

Concept Inventories

Project Mosaic Kick-Off Workshop

Daniel Kaplan

July 4, 2010

What's a Concept Inventory?

*Multiple-choice test, **few or no calculations.***

*Designed to detect a **working knowledge of a concept.***

*Distractors based on commonly held **misconceptions.***

Example: The Force Concept Inventory

“The *Force Concept Inventory* (FCI) is a unique kind of “test” designed to assess student understanding of the *most basic* concepts in Newtonian physics. It can be used for several different purposes, but the most important one is to evaluate the effectiveness of instruction. For that purpose, the FCI is probably the most widely used instrument in physics education today.” — Hestenes and Halloun, [5]

- ▶ Published in 1992. [7]
- ▶ Taken by tens of thousands of physics students.
- ▶ Students do surprisingly poorly.

“When he first heard about the FCI, applied physicist Eric Mazur of Harvard University in Cambridge, Massachusetts, assumed that his elite students would perform perfectly well in the traditional lecture setting. So when they received an average FCI score of 70, where 80 is considered a pass, he got ‘a slap in the face’.”

An Example from the Force Concept Inventory

FIGURE REDACTED to avoid publishing a question that's in use.

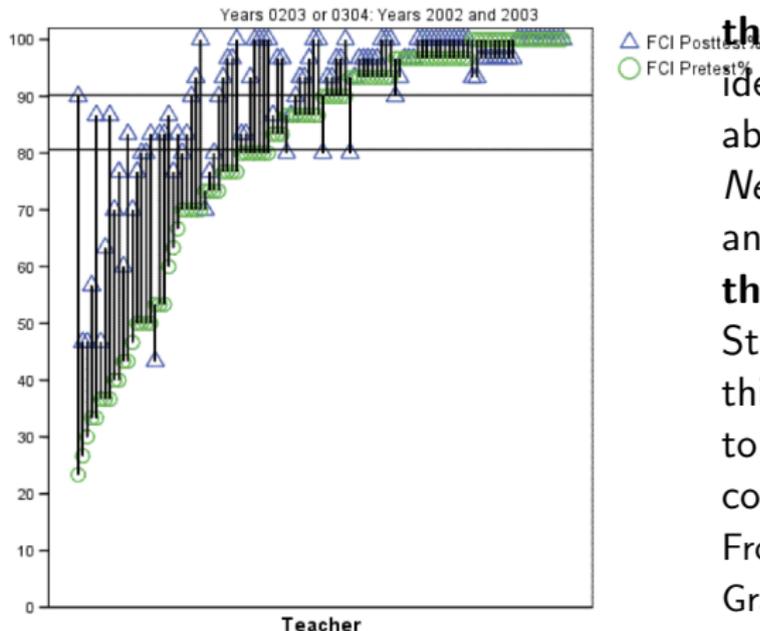
The figure shows a three-quarter circular channel in which a ball can glide, along with 4 different paths that the ball might take on exiting the channel.

Which path in the figure above would the ball most closely follow after it exits the channel at "r" and moves across the frictionless table top?

Concepts, not Technical Vocabulary

“Physicists have developed a technical language for precise expression of scientific concepts and unambiguous description of physical situations. Unfortunately, until students are privy to their special meanings, technical terms can be a barrier rather than a help to understanding. Consequently, on examinations students often respond to the form of the technical language rather than its meaning. For example, for a typical University Physics course we found that nearly 80% of the students could state Newtons Third Law at the beginning of the course, while FCI data showed that less than 15% of them fully understood it at the end. In designing FCI questions we tried to avoid technical language in order to get closer to what students really think. We reasoned that Newtonian thinkers would be able to resolve the consequent imprecision and ambiguities.” — David Hestenes, [4]

How High-School Instructors Perform



rs 2002 and 2003 Mean (SD): Pretest%: 80.6 (21.2) Posttest%: 90.2 (13.0)

From [6]

“We interpret an FCI score of 85% as the Newtonian **Mastery threshold**. We are confident in identifying students with scores above this as *confirmed Newtonian thinkers*. We suggest an FCI score of 60% as the **entry threshold** to Newtonian physics. Students who have just reached this threshold have barely begun to use Newtonian concepts coherently in their reasoning.” [5]

From: Findings of the ASU Summer Graduate Program for Physics

Teachers(2002-2006) -a section in the Final Report submitted to the NSF in 2006, by David Hestenes and Jane Jackson

Some Other Concept Inventories

- ▶ Calculus
- ▶ Statistics
- ▶ Chemistry
- ▶ Astronomy
- ▶ Basic Biology
- ▶ Natural Selection
- ▶ Genetics

An Example from the Calculus Concept Inventory

Developed by Jerome Epstein [3, 2].

If you know that a function $f(x)$ is positive everywhere, what can you conclude from that about the derivative $f'(x)$?:

1. the derivative is positive everywhere
2. the derivative is increasing everywhere
3. the derivative is concave upward,
4. you cant conclude anything about the derivative

[Example question published at http://www.flaguide.org/tools/diagnostic/calculus_concept_inventory.php]

I'll hand out a paper copy of the CCI, following Dr. Epstein's requests regarding security.

Statistics Concept Inventory

From an early publication: [9]

The following are temperatures for a week in August: 94, 93, 98, 101, 98, 96, and 93. By how much could the highest temperature increase without changing the median?

1. Increase by 8°
2. Increase by 2°
3. It can increase by any amount
4. It cannot increase without changing the median.

Statistics Concept Inventory

A researcher performs a t-test to test the following hypotheses:

$$H_0 : \mu \leq \mu_0$$

$$H_1 : \mu > \mu_0$$

He rejects the null hypothesis and reports a p-value of 0.1. Which of the following must be true?

1. The test statistic fell within the rejection region at the $\alpha = 0.05$ significance level.
2. The power of the test statistic used was 90%.
3. There is a 10% possibility that the observed value is due to chance.
4. The probability that the null hypothesis is not true is 0.1.
5. The probability that the null hypothesis is actually true is 0.9.

But ... is this a concept inventory problem?

Problems with the hypothesis-testing example

- ▶ It's based on technical vocabulary (but veers to informal, e.g., “possibility”)
- ▶ It's abstract. Saying “A researcher performs ...” does not allow intuition to apply.
- ▶ It aims at technical misconceptions (e.g., “the prob. that the null hypothesis is not true”) rather than misconceptions that prevent effective work.
- ▶ It's rooted in the particular approach used to teach hypothesis testing (inference on means) rather than the root concepts of hypothesis testing.

To modify Hestenes: “Physicists Statisticians have developed a technical language for precise expression of scientific statistical concepts and unambiguous description of physical situations statistical methods. Unfortunately, until students are privy to their special meanings, technical terms can be a barrier rather than a help to understanding. Consequently, on examinations students often respond to the form of the technical language rather than its meaning.”

The CAOS Statistics Test I

A 40 item multiple-choice test for students enrolled in a college-level non-mathematical first course in statistics. [1]

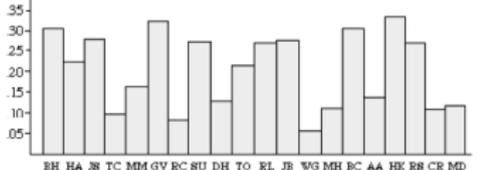
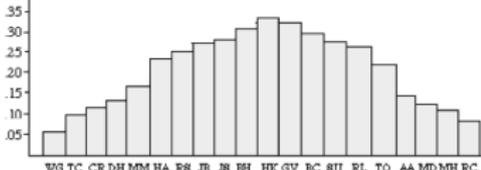
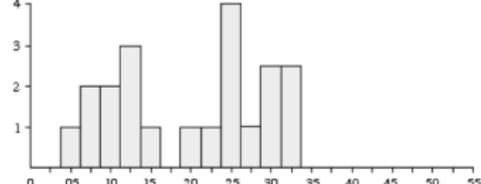
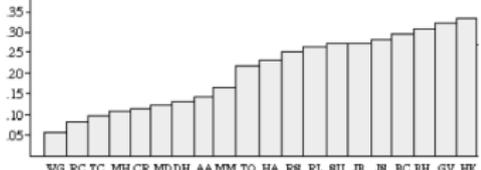
6. A baseball fan likes to keep track of statistics for the local high school baseball team. One of the statistics she recorded is the proportion of hits obtained by each player based on the number of times at bat as shown in the table below. Which of the following graphs gives the best display of the distribution of proportion of hits in that it allows the baseball fan to describe the shape, center and spread of the variable, proportion of hits?

Player	Proportion of hits
BH	0.305
HA	0.229
JS	0.281
TC	0.097
MM	0.167
GV	0.333
RC	0.085

Player	Proportion of hits
SU	0.270
DH	0.136
TO	0.218
RL	0.267
JB	0.270
WG	0.054
MH	0.108

Player	Proportion of hits
BC	0.301
AA	0.143
HK	0.341
RS	0.261
CR	0.115
MD	0.125

CAOS Question Choices and Results I

RESPONSE OPTION	Percent (N = 10,729)
<p>A</p>  <p>EH HA NS TC MM GV RC SU DH TO RL JB WG MH BC AA HK RS CR MD</p>	<p>11.7%</p>
<p>B</p>  <p>VG TC CR DH MM HA RS JB NS EH HK GV BC SU RL TO AA MD MH RC</p>	<p>55.2%</p>
<p>C</p>  <p>0 05 10 15 20 25 30 35 40 45 50 55</p>	<p>21.5%</p>
<p>D</p>  <p>VG RC TC MH CR MD DH AA MM TO HA RS RL SU JB NS BC EH GV HK</p>	<p>11.6%</p>

CAOS Question Choices and Results II

Note how most students make a category mistake — they don't recognize that in a histogram display the “cases” go on the vertical axis and the value for each case are on the horizontal axis.

They have been taught to look for bell-shaped distributions and that's what they do.

NOTE: I actually like graph D as the most informative. But I know how to read from a cumulative distribution.

The CAOS Student Learning Goals I

1. Ability to describe and interpret the overall distribution of a variable as displayed in a histogram, including referring to the context of the data.
2. Ability to recognize two different graphical representations of the same data (boxplot and histogram).
3. Ability to visualize and match a histogram to a description of a variable (negatively skewed distribution for scores on an easy quiz).
4. Ability to visualize and match a histogram to a description of a variable (bell-shaped distribution for wrist circumferences of newborn female infants).
5. Ability to visualize and match a histogram to a description of a variable (uniform distribution for the last digit of phone numbers sampled from a phone book).

The CAOS Student Learning Goals II

6. Understanding to properly describe the distribution of a quantitative variable (shape, center, and spread), need a graph like a histogram which places the variable along the horizontal axis and frequency along the vertical axis.
7. Understanding of the purpose of randomization in an experiment.
8. Understanding that boxplots do not provide accurate estimates for percentages of data above or below values except for the quartiles.
and so on.

Some CAOS Identified Misconceptions

1. Random assignment is confused with random sampling or thinks that random assignment reduces sampling error.
2. When comparing histograms, the graph with the largest number of different values has the larger standard deviation (spread not considered).
3. Causation can be inferred from correlation.
4. Grand totals are used to calculate conditional probabilities.
5. Rejecting the null hypothesis means that the null hypothesis is definitely false.

From [1].

The FCI Concepts

Table I. Newtonian Concepts in the Inventory.

	<u>Inventory Item</u>
0. Kinematics	
Velocity discriminated from position	20E
Acceleration discriminated from velocity	21D
Constant acceleration entails parabolic orbit	23D, 24E
changing speed	25B
Vector addition of velocities	(7E)
I. First Law	
with no force	4B, (6B), 10B
velocity direction constant	26B
speed constant	8A, 27 A
with cancelling forces	18B,28C
2. Second Law	
Impulsive force	(6B), (7E)
Constant force implies constant acceleration	24E, 25B
3. Third Law	
for impulsive forces	2E, 11E
for continuous forces	13A, 14A
4. Superposition Principle	
Vector sum	19B
Cancelling forces	(9D), 18B, 28C
5. Kinds or Force	
5S. Solid contact	
passive	(9D), (12 B.D)

Each concept is keyed to one or more specific answers in the FCI.[7]

The FCI Misconception Taxonomy

Table II. A Taxonomy of Misconceptions Probed by the Inventory. Presence of the misconceptions is suggested by selection of the corresponding Inventory Item.

	Inventory Item
0. Kinematics	
K1. position-velocity undiscriminated	208,C,D
K2. velocity-acceleration undiscriminated	20A; 21B,C
K3. nonvectorial velocity composition	7C
1. Impetus	
I1. impetus supplied by "hit"	9B,C; 22B,C,E; 29D
I2. loss/recovery of original impetus	4D; 6C,E; 24A; 26A,D,E
I3. impetus dissipation	5A,8,C; 8C; 16C,D; 23E; 27C,E; 29B
I4. gradual/delayed impetus build-up	6D; 8B,D; 24D; 29E
I5. circular impetus	4A,D; 10A
2. Active Force	
AF1. only active agents exert forces	11B; 12B; 13D; 14D; 15A,B; 18D; 22A
AF2. motion implies active force	29A
AF3. no motion implies no force	12E
AF4. velocity proportional to applied force	25A; 28A
AF5. acceleration implies increasing force	17B
AF6. force causes acceleration to terminal velocity	17A; 25D
AF7. active force wears out	25C,E
3. Action/Reaction Pairs	
AR1. greater mass implies greater force	2A,D; 11D; 13B; 14B
AR2. most active agent produces greatest force	13C; 11D; 14C
4. Concatenation of Influences	
CI1. largest force determines motion	18A-F; 19A

Each misconception is keyed to one or more specific answers in the FCI.[7]

The Process of Developing a Concept Inventory I

General observations:

- ▶ It's not easy.
- ▶ It's expensive and time consuming.

Calculus concept inventory was developed in part with a grant from NSF of \$350K, similar to the entire project grant for MOSAIC.

“The process begins with the identification of the *fundamental constructs* the test is designed to measure. One then puts together a panel of item writers with expertise in the subject matter and, hopefully, some knowledge of the principles of good assessment.”

[2]

The Process of Developing a Concept Inventory II

“*Cognitive Laboratories* are of great help in knowing what test items are really measuring, and they were used in the validation of the CCI. Scores on items and on tests can tell a lot when properly analyzed, but it is surely true that students get right answers for wrong reasons and can get wrong answers that are at least in part the fault of the item. Cognitive labs (sometimes called “analytic interviews”) are a marvelous technique to discover this. They are a highly structured interview technique where individual students are asked to think out loud as they work on a problem. Probing questions are then used to access the student mental process (not to tutor the student!). These probing questions for each item are contained in a carefully designed protocol. It is subtle to design this protocol. We utilized consultant services to do this for the CCI. One wants to use a Lab on an item with poor discrimination (good students got it wrong and/or poor students got it right), but also on a few items that perform well, to be sure that students are not getting right answers for the wrong reasons or getting wrong answers due to a problem in wording the item.” [2]

The Process of Developing a Concept Inventory III

“We gave the first pilot test of the CCI to about 250 students at 6 schools in the Spring of 2005. There was *no gain anywhere*, and scores were near the random guess level of 20% (even at post-test). This shocked us. Extensive discussion among the panel led to a significant modification to the test, and to making it considerably easier. The conclusion was that if most faculty believe the test is trivial, we are probably about right.” [2]

Purposes of Concept Inventories

- ▶ Investigate effectiveness of pedagogical techniques. E.g., traditional lecture versus “active” learning. [8]
“Mathematics education is mired in the ‘math wars’ between ‘back-to-basics’ advocates and ‘guided-discovery’ believers. There is no possibility of any resolution to this contest between competing faiths without scientific evidence of what works and what doesn’t.
- ▶ Investigate effect of intervention, e.g., summer workshops for instructors.[6]

These purposes require that the tests be psychometrically validated.

- ▶ Demonstrate to instructors that something is wrong.

This requires face validity — that the questions seem sensible to instructors and that an instructor believes that the questions are easy. So failure to answer correctly is evidence for a problem.

My Purpose for a Modeling Concept Inventory

Newtonian physics and Calculus are both areas with a long-established canon of what should be taught, standard textbooks, and extensive experience for most instructors.

In modeling, it's still unclear what are the central concepts, and there is little consensus about how to teach them.

- ▶ Use the development of a concept inventory to organize a list of concepts.
- ▶ Use it to identify common misconceptions so that they can be addressed.
- ▶ Establish a case that modeling should be taught.
- ▶ Help to address the problem of empty claims of teaching modeling, that using models as a context is teaching modeling.

Agenda

1. In Small Groups: Identify a couple of misconceptions that students commonly have about modeling. Draft a question for each of these that probes whether students hold those misconceptions.
2. Report back on those questions.
3. Discuss what is an appropriate goal regarding a “modeling concept inventory.” What is realistic to accomplish? What process will help us the most?

References I

-  Robert delMas, Joan Garfield, Ann Ooms, and Beth Chance. Assessing students' conceptual understanding after a first course in statistics.
In Annual Meetings of the American Educational Research Association, 2006.
-  Jerome Epstein.
Development and validation of the calculus concept inventory.
via Internet.
-  Jerome Epstein.
The calculus concept inventory.
In Donald Deeds and Bruce Callen, editors, Proceedings of the National STEM Assessment Conference, pages 60–67. Drury University, October 2006.

References II



David Hestenes.

Who needs physics education research!?

American Journal of Physics, 66:465–467, 1998.



David Hestenes and Ibrahim Halloun.

Interpreting the force concept inventory.

The Physics Teacher, 33:502–506, 1995.



David Hestenes and Jane Jackson.

Findings of the asu summer graduate program for physics teachers (2002-2006).

Technical report, Arizona State University, 2006.

a section in the Final Report submitted to the National Science Foundation.



David Hestenes, Malcolm Wells, and Gregg Swackhamer.

Force concept inventory.

The Physics Teacher, 30:141–158, March 1992.

References III



Kendall Powell.

Science education: Spare me the lecture.

Nature, 425:234–236, 2003.



Andrea Stone, Kirk Allen, Teri Reed Rhoads, Teri J Murphy,
Randa L Shehab, and Chaitanya Saha.

The statistics concept inventory: A pilot study.

In *ASEE/IEEE Frontiers in Education Conference*, pages
T3D1–6, 2003.