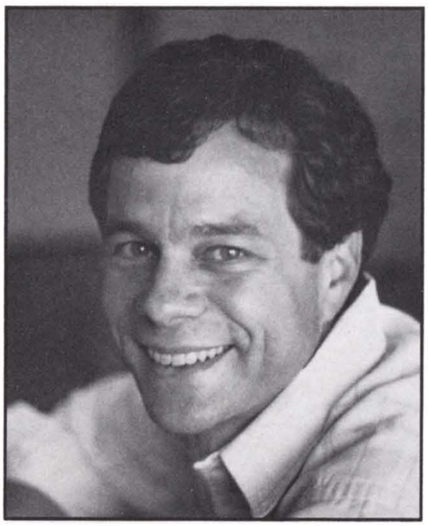
Teacher Predictions Versus Actual Student Gains

By Alan Lightman and Philip Sadler



Alan Lightman *is professor of science and writing, and senior lecturer in physics, at MIT (Program in Writing and Humanistic Studies, Room 14E-303, Cambridge, MA 02139-4307). He is also on the staff of the Harvard-Smithsonian Center for Astrophysics. He received his A.B. from Princeton and Ph.D. from the California Institute of Technology in physics. Some of his most recent books include* Ancient Light *and* Great Ideas in Physics.

For some time science educators have realized that meaningful learning in science involves the mastery of concepts rather than facts (e.g., Ref. 1). In turn, research on conceptual learning in science has shown that students enter the classroom with a wide range of deep-seated misconceptions about the natural world, and unless the science teacher directly confronts these erroneous prior beliefs, no amount of teaching the "correct" explanations of natural phenomena will have staying power beyond the last class quiz.[2-7] For example, since early childhood many students have firmly believed that forces are required to keep an object in motion, that heavy objects always fall faster than light ones, and that light is a condition rather than a discrete entity that moves through space. In short, students are not empty vessels waiting to be filled with knowledge.

Unfortunately, the majority of science teachers may still not appreciate the powerful grip of prior beliefs and misconceptions, or even their existence. Some studies have shown that teachers react with surprise when they discover that their students have basic misconceptions about science concepts they have already taught.[8-9] Several years ago, teachers and educators were stunned by the short documentary film "Private Universe,"[10] which showed the basic astronomical misconceptions of Harvard seniors.

We were thus motivated to investigate how well teachers can predict students' abilities to overcome science misconceptions through classroom instruction. In particular, we report here on a study of teachers' predictions of students' gains in knowledge after a traditional astronomy course. The knowledge level was sampled by 16 questions, many of which concern misconceptions. We then compare the teachers' predictions to the actual gains in a similar population of students.

Our results are clearly related to teacher effectiveness. If science teachers are indeed unaware of the misconceptions of their entering students and do not take these misconceptions into account, they may have overly optimistic expectations of the effectiveness of their teaching.

There has been a good deal of research by educational psychologists on how teachers' beliefs and expectations affect the behavior and performance of their students.[11] However, we have found practically no studies of teachers' predictions of the effectiveness of their own teaching. Bruce and Lawrenz[12] recently compared 50 chemistry teachers' predictions versus actuality of the mathematical knowledge of 276 high-school students in Minnesota. These teachers overestimated by an average of 24% the fraction of students who would be proficient on a test administered to them at the *beginning* of a chemistry course. In our study, we asked astronomy teachers to predict student performances on an astronomy test *both* on the first day of class and on the last day. In this way, we could directly measure the predicted *gains* in knowledge, and not simply the predicted initial level of knowledge.

Design and Administration of the Test Instrument

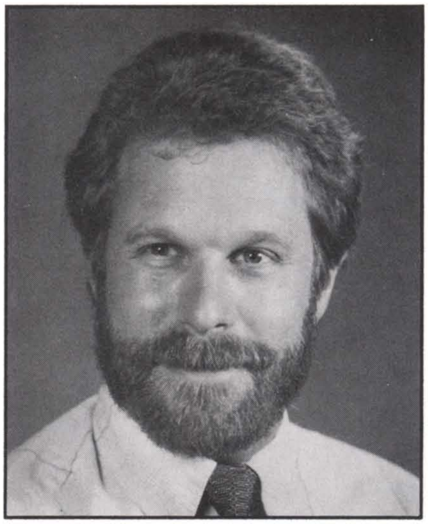
Our test instrument consisted of 16 questions in astronomy and related subjects, with multiple-choice responses. The test was not constructed to conform to any particular high-school astronomy course; rather it was designed around a number of ideas for which students might have strong prior misconceptions, and those misconceptions were included among the possible responses. The 16 items on the test evolved from our own interviews and tests of students and from previous research by others in student misconceptions in science.[13-21] The test instrument is shown in Appendix I (see p. 167).

For comparison with the conceptual questions, we also included some factual and quantitative questions. In particular, questions 1, 3, 4, 6, and 9 involve factual or quantitative knowledge, whereas the remaining questions require conceptual reasoning. Question 12, on astrology, doesn't easily fit into these categories.

We performed a "distractor" analysis to confirm the belief that many of the questions and responses indeed contain prevalent misconceptions. A distractor is an incorrect choice. Among the subset of students who choose an incorrect answer, if a particular distractor is chosen more frequently than would be expected by chance, then that question and distractor represent a misconception. Using the Gaussian approximation to the binomial distribution, we computed the probabilities of choosing distractors if those choices were made at random. The "maximum distractor" is that distractor chosen the largest number of times. The filled circles in Fig. 1 give that fraction of students choosing the maximum distractor above which the probability would be less than 1% if such a choice were made at random from among the four choices. The open circles give the actual fraction of students choosing the maximum distractor. Any question for which the open circle is higher than the closed circle is a misconception question. As can be seen, the incorrect choices for the question on gravity, number 16, appear to have been chosen at random. Almost all the remaining questions involve misconceptions, many of them (such as the question about light and filters, number 13), very strongly. The test was intentionally designed for this outcome.

In the spring of 1991, we sent half the test (8 questions) to 600 astronomy teachers and the other half to another 600; the teachers were drawn from our own STARNews mailing list, compiled from national workshops on astronomy teaching. We received responses from about 20% of teachers. After eliminating teachers with students outside of our grade range, there remained responses from 66 teachers on each half of the test, so that each of the 16 questions has predictions by 66 teachers. Most of the respondents were located in New York, Minnesota, Illinois, California, and Ohio. The great majority of teachers taught in public schools. Their students ranged from eighth grade to twelfth, with an average grade level of 9.5. Forty-six percent of the respondents taught a one-semester course in astronomy and 54% taught a two-semester course. For each question, the teachers were asked to predict the percentage of students who would get the correct answer on the first day of class and also on the last day of class.

For comparison of the teacher predictions, we administered the 16-question test to astronomy students chosen from the classrooms of a random subset of the teacher mailing list. The students came from public schools in Wisconsin, Minnesota, Massachusetts, Indiana, New Hampshire, Missouri, and Ohio. The students were given the test twice—once at the beginning of their course in astronomy and once at the end—over the academic year 1990-91. Fifty-two percent of the students were male and 48% female. A total of 1414 students took the test at the beginning of their astronomy course; of these, 330 went on to take a traditional astronomy course and were tested a second time at the end of their course. The post-course test results we report on here refer to these 330 students. (The other 1084 students took an innovative astronomy course, called Project STAR, that incorporates new research on misconceptions and educational strategies.) The average grade levels of the students taking the test before and after the astronomy course were 10.0 and 9.0, respectively. (Compare with the average grade taught by the teachers making the predictions: 9.5.) As we will see, these differences in grade level are insignificant.



Philip M. Sadler *received his B.S. in physics from MIT in 1973 and his M.A. in 1974 and Ed.D. in 1992 both from Harvard University. In 1977, he founded Learning Technologies Inc. to manufacture the Starlab Portable Planetarium and other educational equipment. In 1981, he helped found Computer Products Marketing Inc. which markets computer hardware and software throughout New England. He is currently the Frances W. Wright Lecturer on Celestial Navigation in the department of astronomy at Harvard University, the director of education projects at the Harvard-Smithsonian Center for Astrophysics (60 Garden Street, MS 71, Cambridge, MA 02138), and assistant professor of education at the Harvard Graduate School of Education. He has worked with his colleagues at the Center for Astrophysics on Project STAR, Project SPICA, MicroObservatory, and InSIGHT.*

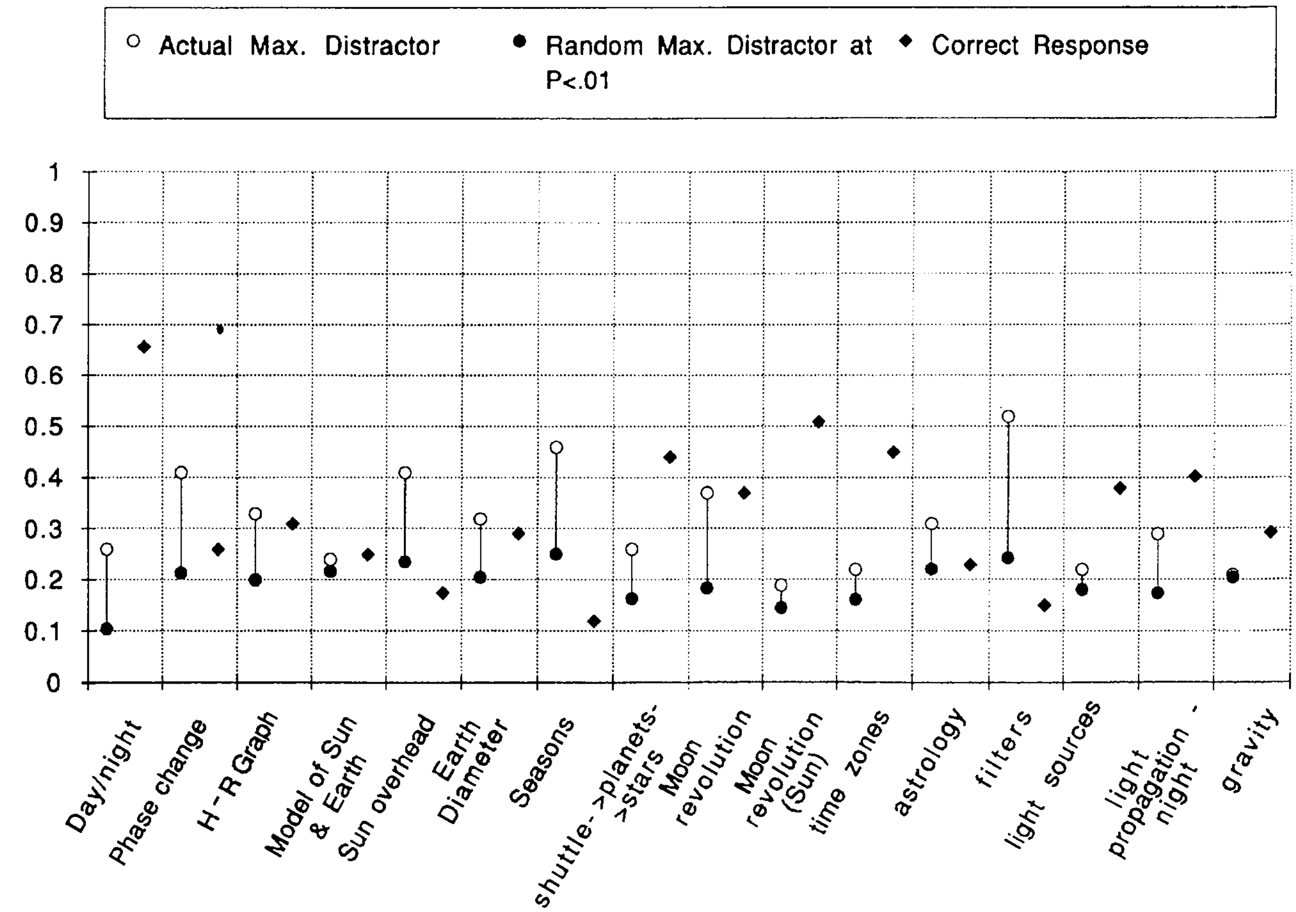


Fig. 1. A distractor analysis of the 16 test questions. All analysis is done for students' test scores before a course in astronomy. The filled diamonds give the fraction of students choosing the correct answer to the question. The open circles give the fraction of students choosing the most popular wrong answer (maximum distractor). The filled circles give that fraction of students choosing the maximum distractor above which the probability would be less than one percent if such a choice were made at random among the four wrong answers.

Results and Discussion

Figure 2 gives the results of both the teacher predictions and the actuality of student performance before and after a traditional course in astronomy. The last column in the figure gives the actual and predicted performances averaged over all 16 questions. Figure 3 gives the 95% confidence intervals for the student performances and teacher predictions averaged over the 16 questions, taking into account finite sampling errors.

Three results can be immediately seen from the last column in Fig. 2 or in Fig. 3: (1) A course in astronomy did not result in an overall gain of astronomical knowledge as measured by the test. (2) On average, the teachers did a good job of predicting their students' initial scientific knowledge; in other words, the percentage of students expected to get a correct answer. (3) The teachers vastly overestimated the gain in knowledge their students would achieve after their astronomy course (on average, a predicted gain of about double, from 36 to 73%).

The teacher predictions did not vary greatly with the ages of the students. Teachers of younger students predicted a slightly lower *initial* state of knowledge than did the teachers of older students, but even these small differences vanished completely for the predicted state of knowledge after the course. In particular, teachers of eighth-graders predicted that, on the average question, 33% of their students would get the correct answer on the first day of class; teachers of twelfth-grade students predicted 42%. (Teachers of students intermediate between these two extremes predicted intermediate fractions.) Teachers of all students, from eighth grade to twelfth grade, predicted that 73% of their students would get the correct answer to the average question on the last day of class. The consensus of the predicted post-course knowledge state, independent of the ages of students, is remarkable.

Student performance, either pre-course or post-course, did not vary much with grade level, ranging from 29% for eighth-graders to 40% for twelfth-graders. The difference between ninth-grade performance (33%) and tenth-grade (35%) was small.

If we look more closely at the actual student performances, we find that students achieved a definite positive gain in the factual questions (1,3,4,6,9) after taking their astronomy courses but a *negative* gain in the conceptual questions (2,5,7,8,10,11,13,14,15,16). More quantitatively, if we divide the total number of correct responses (summed over all questions in a certain category) after the course by the corresponding total before the course, then this ratio is 1.26 for the factual questions and 0.91 for the conceptual questions, leading to a fractional gain of 0.26 for the factual questions and of -0.09 for the conceptual questions. The consequence for teacher predictions was that although the teachers overestimated students' gains in both the factual and conceptual categories, their overestimates were much farther off the mark for the conceptual questions.

We emphasize that our 16-question test is only a small subset of the knowledge that astronomy teachers teach. No finite test could adequately reflect the material in a semester course. However, the precise coverage of the test is irrelevant when comparing teachers' predictions and students' performances. The teachers evidently believe that much of the material on the test is covered in their courses and covered well, because they predict significant improvements on the test after their course is over. As seen by the results, these predicted improvements are much overestimated. In particular, teachers are farthest off the mark for conceptual knowledge, for which students often have strong, underlying misconceptions when they enter the classroom.

We suggest that science teachers might be more effective if they understood the obstacles to conceptual learning (particularly the strong grip of prior misconceptions and the resistance to conventional instruction) and if they became familiar with the educational research and strategies dealing with those misconceptions. Other researchers in science education have also recommended that teachers interview their students about their misconceptions at an early stage of instruction.[22-24]

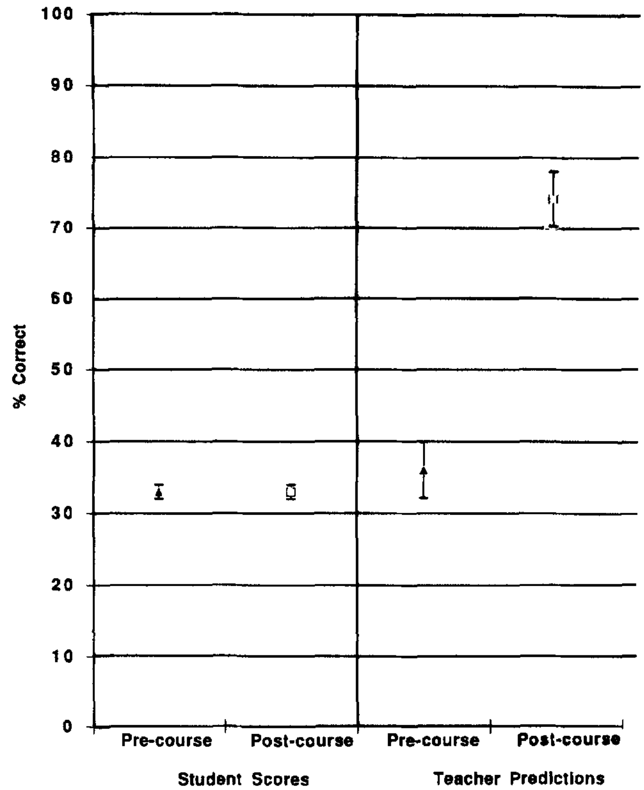


Fig. 3. Ninety-five percent confidence intervals for average student performances and teacher predictions.

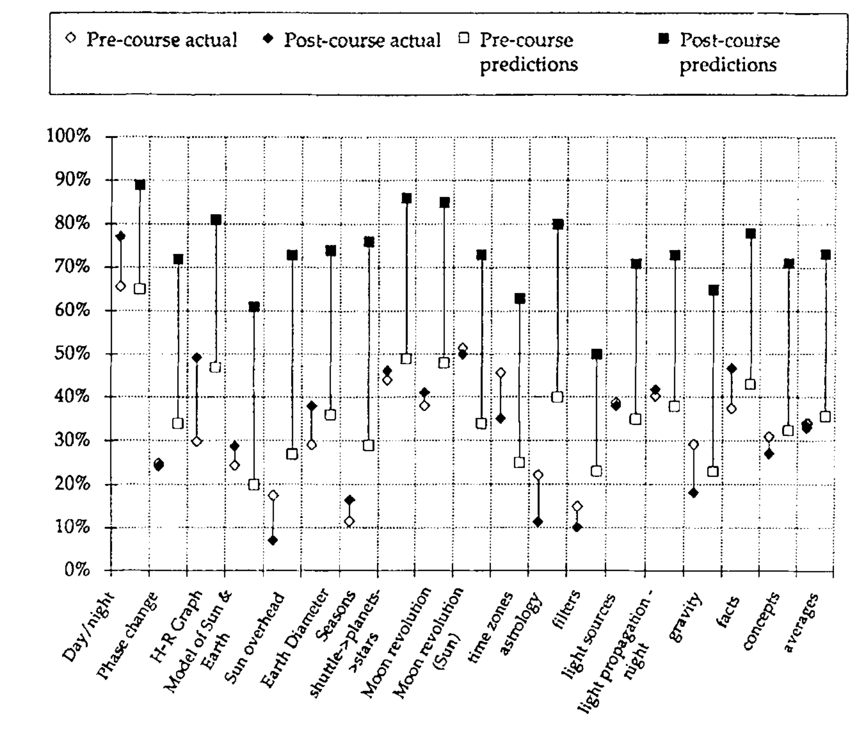


Fig. 2. Comparison of student test scores with teacher predictions. The vertical axis gives the fraction of students getting the correct answer or predicted to get the correct answer on each of the questions indicated on the horizontal axis. Open and filled diamonds refer to actual student performance before and after a course in astronomy, respectively. Open and filled squares refer to teacher predictions of student performance before and after their course in astronomy, respectively.

For the last eight years we have been part of a team of scientists and educators who have developed a new high-school science curriculum, called Project STAR, which has been built on research in misconception theory. Project STAR has now been tested on science teachers and their students across the country, and we will report on the results in a later paper.

Acknowledgements

This work has been supported by NSF MDR 88-50424 and by the Smithsonian Institution.

Appendix I and References follow on pages 166–167.

Appendix I

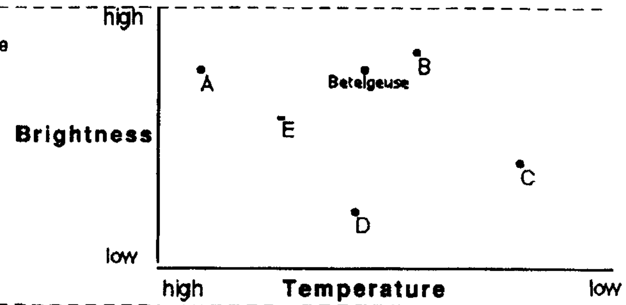
1. What causes night and day?
   A. The Earth spins on its axis.
   B. The Earth moves around the Sun.
   C. Clouds block out the Sun's light.
   D. The Earth moves into and out of the Sun's shadow.
   E. The Sun goes around the Earth.

One night you looked at the Moon and saw this:

A few days later you looked again and saw this:

2. Why did the Moon change shape?
   A. Something passed in front of it.
   B. It moved out of the Earth's shadow.
   C. It moved out of the Sun's shadow.
   D. Its far side is always dark.
   E. None of the above.

3. Which star on the graph has a temperature most like that of Betelgeuse?
   A
   B
   C
   D
   E



4. If you used a basketball to represent the Sun, about how far away would you put a scale model of the Earth?
   A. 1 foot or less   B. 5 feet   C. 10 feet   D. 25 feet   E. 100 feet

5. How often is the Sun directly overhead at noon in your hometown?
   A. Every day.
   B. Only in the summer.
   C. Only for the week of the summer solstice.
   D. Only for one day each year.
   E. Never.

6. Give the best estimate of the Earth's diameter from among the numbers in the column on the right.
   A. 1,000 miles.
   B. 10,000 miles.
   C. 100,000 miles.
   D. 1,000,000 miles.
   E. 10,000,000 miles.

7. The main reason for it being hotter in summer than in winter is:
   A. the Earth's distance from the Sun changes.
   B. the Sun is higher in the sky.
   C. the distance between the northern hemisphere and the Sun changes.
   D. ocean currents carry warm water north.
   E. an increase in "greenhouse" gases.

8. Which answer shows a pattern from closest object to the Earth to farthest from the Earth?
   A. Space Shuttle in orbit --> Stars --> Pluto
   B. Pluto --> Space Shuttle in orbit -->Stars
   C. Stars --> Space Shuttle in orbit --> Pluto
   D. Stars --> Pluto --> Space Shuttle in orbit
   E. Space Shuttle in orbit --> Pluto --> Stars

Choose your best estimates of the times for the events listed. Choices may be used more than once.

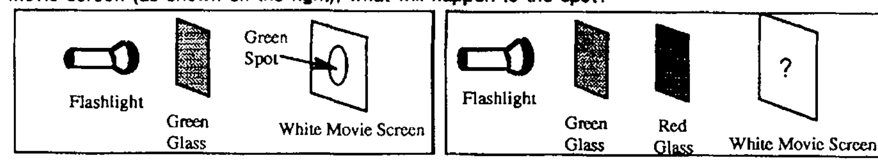
____9. The Moon to go around the Earth.   A. Hour B.Day C.Week D.Month E.Year

____10. The Moon to go around the Sun.   A. Hour B.Day C.Week D.Month E.Year

11. Boston is 90° east of Hawaii. If it is noon in Hawaii, in Boston it would be about:
   A. Sunrise.   B. Sunset.   C. Noon.   D. Midnight.   E. Noon the next day.

12. Most astronomers consider *astrology* to be:
   A. a science.
   B. a good way to determine personality traits.
   C. helpful in predicting world events.
   D. more than one of the above.
   E. none of the above.

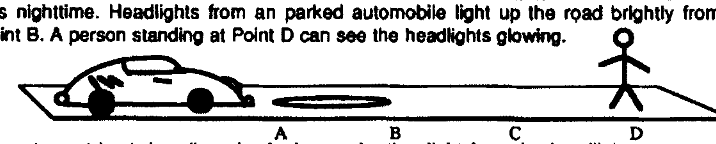
13. When green glass is placed between the flashlight and the white movie screen, a green spot appears on the screen. If green glass and red glass are placed between the flashlight and the movie screen (as shown on the right), what will happen to the spot?



A) It will be green.
B) It will be yellow.
C) It will be brown.
D) It will be red.
E) It will disappear.

14. You are in a completely dark room. There are no lights and no windows. Which group of objects do you believe you might be able to see?
A) bicycle reflectors, a cat's eyes
B) silver coins, aluminum foil
C) white paper, white socks
D) more than one of these groups
E) none of these

15. It is nighttime. Headlights from an parked automobile light up the road brightly from point A to point B. A person standing at Point D can see the headlights glowing.



Which statement best describes the farthest point that light from the headlights can reach?
A) Light does not leave the headlights.
B) The light reaches only as far as point A.
C) The light reaches only as far as point B.
D) The light reaches only as far as point C.
E) The light reaches at least as far as point D.

16. Which of the following would make you weigh *half* as much as you do right now?
A) Take away half of the Earth's atmosphere.
B) Double the distance between the Earth and Sun.
C) Decrease the Earth's rate of spin so that 1 day equals 48 hours instead of 24 hours.
D) More than one of the above.
E) None of these.

References

1. J.D. Novak, *A Theory of Education* (Cornell University Press, Ithaca, 1977).
2. D. Ausubel, *The Psychology of Meaningful Learning* (Holt, Rinehart, and Winston, New York, 1963); *Educational Psychology: A Cognitive View* (Holt, Rinehart, and Winston, New York, 1968).
3. J.L. Keuthe, "Science concepts: A study of sophisticated errors," *Sci. Educ.* **47**, 361 (1963).
4. J. Nussbaum and J.D. Novak, "An assessment of children's concepts of the earth utilizing structured interviews," *Sci. Educ.* **60**, 535 (1976).
5. L.C. McDermott, "Research on conceptual learning in mechanics," *Phys. Today*, 24 (July 1984).
6. I.A. Halloun and D. Hestenes, "The initial knowledge state of college physics students," *Am. J. Phys.* **53**, 1043 (1985).
7. *Misconceptions and Educational Strategies in Science and Mathematics*, edited by J.D. Novak (Cornell University Press, Ithaca, 1987).
8. L. Flick, "Preservice teachers conduct structured interviews with children to improve instructional methods," in Ref. 7, Vol. 1, 133 (1987).
9. A. Lightman and P. Sadler, "The earth is round? Who are you kidding?" *Sci. Child.* **25** (5), 24 (1988).
10. "A Private Universe," produced by Matthew Schnepps, Pyramid Films (1986).
11. T.L. Good, "Two decades of research on teacher expectations: Findings and future directions," *J. Teach. Educ.* **38**, 32 (July-August 1987).
12. C.K. Bruce and F. Lawrenz, "Actual and teacher predictions of the abilities of mathematical high school chemistry students in Minnesota," *School Sci. Math.* **91** (1), 1 (1991).
13. B. Anderson and C. Karrqvist, "How Swedish pupils, aged 12-15 years, understand light and its properties," *Eur. J. Sci. Educ.* **5** (4), 378 (1983).
14. C.A. Camp, *Problem solving patterns in science: Gender and spatial ability during early adolescence*, Ed.D. thesis, University of Massachusetts, Amherst (1981).
15. M.R. Cohen and M.H. Kagan, "Where does the old Moon go?" *Sci. Teach.* **46**, 22 (1979).
16. E. Guesne, "Light," in *Children's Ideas in Science*, edited by R. Driver, E. Guesne, and A. Tiberhien (Open University Press, Philadelphia, 1985), pp. 10-32.
17. R.F. Gunstone and R.T. White, "Understanding gravity," *Sci. Educ.* **65** (3), 291 (1981).
18. C.A. Klein, "Children's concepts of the Earth and the Sun: A cross cultural study," *Sci. Educ.* **65** (1), 95 (1982).
19. M.M. Rollins, et al., "Attainment of selected earth science concepts by Texas high school seniors," *J. Educ. Res.* **77**, 81 (1983).
20. B.F. Stead and R.J. Osborne, "What is gravity? Some children's ideas," *NZ Sci. Teach.* **30**, 5 (1981).
21. D.F. Treagust, D.F. Smith, and C.L. Smith, "Secondary students understanding of the solar system: Implications for curriculum revisions," in *GIREP—Cosmos and Educational Challenge in Copenhagen* (European Space Agency, 1986), pp. 363-369.
22. E. Duckworth, *The Having of Wonderful Ideas and Other Essays on Teaching and Learning* (Teachers College Press, New York, 1987).
23. P. Langford, *Children's Thinking and Learning in Elementary School* (Technomic, Lancaster, 1989).
24. S. Novick and J. Nussbaum, "Using interviews to probe understanding," *Sci. Teach.* **45** (8), 29 (November 1978).

Physics Trick of the Month

The Gorilla Effect

Before showing this trick, secretly rub your wet *index* fingertip on a bar of soap.

Fill a shallow dish, a saucer will do, with water and scatter black pepper over the surface. The water represents a lake. The pepper grains are bathers. Your finger, you explain, models another bather entering the lake. So saying, touch the tip of your *middle* finger to the water at the saucer's rim. Nothing happens to the "bathers." Repeat this a few times to represent other bathers entering the water.

Now, you continue, along comes a gorilla who has escaped from a nearby zoo. Again, your finger models the gorilla as he enters the lake. This time, however, place the tip of your soaped *index* finger on the "lake's" edge. Instantly all the "bathers" flee to the opposite side!

Martin Gardner, *Hendersonville, NC 28739*