

Toward a description of upper secondary physics students' modeling competency

Øystein Guttersrud
oysteg@fys.uio.no

Abstract

An achievement test assessing abilities to reason scientifically and interchange between representations of physical phenomena was developed to describe physics students' modeling competency. Students' competency is described using four proficiency levels. Results indicate that aspects of the nature of science (NOS) and the learning strategy "elaboration" are positively associated with performance on the achievement test.

Introduction

Physics has a long history of being perceived as the most incomprehensible (Osborne & Collins, 2000) and hardest of school subjects (Orton & Roper, 2000). In physics, phenomena are described using multiple representations (e.g. concepts, graphs and formulas). Dolin (2002) suggests the continuous interchanging between the different representations as an important reason why students perceive physics as so demanding.

Project PHYS 21 set focus on the key competencies empirical-mathematical modeling and scientific reasoning. Schools which took part in the project followed a modified version of the regular Norwegian upper secondary physics curriculum (Angell, Guttersrud, Henriksen, & Kind, 2006). As part of this project the PHYS 21 achievement test was developed to profile physics students' modeling skills and relates outcomes to students' epistemological beliefs, learning styles and use of multiple representations of physical phenomena.

The aim of this paper is, based on the achievement test which was part of the PHYS 21 student assessment program (PHYSAP), to give a description of physics students' skills to reason scientifically and interchange between multiple representations of physical phenomena (graphs, formulas, concepts etc.). These two dimensions are used to describe physics students' *modeling competency*. A glance at students' epistemological beliefs (NOS) and learning strategies and how these relate to performance on the PHYS 21 achievement test is also presented.

Method

Population and sample

Students attending the regular and the modified version (PHYS 21) of the Norwegian upper secondary physics course constitute the target population. All 289 PHYS 21 students at six schools and 240 students at nine other schools were sampled. The sampling procedure makes it impossible to generalize from the sample to the population of all physics students in Norway. The data collected can only describe those individuals assessed, but may offer some insight into physics students' knowledge and learning.

Instruments part A: Unit contexts and item formats

Instruments were designed to collect valid data according to the research question. A field trial had been conducted in 2004 - 2005 to investigate the achievement items' psychometric properties.

Students' views of the nature of science, components of self-regulation and the use of interchanges between different forms of representations during physics lessons were assessed by a questionnaire.

Each assessment item was part of a unit made up of a stem and 3 – 9 items related to the theme of the stem. The stem may e.g. describe a current issue related to scientifically investigable questions in a global (e.g. environmental), social (e.g. power supply) or personal (everyday experiences) context.

Four types of item formats were used. The first type of selected response items in figure 1 is *vector* items. Vector items ask students to agree or disagree on e.g. a set of assertions presented in a table. Students are supposed to select one out of two given choices for each assertion: “yes”/“no” or “agree”/“disagree” etc. The second type is the *multiple choice* (MC) items which offer four alternatives for the students to choose from. The first type of constructed response items is the *short constructed response* which students respond to by writing a single word or number. The second type is the *extended constructed response* where students have to write an answer over one or more lines in their own words.

Figure 1: Distribution of item formats

<i>Classes</i>	<i>Formats</i>	<i>Items</i>
selected response	vector	8
	multiple choice	6
constructed response	short constructed response	8
	extended constructed response	7

Students’ total score was calculated from response to the 29 items in figure 1. Of these 17 were scored dichotomously, while 12 were scored 0, 1, and 2 points resulting in a maximum score of 41 points. Mean score was 19.6 points with a standard deviation of 6.8 points.

Instruments part B: how modeling competency is measured by PHYSAP

The concept of modeling in PHYS 21 has two dimensions, which directed the development of the assessment items: a representational dimension and a reasoning dimension. The first of these dimensions reports on students’ ability to interchange between multiple representations of physical phenomena, while the second dimension measures students’ ability to reason scientifically. The five forms of representations taking part of these interchanges are, together with the five scientific reasoning processes, described shortly in figure 2.

The five forms of representations in figure 2a constitute ten categories of *interchanges* between pairs of representations (figure 3a). Each test item assessed one of these interchanges and concurrently one of the reasoning processes (figure 3b). For example, four items assessed interchange between the graphical and the experimental representation (element 1.1 in figure 3a). These items account for five of the points scored on the PHYS 21 achievement test.

Analyses

Sampling procedures decide the inferential statistics used when analyzing data. In this study it was not just convenient, but necessary to study subjects in naturally occurring groups, or clusters. Distribution-free techniques, so-called non-parametric tests, were used solely as such tests do not presuppose a normal distribution of randomly selected individuals.

Likert scales (Crocker & Algina, 1986) with four and five numerical values were used for all single items in the questionnaire. Each single item represents a discontinuous ordinal variable and was for this reason analyzed by comparing observed and expected count using Pearson Chi-square test.

Results

Defining proficiency levels

The representational and reasoning dimensions were merged into a combined scale reporting on students’ modeling skills. The score on the scale is interpreted to represent degree of competency to model physical phenomena.

Figure 2a: forms of representations

Representations
<i>Graphical representation</i> refers to graphs and other descriptive representations of variables.
<i>Pictorial representation</i> refers to all kinds of figurative descriptions except graphs.
<i>Mathematical representation</i> includes equations and the mathematical operations on these.
<i>Experimental representation</i> refers to all practical approaches.
<i>Conceptual representation</i> deals with the concepts used to describe phenomena inclusive verbal descriptions of phenomena using <i>scientific</i> concepts.

Figure 2b: scientific reasoning processes

Reasoning processes
<i>Analyze/categorize</i> : Analyze problems and categorize data to determine relationships between physical quantities.
<i>Generalize</i> : Describe physical relationships using general mathematical expressions and describe shared properties of physical formulae (e.g. linearity).
<i>Make decisions</i> : Select from alternative solutions and explanations in relation to evidence and data provided.
<i>Evaluate assertions</i> : Evaluate scientific claims in relation to evidence and data provided.
<i>Predict; justify; conclude; and, communicate</i> : Make predictions about effects of changes in physical systems, use evidence to justify problem solutions and draw and communicate valid conclusions.

Figure 3: The representational dimension's ten categories of interchange between pairs of representations (part a). The first digit in the parenthesis refers to the number of items (total 29) and the second to scored points (total 41) across the different categories of test items. Part b summarizes the five process categories of the reasoning dimension. The number of items and scored points are reported in separate columns.

Figure 3a	<i>Exper</i>	<i>Graphical</i>	<i>Math</i>	<i>Conceptual</i>
<i>Graphical</i>	(4,5)			
<i>Math</i>	(1,1)	(6,10)		
<i>Conceptual</i>	(3,4)	(8,11)	(2,3)	
<i>Pictorial</i>	(2,3)	(1,1)	(1,1)	(1,2)

Figure 3b	<i>Items</i>	<i>Points</i>
Analyze/categorize	5	6
Generalize	7	10
Make decisions	5	5
Evaluate assertions	7	11
Predict; justify; conclude; communicate	5	9

Four proficiency levels were empirically defined on the basis of student performance. The divisions between the proficiency levels were chosen to be as close as possible to the following distribution of students: 10%; 25%; 30%; 25%; and, 10%. The intermediate level, level 2, was designed to be symmetric around the mean score. In practice level 2 accounted for 31% of the students, while level 1 (low) and level 3 (high) accounted for 24% of the students each. Only 8% of the students were assigned level 4 (advanced), while 13% were

consigned to the “below level 1” group. Students who do not reach level 1 should not be interpreted as having no modeling skills at all, but they have severe difficulties in applying the multiple representations and reasoning skills considered necessary to acquire understanding of physics.

Anchoring at proficiency levels

Each item was assigned a p-value (percentage correct response) at each of the four proficiency levels. If at least half of the students proficient at a certain level complete an open constructed item, the item is said to “anchor” at that level. Pure guessing makes it 25% probable that a student chooses the key response on a 4-choice MC item. A 4-choice item therefore anchors at a proficiency level if 62.5% ($50 + 50/4 = 62.5$) of the students tick off the key response (Crocker & Algina, 1986). Items anchoring at the same level thus lie in the same difficulty interval.

Example of item

The following item is taken from the unit “Sea level” which raises an environmental issue in a “global” context. The origin of the stem is “some students” who want to examine how the melting of ice around The South Pole and in the areas around The North Pole influence sea level. The stone in glass 2 represents the territories covered with ice at The South Pole (figure 4). The item assessed interchange between an experimental and a mathematical representation of the phenomenon. The pictures of the glasses were used to extend and explain the experimental representation. As the mathematical representations were articulated through general mathematical expressions, the item was categorized as a “generalization” according to the reasoning dimension. Approximately one-third of the students answered this item, which anchored at proficiency level 2, correctly.

Figure 4: Example of an item assessing students’ modeling competency as measured by the PHYSAP instrument. The SPSS output displays mean z-score for the group of N students choosing the key response (A), the three distracters, other type of answers and blank responses.

Before melting

Glass 1 Glass 2

5 cm 5 cm

After melting

Glass 1 Glass 2

5 cm 5.5 cm

Assume that the ice is melting with a constant rate. Which mathematical expression describes the water level (y) in glass 1 and glass 2 while the ice melts?

A Glass 1: $y = b$, glass 2: $y = ax + b$

B Glass 1: $y = ax + b$, glass 2: $y = b$

C Glass 1: $y = b$, glass 2: $y = ax$

D Glass 1: $y = ax$, glass 2: $y = b$

Zscore(SUM)		
Response	Mean	N
A	.23	301
B	-.63	26
C	-.27	68
D	-.34	14
Other	-.42	4
Blank	-.84	33
Total		446

Example of verbal description of a proficiency level

Items anchoring at a certain level have equivalent difficulty and the proficiency associated with that level is described as a summary of what these items require (figure 5). These summaries thus make it, to some extent, possible to describe what students at each proficiency level “can do”. Students at a specific level not only possess the skills associated with that level but also the expertise entailed at levels below. The description in figure 5 reflects the skills assessed by the items anchoring at proficiency level 2 and are related to both the representational dimension and reasoning dimension.

Figure 5: Characteristics of items anchoring at proficiency level 2

Level	Dimension	Description
2 Intermediate	Representational	Interchange between few forms of representations simultaneously, i.e. the line of arguments includes few steps. These interchanges may include first order but no second order mathematical representations. (Relate $y = ax + b$ to changes in physical <i>systems</i> , link $y = k$ to constant graphs and formulae as $F = ma$ and $F = qE$ to straight lines passing the origin).
	Reasoning	Analyze experimental data to test hypothesis; express experimental situations using general first order mathematical representations; and, evaluate straightforward assertions.

Discussion: can physics teachers do something to make a difference?

This report does not claim to provide underlying links between what teachers do and how their students perform, but by identifying factors which interact to influence performance some clues about factors related to success on the PHYS 21 achievement test may be presented to physics teachers and teacher educators.

The importance of being self-regulated

The multifaceted construct “Self-regulated learning” has been defined and described in different studies (see e.g. Boekaerts, 1999; Pintrich, 2000). The PHYSAP instrument operationalised among others the learning strategies *memorizing* and *elaboration*. Frequent use of elaboration strategies has a tendency to be positively associated with performance on the PHYS 21 achievement test.

The construct *elaboration* is derived from responses to items asking for the frequency with which the student *relates new physics knowledge to: prior general knowledge; prior knowledge in physics; and, prior knowledge in mathematics*. How new physics knowledge might be used in the “real” world was also part of the construct. The Likert scale 1= Almost never; 2= Sometimes; 3 = Often; and, 4 = Almost always were used for all items operationalising the two learning strategy constructs.

The importance of possessing sound epistemological beliefs

The concept of “nature of science” (NOS) has repeatedly been described in the literature (see e.g. Abd-El-Khalick & Lederman, 2000; Lederman, Wade, & Bell, 1998). The PHYSAP instrument assessed some primary components of NOS using a questionnaire. Students were asked to consider more than 30 statements regarding science’s tentative; creative; empirical; and, objective nature. Students’ beliefs about scientific laws, theories and models were also explored. Displaying sophisticated NOS-views correlated positively with elaboration strategies.

It is not an easy task to decide whether *knowledge about science* activates *elaborating strategies* or if “clever” reflected students develop proper and consistent ideas about science as they gain insight into science’s products (i.e. scientific concepts, laws and theories) and processes (i.e. science’s methods, techniques and procedures). An important part of project PHYS 21 was to make students work scientifically and build “mathematical models” (i.e.

make mathematical representations from experimental data) of physical systems and thereby gain insight into science's processes and secondly arrive at more sophisticated epistemological beliefs. Data do not imply such relationships as very few items measuring knowledge about the nature of science correlate "strongly" with PHYS 21 attendance. Epistemological beliefs may however assist physics students' assessing their own learning assuming "proper" ideas about science have been taught and acquired.

References

- Abd-El-Khalick, F., & Lederman, N. G. (2000). Improving science teachers' conceptions of the nature of science: A critical review of the literature. *International Journal of Science Education*, 22(7), 665 - 701.
- Angell, C., Guttersrud, Ø., Henriksen, E. K., & Kind, P. M. (2006). *Rationale for and implementation of an empirical-mathematical modelling approach in upper secondary physics in Norway*. Paper presented at the GIREP conference, Amsterdam.
- Boekaerts, M. (1999). Self-regulated learning: where we are today. *International Journal of Educational Research*, 31, 445 - 457.
- Crocker, L., & Algina, J. (1986). *Introduction to Classical and Modern Test Theory* (first ed.). London.
- Dolin, J. (2002). *Fysikfaget i forandring, Læring og undervisning i fysik i gymnasiet med fokus på dialogiske processer, autencitet og kompetenceudvikling*. Roskilde Universitet, Roskilde.
- Lederman, N. G., Wade, P. D., & Bell, R. L. (1998). Assessing understanding of the nature of science: A historical perspective. In M. C. W (Ed.), *The nature of science in science education: Rationales and strategies* (pp. 331 - 350). Dordrecht: Kluwer Academic Publishers.
- Orton, T., & Roper, T. (2000). Science and Mathematics: A Relationship in Need of Counselling? *Studies in Science Education* 35, 123 - 154.
- Osborne, J., & Collins, S. (2000). Pupils' and Parents' View of the School Science Curriculum.
- Pintrich, P. R. (2000). The Role of Goal Orientation In Self-Regulated Learning. In Boekaerts, Pintrich & Zeidner (Eds.), *Handbook of self-regulation* (pp. 451-502) (pp. 451 - 502). San Diego: Academic Press.