Two steps forward, one step back

Nathaniel Lasry, Jonathan Guillemette and Eric Mazur

Among physics students there exists a wide variety of misconceptions, generally thought to be robust and resistant to change. But our analysis of the path of progress has changed our conception of how students learn physics.

The French novelist Marcel Proust once wrote: "The real voyage of discovery consists not in seeking new landscapes, but in seeing with new eyes." One of the main goals of science education is to help students change the way they see the world; particularly in physics because students often have views that are at odds with what we try to teach them. Instructors of introductory physics courses routinely assess their students' conceptual understanding by giving tests such as the Force Concept Inventory[1], or FCI. The FCI consists of 30 multiple-choice questions phrased in simple language and designed so that the solutions don't require any calculations. All of the incorrect answers among the multiple choices are based on those most frequently given by students in interviews. Hence, students are less likely to guess an answer because they usually find one among the possibilities that matches their line of thinking.

For example, one of the questions requires students to complete the following: "A large truck collides head-on with a small compact car. During the collision…". In addition to the correct answer, the FCI presents common incorrect choices such as "the truck exerts a larger force on the car than the car on the truck". However, the FCI also provides the choice "neither exerts a force on the other, the car gets smashed simply because it gets in the way of the truck" — an answer given by students during interviews, which most instructors would never have thought of, let alone consider putting in a test.

To physicists, questions on the FCI seem trivial. Yet, many instructors are puzzled by how poorly their students perform. When one of us[2] first used the FCI at Harvard, a student asked: "How should I answer these questions? According to what you taught me, or according to the way I usually think about these things?" The counter-intuitive simplicity and documented reliability[3] of the FCI have contributed to making it the most widely used, cited and researched instrument in physics education.

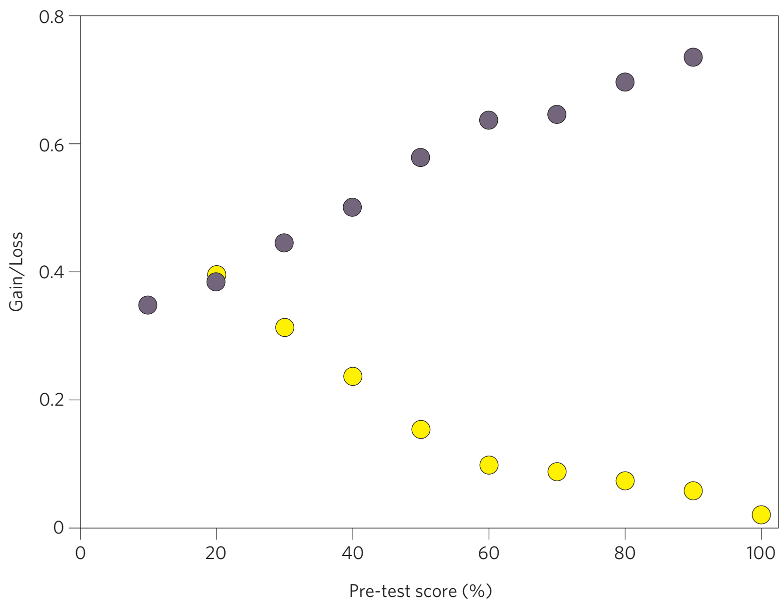

Figure 1 | The rich get richer. Gain (purple points) and loss (yellow points) as a function of pre-test score. The more prior knowledge students have, the more concepts they gain during their course of study and the fewer they lose.

Conceptual gains

Diagnostic tests like the FCI are frequently given both at the beginning and at the end of introductory physics courses to measure how much students have learned. This is often referred to as 'pre–post testing'. In a landmark study[4], Richard Hake measured the pre–post normalized gains for more than 6,000 college and university physics students tested using the FCI. He defined the normalized gain (g) as the difference between the post-test score (S_{post}) and the pre-test score (S_{pre}) divided by the maximum possible increase in score (or equivalently, the number of incorrect responses made in the pre-test): $g = (S_{post} - S_{pre})/(1 - S_{pre})$.

Hake found that students exposed to traditional teaching (such as lecturing) have significantly lower conceptual gains than students on courses that involve more active learning. The 'Hake gain' is widely used by instructors, who want to see how much conceptual learning their students achieve, and by researchers, who want to compare the difference in conceptual learning between groups exposed to different pedagogies.

A few years ago, we analysed several years of FCI data collected at Harvard and were puzzled by a curious finding — year after year many students were answering some FCI questions incorrectly at the end of the course, even though they had answered those questions correctly at the beginning of the course. This finding prompted us to systematically look at gains and losses for each FCI question. We defined gain (G) as the proportion of incorrect answers (I) on the pre-test that were changed to correct answers (C) on the post-test, or $G = IC/I_{pre}$. Similarly, we defined loss (L) as the proportion of correct answers on the pre-test that were changed to incorrect answers on the post-test, or $L = CI/C_{pre}$.

Table 1 | Functional Concept Inventory pre-test score (S_{pre}) and the gain and loss of student conceptual understanding for a range of institution types.

	Students tested	$\langle S_{pre} \rangle$	Gain (%)	Loss (%)
High school	10,007	26.7 ± 0.2	39.4 ± 0.1	30.4 ± 0.2
Public university	1,560	31.3 ± 0.4	32.5 ± 0.3	34.6 ± 0.4
Two-year college	971	35 ± 2	43 ± 1	28 ± 1
Top-tier university	884	47 ± 2	59 ± 3	15 ± 2

We wondered whether these losses were particular to Harvard students, because we expected Harvard students to have more initially correct answers to lose than most students. So we pooled data from more than 13,000 students who had taken the FCI at the beginning and at the end of an introductory physics course at a variety of institutions: US high schools, three Canadian two-year colleges, a US public university and three top-tier private universities. Averaging across all students in this sample, we found that 30% of all correct answers on the pre-test were changed to incorrect answers on the post-test. Students do realize gains — on average, 46% of all incorrect answers on the pre-test are changed to correct answers on the post-test. However, conceptual change in physics is not only about gains; losses must be considered too.

> "How should I answer these questions? According to what you taught me, or according to the way I usually think about these things?"

We consider that the gains and losses depend on the type of institution and on how much prior physics knowledge students have, which can be determined from S_{pre}. Table 1 shows the mean FCI pre-test score, as well as the proportion of gains and losses for each type of institution. Although the gains fluctuate widely across institutions (ranging roughly between 30% and 60%), the losses fluctuate little. Almost one-in-three correct answers on the pre-test are changed on the post-test in all institutions, except top-tier institutions, where the figure is much lower (15%). Figure 1 shows a plot of gains and losses as a function of the pre-test score. As S_{pre} increases, the gains increase whereas the losses decrease, both nearly linearly.

Diagnostic tests such as the FCI are designed to deter guessing as each incorrect answer is based on actual answers provided by students. Nevertheless, it is tempting to attribute the 30% loss to guessing on the pre-test. The amount of guessing in a test can be gauged using 'item response theory' by determining the likelihood that someone with no ability would get a correct answer. Such an analysis[5] of the FCI gives an average guessing rate of 14% — a value that is well below the 30% reported in Table 1. The instability of initially correct answers seems to be larger than the noise due to guessing.

Changing conceptions

The magnitude of the conceptual losses in Table 1, although surprising, can be understood in three ways. First, instruction may sometimes cause students to change from a Newtonian to a non-Newtonian conception. Students in introductory physics courses are routinely asked to challenge their preconceptions; some students may also challenge and set aside correct conceptions. For example, in a recent study[6] students were asked which would reach the ground first: a ball thrown downwards or an identical ball thrown horizontally from the same height? Only students who had taken physics got this wrong by answering that both would reach the ground at the same time. Therefore, part of the conceptual loss may be instruction-induced.

Instruction-induced loss, however, is probably the exception, not the rule. The second way to understand conceptual loss is by thinking of the nonlinear path that leads a novice to expertise — becoming an expert in a discipline is a complex process that does not progress linearly. The dependence of loss on prior knowledge, as shown in Fig. 1, suggests that expert-like Newtonian concepts are unstable in the minds of novice students. In fact, Fig. 1 indicates that the less experienced the student — that is, the lower the FCI pre-test score — the greater the loss and therefore the more unstable Newtonian concepts are for the student. This challenges the notion that student conceptions are robust and resistant to change, and argues in favour of a model[7] in which constructing expert conceptions is a fragile and context-dependent process.

The third way to understand conceptual loss is by relating it to prior knowledge. As shown in Table 1, students with higher prior knowledge have higher gain and smaller loss. This observation might be another instance of the so-called Matthew effect — the rich get richer and the poor get poorer. The Matthew effect is consistent with the idea[8] that concepts self-organize into conceptual networks that grow through preferential attachment. Therefore, the more concepts one starts with, the more probable that a new concept is integrated into the network and the less probable that it is lost. Our FCI data thus support the idea of conceptual change as the growth of a self-organizing network of concepts.

Indeed, the data paint an intriguing portrait of how students in introductory physics courses change their conceptions. Constructing an expert view of the physical world seems to be a nonlinear and fragile process, particularly when students enter a course with less expert knowledge. Correlating conceptual gains and losses with instructional approach — something we could not do with the present dataset — could provide interesting insight into how students learn and could highlight the pedagogies that are most efficient in helping students to gain more and lose less. ❐

Nathaniel Lasry and Eric Mazur are in the Department of Physics and School of Engineering and Applied Sciences, Harvard University, 29 Oxford Street, Cambridge, Massachusetts 02138, USA. Lasry and Jonathan Guillemette are in the Physics Department, John Abbott College, Montreal, Quebec H9X 3L9, Canada.
e-mail: mazur@seas.harvard.edu

References

1. Hestenes, D., Wells, M. & Swackhamer, G. *Phys. Teach.* **30,** 141–158 (1992).
2. Mazur, E. *Science* **323,** 50–51 (2009).
3. Lasry, N., Rosenfield, S., Dedic, H., Dahan, A. & Reshef, O. *Am. J. Phys.* **79,** 909–912 (2011).
4. Hake, R. R. *Am. J. Phys.* **66,** 64–74 (1998).
5. Wang, J. & Bao, L. *Am. J. Phys.* **78,** 1064–1070 (2010).
6. Hutchison, P. & Elby, A. *AIP Conf. Proc.* **1513,** 194–197 (2013).
7. Hammer, D., Elby, A., Scherr, R. E. & Redish, E. F. in *Transfer of learning from a modern multidisciplinary perspective* (ed. Mestre, J.) 89–120 (Information Age Publishing, 2005).
8. Barabási, A.-L. & Bonabeau, E. *Sci. Am.* **288,** 50–59 (2003).

Acknowledgements

The authors thank their collaborators in this work: M. Dugdale, S. Bhatnagar, C. Whittaker, E. Charles, K. Lenton, K. Ragan, G. Morris, T. Mzoughi and K. Miller; and D. Hestenes for providing the high-school FCI data.