FMCEJ to 37 points. The basic statistical values of the SSS's and the SNS's are similar (Table I), and they are well correlated ( $r = 0.99$ ).

**Table I.** Basic statistics for three scoring systems

	SSS	SNS
<i>M</i>	16.05 (34%)	11.35 (31%)
<i>SE</i>	0.29	0.22
<i>Mdn</i>	13	9
<i>SD</i>	9.51	7.32
Max. Score	47	37
CI (95.0%)	0.56	0.43

CI, confidence interval; *M*, mean; Max. Score, maximum possible score; *Mdn*, median; *SD*, standard deviation; *SE*, standard error; SNS, single number score; SSS, simple sum score.

## RESULTS

### Classical Test Theory Indices

All 47 questions were used to analyze statistically the FMCEJ, based on classical test theory [4]. The reliability index (KR-20) and the discriminatory power of the test (Ferguson's  $\delta$ ) were 0.91 and 0.98, respectively, indicating that the FMCEJ is adequate in evaluating individuals' understanding of force and motion concepts (Table II).

**TABLE II.** Evaluation of the FMCEJ

Evaluation Measure	FMCEJ (47 pts)	Desired Values
p	Average of 0.34	0.3–0.9
D	Average of 0.46	$\geq 0.30$
PBC	Average of 0.49	$\geq 0.2$
KR-20	0.91	$\geq 0.7$
Ferguson's $\delta$	0.98	$\geq 0.9$
No. of students	1095	N/A
Average	16.05 $\pm$ 0.29	N/A
<i>SD</i>	9.51	N/A
CI (95.0%)	0.564	N/A

CI, confidence interval; D, discrimination index; FMCEJ, Force and Motion Conceptual Evaluation–Japanese; KR-20, reliability index; p, difficulty index; PBC, point biserial coefficient; *SD*, standard deviation.

Table III shows each item's difficulty index (p), point biserial coefficient (PBC), and discrimination index (D) based on data from the top 25% of scorers and the bottom 25% of scorers. The columns labeled "MC" and "NA" in Table III show the FMCEJ distractor symbols for the most commonsense concepts (MC's) and the proportion of no correct answers (NA's), respectively. The 10 blank spaces in the MC columns of Table III indicate that no MC's were probed. The first 7 blank spaces correspond to the seven items regarded by Thornton and Sokoloff [2] as

inaccurate for the assessment. The last 3 blank spaces correspond to velocity graph questions. Because the correct response rate was near 90% for these questions, the other responses were too minute to be listed. Almost all MC's of Japanese students agree with those of American students. The items with different MC's from those of the U.S. counterparts are items 23, 46, and 47 (presented in bold type in Table III). Item 23 regards the corresponding acceleration for a toy car with decreasing speed. The surveyed Japanese students most commonly used the  $F \propto v$  model (52%), whereas Smith and Wittmann [5] found that American students most commonly used the brake model. In the brake model, the opposite force to the velocity direction causes the change in velocity. Items 46 and 47 refer to the velocity of a sled at the bottom of a hill when released from rest at the top of two kinds of hills. One

hill is higher and less steep than the other hill. The most common incorrect response (around 20%) given by the Japanese students was that there was insufficient information to answer these questions. By contrast, the most common incorrect response given by American students was slope dependent; that is, they incorrectly attributed the higher speed of the sled to the steepness of the hill.

**TABLE III.** Difficulty index (p), point biserial coefficient (PBC), discrimination index (D), and most commonsense concept (MC)

Item #	p	PBC	D <sup>a</sup> (25%)	MC <sup>b</sup>	NA <sup>c</sup>	Item #	P	PBC	D (25%)	MC	NA
1	0.24	0.61	0.58	a	0.01	25	0.31	0.68	0.74	f	0.03
2	0.16	0.56	0.45	b	0.04	26	0.38	0.66	0.83	a	0.02
3	0.35	0.57	0.66	b	0.02	27	0.22	0.68	0.66	g	0.00
4	0.26	0.54	0.52	g	0.01	28	0.16	0.63	0.49	d	0.01
5	0.48	0.47	0.62		0.04	29	0.29	0.57	0.62	b	0.00
6	0.14	0.54	0.39		0.10	30	0.17	0.29	0.24	a	0.01
7	0.41	0.53	0.65	e, a	0.02	31	0.18	0.27	0.25	f	0.03
8	0.11	0.53	0.32	g	0.01	32	0.35	0.41	0.52	b, f	0.13
9	0.09	0.54	0.27	d	0.02	33	0.88	0.21	0.19		0.04
10	0.19	0.48	0.41	b	0.01	34	0.34	0.44	0.54	b	0.11
11	0.16	0.67	0.51	g	0.00	35	0.32	0.37	0.39		0.02
12	0.14	0.62	0.41	d	0.01	36	0.06	0.35	0.15	c	0.05
13	0.25	0.57	0.54	b	0.00	37	0.44	0.33	0.37		0.17
14	0.13	0.67	0.42	a	0.05	38	0.07	0.30	0.14	b	0.20
15	0.96	0.10	0.05		0.02	39	0.47	0.45	0.59		0.01
16	0.18	0.71	0.53	c	0.01	40	0.93	0.10	0.08		0.01
17	0.11	0.67	0.39	b	0.07	41	0.80	0.23	0.27	g	0.06
18	0.12	0.68	0.40	h	0.06	42	0.87	0.18	0.17		0.02
19	0.15	0.69	0.48	d	0.06	43	0.89	0.16	0.15		0.07
20	0.23	0.53	0.48	f	0.14	44	0.37	0.53	0.66	a	0.01
21	0.19	0.67	0.57	f, h, a	0.12	45	0.55	0.45	0.61	a	0.01
22	0.39	0.68	0.83	e, a	0.01	<b>46</b>	0.42	0.46	0.61	<b>d, c</b>	0.02
<b>23</b>	0.26	0.70	0.72	<b>g</b>	0.03	<b>47</b>	0.52	0.45	0.61	<b>d, c</b>	0.01
24	0.34	0.67	0.79	b	0.07						

<sup>a</sup> Discrimination index (D) is based on data from the top 25% of scorers and the bottom 25% of scorers. <sup>b</sup> The 10 blank spaces in the MC columns of Table III indicate that no MC's were probed. The first 7 blank spaces correspond to the 7 items (5, 6, 15, 33, 35, 37, and 39) regarded by Thornton and Sokoloff [2] as inaccurate for the assessment. The last 3 blank spaces correspond to velocity graph questions (items 40, 42, and 43); the correct response rate was nearly 90% for these items, and, therefore, these responses were too minute to be listed. <sup>c</sup> NA, no correct answer.

## CONCLUSIONS

### Cluster Analysis

Smith and Wittmann [5] used the following clusters to analyze FMCE data: Force Sled, Reverse Direction, Force Graph, Acceleration Graph, Newton's Third Law, Velocity Graph, and Energy. I used the same clusters to analyze the present study's data. The responses in the clusters of Force Sled, Reverse Direction, Force Graph, and Acceleration Graph reveal students' views of Newton's First Law and Second Law. The apparent proportions of the models used are roughly in the range of 10% to 20% for the  $F \propto a$  model, 40% for the  $F \propto v$  model, and 40% for the brake model. The percentage of students who used a particular model depended on the context of the question asked. However, the percentage of students who consistently used the  $F \propto a$  model is estimated to be less than 5%.

The responses in the cluster of Newton's Third Law reveal that the most predominantly used commonsense concept models are the mass-dependent model and the action-dependent model. The percentage of students who consistently used the Newton's Third Law model is estimated to be less than 5%.

The responses in the cluster of Energy Conservation provide information about students' views of the law of conservation of mechanical energy. The most frequently used commonsense concept model is the steepness dependency model (three quarters of the Japanese students used this model). The percentage of students who used the energy conservation model is estimated to be less than 25%.

The FMCEJ data indicate that less than 5% of Japanese students at a mid-level engineering school have a firm grasp of the concepts of Newtonian forces and that less than 25% understand energy conservation concepts.

This study shows that the FMCEJ, a Japanese translation of the FMCE, is a high-quality conceptual multiple-choice test with levels of reliability and discrimination that are adequate for assessing Japanese students' understanding of the concepts of force and motion, based on the classical test theory. The item analysis shows the majority of items with a discrimination index of more than 0.30 and a point biserial correlation of more than 0.20. A comparison of Japanese data and American data shows that Japanese students and American students provide the same most common answers. According to the cluster analysis, no intrinsic differences exist between Japanese and American students.

## REFERENCES

1. D. Hestenes, M. Wells, and G. Swackhamer, *Phys. Teach.* **30**, 141-158 (1992).
2. R. K. Thornton and D. R. Sokoloff, *Am. J. Phys.* **66**, 338-352 (1998).
3. E. F. Redish, *Teaching Physics with the Physics Suite*, WWW Document, <http://www2.physics.umd.edu/~redish/Book/>.
4. P. V. Englehardt, "An Introduction to Classical Test Theory as Applied to Conceptual Multiple-choice Tests," in *Reviews in PER Volume 2: Getting Started in PER*, edited by C. Henderson and K. A. Harper (American Association of Physics Teachers, College Park, MD, 2009), WWW Document, (<http://www.per-central.org/items/detail.cfm?ID=8807>).
5. T. I. Smith and M. C. Wittmann, *Phys. Rev. ST Phys. Educ. Res.* **4**, 020101 (2008).
6. National Center for University Entrance Examinations, WWW Document, ([http://www.dnc.ac.jp/modules/center\\_exam/content056\\_5.html](http://www.dnc.ac.jp/modules/center_exam/content056_5.html)).
7. R. K. Thornton, D. Kuhl, K. Cummings, and J. Marx. *Phys. Rev. ST Phys. Educ. Res.* **5**, 010105 (2009).