FCI-based Multiple Choice Test for Investigating Students' Representational Coherence

Antti Savinainen, Pasi Nieminen, Jouni Viiri, Jukka Korkea-aho and Aku Talikka

Department of Teacher Education, P.O. Box 35, FI-40014 University of Jyväskylä, Finland

Abstract. We present the Representation Test derived from the FCI for evaluating students' representational coherence on some aspects of gravitation and Newton's third law. The test consists of 23 questions addressing verbal, graphical, bar chart, and vectorial representations. Matched high school student data (n = 54) on the pre- and post-test are analyzed in terms of representational coherence and scientific correctness.

Keywords: multiple representations, representational coherence, the Force Concept Inventory
PACS: 01.40.Fk

INTRODUCTION

Many studies have been published on multiple representations (such as texts, pictures, diagrams, graphs, or mathematical ones) in physics education (e.g., [1], [2], [3], [4]). Van Heuvelen and Zou [1] offer several reasons why multiple representations are useful in physics education: they foster students' understanding of physics problems, build a bridge between verbal and mathematical representations and help students develop images that give meaning to mathematical symbols. These researchers also argue that one important goal of physics education is helping students to learn to construct multiple representations of physical processes, and to learn to move in any direction between these representations. Furthermore, it has been pointed out that in order to thoroughly understand a physics concept, the ability to recognize and manipulate that concept in a variety of representations is essential [2].

Several research-based multiple choice tests have been developed for evaluating students' conceptual understanding in the domain of introductory mechanics, the most widely used being perhaps the Force Concept Inventory (FCI) [5]. The FCI addresses several representations in a variety of contexts but it does not provide a systematic evaluation of students' ability to use multiple representations when a context is fixed. It is important to note that both the context and the representation affect on students' responses: the student might be able to apply a concept in a familiar context using a certain representation but fail when the context or the representation is changed [6].

The existing tests are limited in that they do not permit comprehensive evaluation of students' skills in using multiple representations. This is why we have developed a multiple-choice test – the Representation Test - to evaluate students' representational coherence: i.e., their ability to use different representations consistently when the context is kept as constant as possible.

In this paper we present the rationale and structure of the Representation Test. High school students' pre- and post-test data on the Representation Test are analyzed from the point of view of representational coherence. We also analyze students' written responses explaining their choices in the post-test to support the validity of the Representation Test.

THE REPRESENTATION TEST

The Representation Test is based on the Force Concept Inventory [5]. Firstly, FCI questions which might be easy to transform into different representations were identified. Secondly, after inspection, FCI questions 1, 4, 13, 28 and 30 were chosen: questions 1, 13 and 30 deal with gravity, and questions 4 and 28 deal with Newton's third law. We could not use all the FCI questions since the questionnaire would then have been much too long. Therefore, the Representational Test does not deal with all the aspects of the force concept addressed by the FCI.

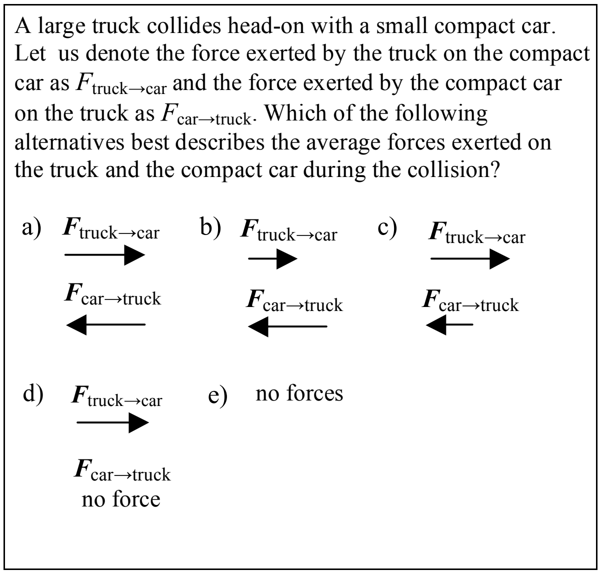
The questions chosen for the Representational Test questions are in verbal form in the original FCI. The description of the context, the question and the different alternative multiple choice items are verbal. Each question has five alternative answers of which the student must choose one. In developing the new questions we tried to keep the contexts and alternative forms of the answers as similar as possible to the original questions. For each of the selected five FCI questions, two to four new questions were formulated in different representations. We made some changes after the pilot phase [7] on the basis of the data gathered and feedback from an expert [8]: there are 23 questions altogether in the new version (one question has three parts).

We use the term *theme* for the question set developed from an FCI question and formulated in different representations (see Table 1). The representations are verbal, graphical, bar chart, and vectorial. In every question the description of the question situation is verbal, but the different multiple choice alternatives are described in different representations. For instance, the theme FCI4 corresponds to the question set developed from FCI question number 4; the alternatives in the theme FCI4 are in verbal, graphical, vectorial and bar chart representations (see Table 1). The questions of each theme were split across the Representation Test.

TABLE 1. Themes of the Representation Test, the concepts the items deal with, and the context of the theme.

Theme	Concept	Context	Representations
FCI1t	gravitation (time of falling)	falling ball	verbal, pictorial
FCI1f	gravitation (gravitation force)	falling ball	verbal, graphical, vectorial
FCI1a	gravitation (acceleration)	falling ball	verbal, bar chart
FCI4	Newton III	collision of cars	verbal, graphical, vectorial, bar chart
FCI13	gravitation (forces)	a steel ball is thrown vertically upwards	verbal, graphical, vectorial
FCI28	Newton III	students sitting on office chairs push each other off	verbal, graphical, vectorial, bar chart
FCI30	gravitation (forces)	a tennis ball passes through the air after being struck	verbal, bar chart, vectorial

A sample question in the vectorial representation is provided in Figure 1: it is derived from the FCI question 4 addressing a collision of a car and a truck (the original alternatives are in verbal representation). Additional information had to be included in the question to enable sensible formulations in multiple representations. Naturally, these additions did not change the physics involved.



A large truck collides head-on with a small compact car. Let us denote the force exerted by the truck on the compact car as $F_{\text{truck}\to\text{car}}$ and the force exerted by the compact car on the truck as $F_{\text{car}\to\text{truck}}$. Which of the following alternatives best describes the average forces exerted on the truck and the compact car during the collision?

FIGURE 1. A sample question from the Representation Test.

DATA COLLECTION AND ANALYSIS

The data for this study were collected from Finnish high school students (aged 17). The students were taught mechanics in two groups (n = 54 altogether), following the same textbook with the same teacher (author AS). Hence, the groups are combined in the following analysis. The students had had an introduction to both kinematics and the force concept before the course in mechanics. The teaching of the mechanics course was geared to foster conceptual understanding and made use of multiple representations.

The Representation Test was administered before and after teaching kinematics and the force concept to the students. The students also had to provide a written explanation in the post-test of why they selected one of the multiple choice alternatives in every item. This was done to investigate the validity of the Representation Test, i.e. to detect possible false positives (choosing the correct answer for the wrong reason) and false negatives (choosing a wrong answer while understanding the idea in question correctly).

The coherence of a student's understanding in each of the themes was determined by studying how consistently he/she had answered the different questions belonging to each theme. To answer consistently the student had to choose the corresponding multiple choice alternatives in all questions belonging to the same theme. Students' representational coherence was evaluated from two points of view. Firstly, students' answers were investigated solely from the point of view of the representational coherence regardless of the scientific correctness: their answers were considered to exhibit representational coherence if they chose the corresponding alternatives in different representations within the themes, i.e. answered all the questions in a given theme in the same way. Secondly, students' answers were investigated from the point of view of how consistent they were between representations in a given theme and also whether their answers were scientifically correct.

Students' answers in a given theme were graded in the following way:

- two points, if they had answered every question in the theme in the same way.
- one point, if they had answered every question except one in the theme in the same way.
- zero points, if they had answered two or more questions in the theme differently.

We used the same criteria in analysing students' responses when both representational coherence and scientific correctness were required: in this analysis 'answering the same way' means only scientifically correct responses.

Student's points in seven themes (see Table 1) were added up so they got from zero to fourteen points in the whole test. On the basis of this grading system students' representational coherence was categorized into three coherence classes:

- Class I: 13–14 points indicate that representational thinking was *coherent.*
- Class II: 11–12 points indicate that representational thinking was *moderately coherent.*
- Class III: 0–10 points indicate that representational thinking was *incoherent*.

The first analysis involves representationally coherent classes with or without scientific correctness: these classes are *representationally correct*. The second analysis also demands scientific correctness: these classes are *scientifically correct.*

RESULTS

To provide a general view of the data the students' matched pre- and post-test results (averages and standard deviations) in the Representation Test are presented in Table 2. There are 23 questions in the Representation Test so the maximum score is 23: one point was awarded for each scientifically correct answer.

TABLE 2. Students' (n =54) results in the Representation Test. Standard deviations are given in the parentheses.

Pre-test (%)	Post-test (%)	Hake's gain	Cohen's d
52.7 (26.7)	78.3 (19.9)	0.54	0.77

Hake's gain (the average normalized gain) and Cohen's d (effect size) indicate that the students did improve their command of the concepts and representations addressed by the Representation Test reasonably well. There was almost no correlation (r = -0.06) between the pre-Representation Test scores and individual students' normalized gains. This indicates that the students' initial knowledge state regarding the gravitation and force concepts and their multiple representations did not affect students' conceptual gains.

Students' Representational Coherence

Students' representational coherence results in the pre-test and post-test are presented in Table 3 and in Table 4, respectively.

TABLE 3. Students' (n = 54) results in the matched pre- and post-tests in terms of representationally correct classes.

Representationally correct	I (%) coherent	II (%) moderately coherent	III (%) incoherent
Pre-test	20.4	42.6	37.0
Post-test	59.3	29.6	11.1

TABLE 4. Students' (n = 54) results in the matched pre- and post-tests in terms of scientifically correct classes.

Scientifically correct	I (%) coherent	II (%) moderately coherent	III (%) incoherent
Pre-test	7.4	5.6	87.0
Post-test	24.1	25.9	50.0

There is a clear shift from the pre- to post-test results towards representational coherence in both classes. Students' coherent use of representations increased quite substantially (from 20.4% to 59.3%). However, the change towards both representationally and scientifically correct understanding is not so prominent (from 7.4% to 24.1%).

Differences in Students' Ability to Use Various Representations

There were some statistically significant differences in students' ability to use different representations with scientific correctness. In the pre-test in theme FCI28 (Newton's third law) students had fewer correct answers in graphical representation (31.5%) than in other representations (verbal 66.7%, vectorial 63.0% and diagrammatic 64.8%). The differences were statistically significant (McNemar's test: $p < 0.001$). Also in theme FCI13 (gravitation, forces) students had fewer correct answers in graphical representation (9.3%) than in other representations (verbal 27.8% and vectorial 20.4%). However, only the difference between graphical and verbal representations was statistically significant ($p = 0.013$).

In the post-test there was only one theme in which differences in students' ability to use different representations with scientific correctness were statistically significant. In theme FCI13 83.3% of students chose the scientifically correct answer in the verbal question whereas 61.1% of students did so in the graphical question ($p = 0.04$).

Validity of the Representation Test

We were particularly interested in finding out how well the students could justify their answers. For this purpose, the written responses for each multiple choice were examined. The results indicate that 91% of correct answers were accompanied by correct explanations (7% had partially correct explanations). Hence, the number of clear false positives is very small (2% of all correct answers). However, the number of false negatives is surprisingly high (14% of all incorrect answers). One possible reason for some false negatives might be that some students made mistakes in writing down the answers on the answer sheets; this seems likely in some cases where the verbal explanation was perfect and did not match the chosen answer at all.

CONCLUSIONS

The analysis of students' representational coherence in the post-test revealed that over half of the students could consistently answer all questions in each theme in the same way after the teaching, i.e. exhibited representationally correct coherence, either scientifically correct or not. These results suggest that students can use representations quite consistently even if they do not understand the underlying physics correctly. The scientifically correct coherence class was, however, significantly smaller: about a quarter of the students mastered both representations and the concepts addressed by the post-test. It should be noted that the results provide genuinely new information which could not be deduced from the students' averages in the Representation Test.

Our pre-test results show that there were statistically significant differences between the representations in the three themes, graphical representation being harder for the students than other representations. Our findings on the pre-test results provide evidence that students' performance may vary with the representation even if the context of a problem is very similar. This is well in line with the findings of Meltzer [2] and Kohl & Finkelstein [3]. However, in the post test only one theme had statistically significant differences between the representations: the verbal question was easier for the students than the corresponding graphical question. This suggests that teaching can decrease students' difficulties with multiple representations at least to some extent. The post-test results also suggest that the students' command of physics was more constrained in terms of scientific correctness than in terms of multiple representations.

The written responses provided support for the validity of the test: they show that the test scores do reflect the students' understanding very well. This suggests that the Representation Test is a useful instrument for both researchers and teachers in determining students' ability to use multiple representations in the context of gravitation and Newton's third law.

REFERENCES

1. A. Van Heuvelen and X. Zou. "Multiple Representations of Work-energy Processes", *Am. J. Phys.* **69**(2). 184-194 (2001).
2. D.E. Meltzer. "Relation between students' problem-solving performance and representational format", *Am. J. Phys.* **73**(12). 463-478 (2005).
3. P.B. Kohl and N.D. Finkelstein. "Student representational competence and self-assessment when solving physics problems", *Phys.Rev.ST PER* **1**, 010104 (2005).
4. P.B. Kohl and N.D. Finkelstein. "Effect of instructional environment on physics students' representational skills", *Phys.Rev.STPER* **2**, 010102 (2006).
5. I. Halloun, R.R. Hake, E. Mosca and D. Hestenes. Force Concept Inventory (Revised 1995). Password protected at <http://modeling.asu.edu/R&E/Research.html>.
6. A. Savinainen. "High school students' conceptual coherence of qualitative knowledge in the case of the force concept". Ph.D thesis, University of Joensuu, 2004.
7. P. Nieminen, A. Savinainen and J. Viiri. "Investigating Students' Representational Coherence: Developing a Multiple Choice Test". Submitted for publication.
8. D.E. Meltzer (private communication).