Third Misconceptions Seminar Proceedings (1993)

Paper Title: Teachers' Misconceptions of their Students' Learning
Author: Sadler, Philip M.

Abstract: As startling as many students' misconceptions are, teachers' lack of awareness of their students' ideas are equally confounding. This study has sought to quantify 132 teachers' predictions of students' ideas both before and after a high school science course. Using a subset of Project STAR's 47-item astronomical misconception test, the pre- and post-test responses of 330 secondary school earth science and astronomy students from around the U.S. were compared to teachers' predictions. Teachers were found to have reasonable expectations of students' initial knowledge state, but vastly overestimated student gains as a result of instruction. In fact, students showed no average gain at all on this instrument. This study concludes that traditional courses appear to have little effect on students' understanding of astronomical concepts, yet most teachers believe that their effect on students is large. The questionable validity of teacher-constructed tests for measuring conceptual change and teachers' lack of belief in the enduring nature of misconceptions are discussed as possible explanations for teachers' own misconceptions about how much students learn in science courses.

Keywords: teacher education,testing,concept formation,misconceptions,teachers,teaching for conceptual change,test validity,,
General School Subject:
Specific School Subject:
Students:

Macintosh File Name: Sadler - Teachers
Release Date: 6-13-1994 C, 11-9-1994 I

Publisher: Misconceptions Trust
Publisher Location: Ithaca, NY
Volume Name: The Proceedings of the Third International Seminar on Misconceptions and Educational Strategies in Science and Mathematics
Publication Year: 1993
Conference Date: August 1-4, 1993
Contact Information (correct as of 12-23-2010):
Web: www.mlrg.org
Email: info@mlrg.org

A Correct Reference Format: Author, Paper Title in The Proceedings of the Third International Seminar on Misconceptions and Educational Strategies in Science and Mathematics, Misconceptions Trust: Ithaca, NY (1993).

Teachers' Misconceptions of their Students' Learning

Philip M. Sadler

Harvard-Smithsonian Center for Astrophysics and Harvard Graduate School of Education, USA

As startling as many students' misconceptions are, teachers' lack of awareness of their students' ideas are equally confounding. This study has sought to quantify 132 teachers' predictions of students' ideas both before and after either a secondary school earth science or astronomy course. Using a subset of Project STAR's 47-item astronomical misconception test (Sadler 1992), the pre- and post-test responses of 330 secondary school earth science and astronomy students from around the U.S. were compared to teachers' predictions. Teachers were found to have very reasonable expectations of students' initial knowledge state, but vastly overestimated student gains as a result of instruction. In fact, students showed no average gain at all on this instrument. This study concludes that traditional courses appear to have little effect on students' understanding of astronomical concepts, yet most teachers believe that their effect on students is large. The questionable validity of teacher-constructed tests for measuring conceptual change and teachers' lack of belief in the enduring nature of misconceptions are discussed as possible explanations for teachers' own misconceptions about how much students learn in science courses.

Since my initial involvement filming students ideas in science (Sadler, 1988), I became intrigued by teachers' lack of awareness of the ideas of their own students. How was it possible for teachers (and I must include myself as well) to be totally oblivious of their students misconceptions until the seminal studies in the late 1970's? In spite of the large number of studies that ensued, most classroom teachers still give scant attention to their students ideas. Many teachers react with surprise when they discover that their students have basic misconceptions about concepts they have taught.(Targan, 1987) Most have not changed their teaching strategies to facilitate conceptual change.

Although many teachers and administrators complain of too much testing going in the schools, their reference is typically to standardized testing, not to the tests and quizzes generated by the teachers themselves. The instruments that teachers create themselves are not subjected to reliability, validity, or readability measures. They are only rarely given to other teachers' students. Yet, they are used as a major factor in determining students' grades. Although many of these instruments may be good for measuring student performance, the grades that students get on these tests does not appear to reflect the quality of teaching going on in the classroom. The very worst teachers in a school and the very best teachers in a school cannot be distinguished based upon how students do on the tests that they themselves generate. To what degree do we, as teachers, fool ourselves into believing that our students learn as much as we think, by making up our own assessment instruments? Can instruments that have never been subjected to careful scrutiny have any chance of measuring conceptual change?

My colleague and I have attempted to use methods that have undergone many years of development and use them to examine the astronomical conceptions of students and how well their teachers understand the level of students' understanding.

Earth Notions

A good example of such a possible mismatch is that of teachers' beliefs of their students' notions of cosmography. Studies of students' notions of the earth in space are still unrecognized by most classroom teachers. For example, many 13- and 14-year-old students still believe that the Earth is flat or that we all live inside a hollow Earth.(Klein, 1982; Lightman & Sadler, 1986; Mali & Howe, 1979; Nussbaum, 1979; Nussbaum, 1985; Nussbaum, 1986; Nussbaum & Novak, 1976; Sneider & Pulos, 1983) Of the hundreds of high school teachers I have worked with, only one has a admitted that one of her students has held such a misconception. These notions confounded virtually all of the middle school science teachers and elementary school teachers with whom I have talked, as well. To what degree do students have these beliefs without their teachers being aware of them?

I decided to pursue my own line of inquiry into this matter by asking teachers to predict the Earth notions of their students. For three summers (1988, 1989, 1990), as a part of my presentation of astronomy activities to teachers at the National Science Resources Center Institute and the Independent School Association of Massachusetts, 111 teachers of grades K-8 predicted the distribution of Earth notions among students in their classes (Lightman and Sadler 1986). The level of Earth notion was based on Nussbaum's scale. A simplified explanation of this scale is:

Level 1: The earth is flat with the sky and space above.

Level 2: The earth is a sphere. We live inside with the sky above us.

Level 3: The earth is a sphere. We live only on the top of the sphere.

Level 4: We live all over the earth. Objects inside would fall southward.

Level 5: The earth is spherical with objects inside falling toward its center.

For each grade level, an average score was calculated for both the teachers' predictions and the actual students' performance. Sneider and Pulos carried out testing in grades three to eight for a different, but typical school population. Teacher predictions were made for grades kindergarten to eight. The results were quite surprising (see Figure 1).

If the data collected by Sneider and Pulos are representative of the students in the sample teachers' classrooms, then teachers are woefully unaware of their students' conceptions of the world. Student conceptions in grade 3 (and most likely in K, 1, and 2), are of a flat or hollow Earth. Yet, teachers believe that students' notions are much more sophisticated and closer to their own. Three aspects of teachers' predictions of their students' development level are apparent in Figure 1. First, teachers at every level tested overestimated the sophistication of students notion of the Earth in space. Second, the widest disparity is at the lowest grade level. Middle elementary teachers estimate that their students are operating at levels three to four years more advanced than where they actually are. Lastly, even in the eighth grade, when teachers think that virtually all students are at the highest level of conceptualization, many are still thinking that we live inside the earth or that people who venture to lower latitudes will fall off.

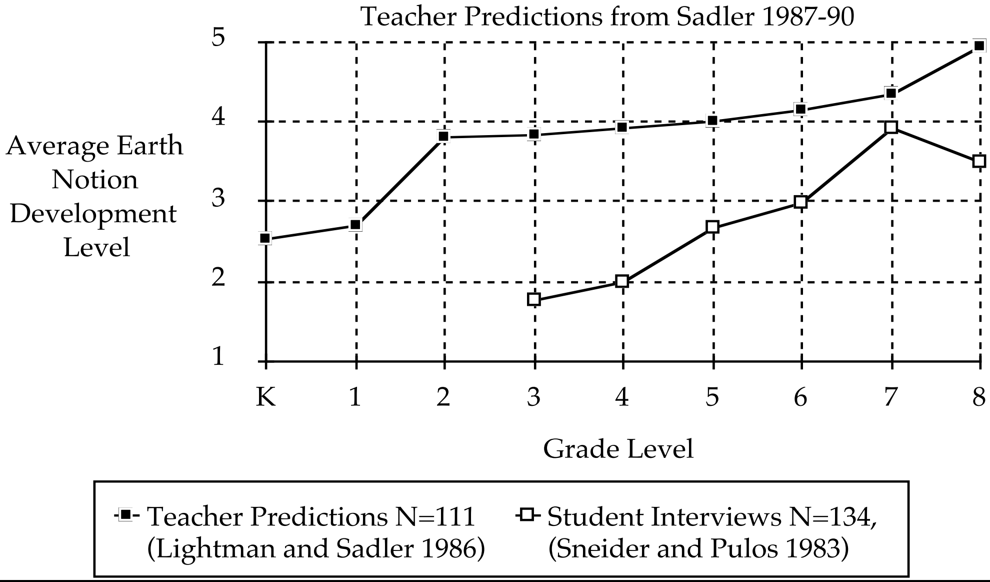


Figure 1, Earth Notions Development Level
Student Interviews from Sneider and Pulos 1983
Teacher Predictions from Sadler 1987-90

Undoubtedly, instruction about space or world geography might be very ineffective at these grades because of the difference between the conceptions students hold and the concepts their teachers think they hold. Many students at the lower primary level even think the globe in the classroom is a model of some other planet in outer space, not the one on which we live.(Vosniadou & Brewer, 1987) How can they make sense of geography or science instruction whenever a globe is used? I concluded that teachers' awareness of their own students' misconceptions might be a fruitful subject for further study. However, what would account for the mismatch? Is it that teachers misjudge their students' initial level of understanding or do they over estimate the effect of their own instruction?

Study Design and Administration

My colleague, Alan Lightman, Professor of Science and Writing at MIT, and I designed a more formal study to investigate how aware teachers are of students misconception both before and after having them in class. We also wished to determine the extent to which student gains related to their teachers' perceptions.

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. (Anderson 1983; Camp 1981; Cohen 1979; Guesne 1985; Gunstone 1981; Klein 1982; Rollins 1981; Stead 1981; Treagust 1986) The test instrument is shown at the end of this article.

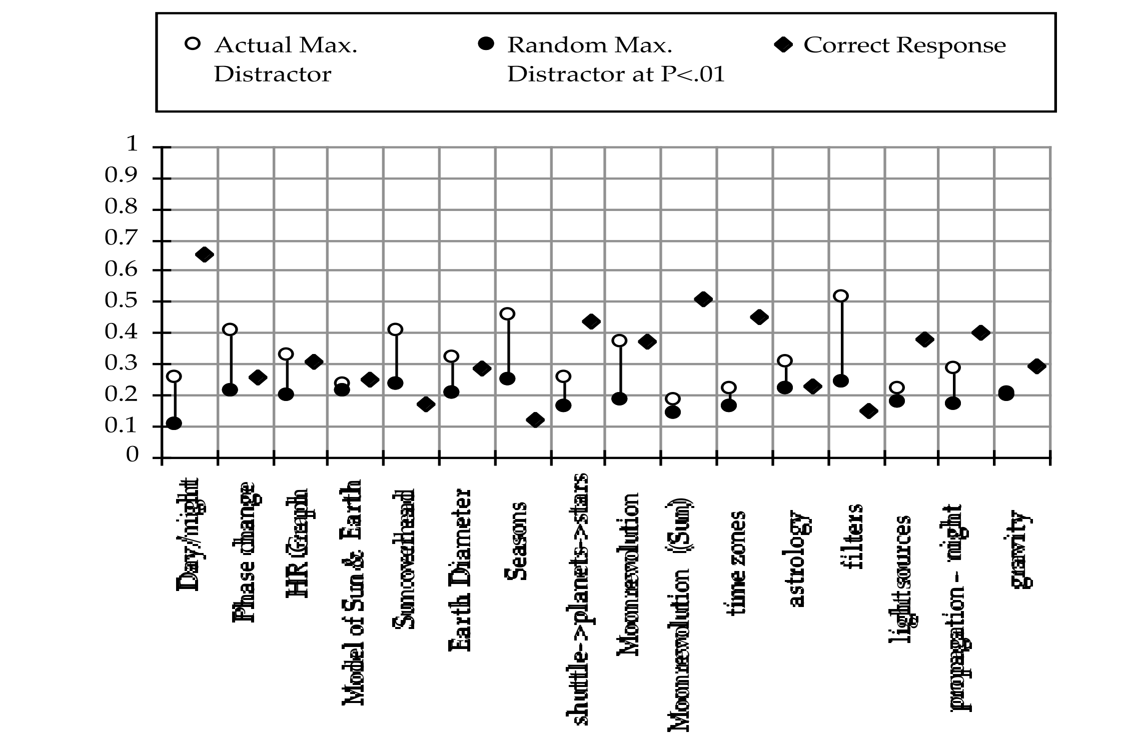
For comparison with the conceptual questions, we also included some factual and quantitative questions. In particular, questions 1, 3, 4, 6, and 9 involve factual or quantitative knowledge, whereas the remaining questions require conceptual reasoning. Question 12, on astrology, doesn't easily fit into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

The reliability of this instrument was calculated using the Kuder-Richardson 20 formula which gives a value of 0.94. Standardized tests, such as those used for college admissions, have reliability coefficients of 0.90 or greater (Hopkins and Stanley 1981). The validity of the test instrument was established in two different ways to determine if the test is a good measure of conceptions in astronomy. Two groups of experts, graduate students in astronomy and astronomy professors took versions of the test to see if they chose the correct answers. Only items on which experts agreed were included. Some questions had to be modified so that the experts would agree on the same answer. Various item characteristics were explored to seek plausible explanations of why students chose not to answer certain questions including:

P-value: the difficulty of the question.
Item #: from the order in which the question appeared.
Picture: whether the problem had an accompanying graphic.
Concept: whether the problem dealt with an astronomical concept.
Fact: whether the problem dealt with an astronomical fact.
Math: whether the problem required the exercise of a math skill.
Readability: Gunning Fog Index.
None of these characteristics were significant at the P=.05 level.

The authors performed a "distractor" analysis to confirm the belief that many of the questions and responses indeed contain prevalent misconceptions. Our distractors are incorrect choices that were drawn from the misconception literature or from our own interviews with children. 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.

Figure 2, Distractor Analysis



The filled circles in Fig. 2 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. All the remaining questions involve misconceptions according to this measure. The test was intentionally designed for this outcome.

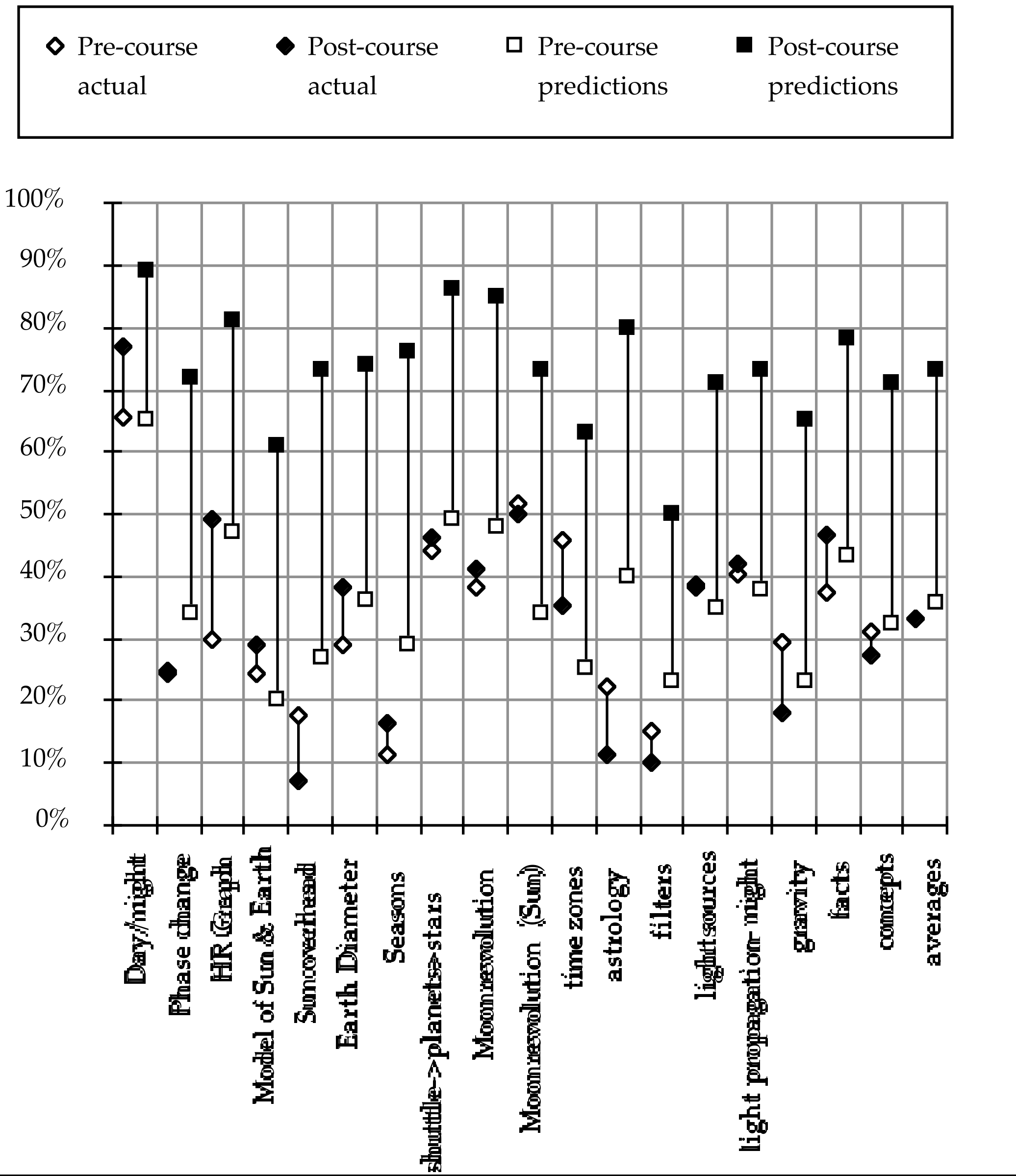
In the spring of 1991, we sent half the test (8 questions) to 600 astronomy teachers and the other half to another 600. These teachers were drawn from the STARNews mailing list, compiled from national workshops on astronomy teaching or those who wrote to us requesting information on our work in astronomy education. 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. The majority of the respondents were located in New York, Minnesota, Illinois, California, and Ohio. Most 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 to the teacher predictions, we administered a 47-question test (of which the 16 questions used in this test were included) 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 or earth science 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 1,414 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 pre-course and 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.

Results and Discussions

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 8 also shows the 95% confidence intervals for the student performances and teacher predictions averaged over the 16 questions, taking into account finite sampling errors.

Figure 3, Comparison of Actual and Predicted Scores

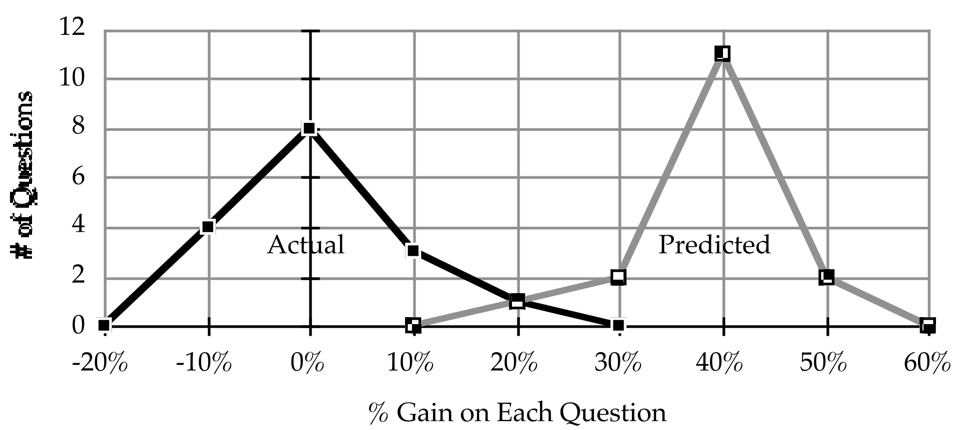


Three results can be immediately seen from the last column in Fig. 2 ; (1) A course in astronomy did not result in an overall gain of astronomical knowledge as measured by the test. (2) On average, the teachers did a good job of predicting their students' initial scientific knowledge; in other words, the percentage of students expected to get a correct answer at the start of the course. (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%).

If we look more closely at the actual student performances, we find that students achieved a definite positive gain in the factual question (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.

A careful examination of individual questions reveals that teachers overestimated student gains on every question. On many items students actually performed worse after their earth science or astronomy course than before. For every example of an increase in score such as the reason for day and night, there is a decrease in score on some other item, such as the use of time zones. Teachers predicted a gain of 40%, while the actual gain was 0%.

Figure 4, Histogram of Actual and Predicted Gains



Item Analysis

A critical analysis of items attempts to reduce the sources of errors in measurement and to gauge the quality of items. One method that has proven fruitful for the task is Classical Test Theory (CTT) . Classical Test Theory is based on the idea that a student's true ability is best measured by her or his *true score* and the *difficulty of a test item* (the fraction of subjects who answer an item correctly) (Hambleton et al. 1991). I calculate the P-value (fraction choosing the correct answer) for each of the test items. I assume that students who have a greater mastery of the material overall will, on average, score higher on individual items than students whose mastery is lesser. Individual items are then characterized by how well each discriminates between high-scoring and low-scoring students (Miller and Erickson 1990). The most popular method for doing this involves separating subjects into several subpopulations (I have chosen five) based on their total score on a test. A graph is then constructed that plots this overall ability versus the average P-value for that group. Good test items show P-values rising monotonically and steeply with ability. Poor items do not fit this profile. The discriminating ability of an item can be estimated as the slope of the curve, but a D-value can also be calculated to determine discriminating ability (Osterlind 1989). This is a measure of the correlation between responses to an individual test item and the total test score. It can be calculated for each individual response but is usually calculated only for the correct answer. I generated graphs and D-values for each item.

Classical Test Theory can be useful for evaluating test items. There is some evidence that it can characterize items dealing with misconceptions as lowering the reliability of the test and hence to be discarded (Narode 1987). Items with low D-values and P-values are often removed from standardized tests. These items are judged as poor because they are difficult and present distractors that are attractive to many high-ability students. I will give examples of a few different types of test questions using this technique. The correct answers are underlined for each item shown.

<u>Item 1, Reason for Day and Night</u>

What causes night and day?

<u>A. The Earth spins on its axis.</u>
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.

The reason for day and night is perhaps the most basic idea assumed by teachers of astronomy of their introductory students. In the prediction survey, participating teachers assumed that 65 percent of their students, on average, would enter their classes with this concept understood and, by the end of the course, 89 percent would leave knowing it. Only half that gain was realized. The Earth turning on its axis causing day and night has been described as one of the "most essential ideas which form the Earth conception" (Nussbaum 1985).

In a study of elementary school students, the reasons stated for day and night included that the Earth revolves around the Sun (B), that the Earth or Sun moves into a shadow (D), or that clouds block out the Sun's light (C) (Vosniadou and Brewer 1987). Another study of second-grade students found that many knew that the Sun was "on the other side of the Earth" at night, but showed no clear preference for whether it was the Earth or the Sun that moves (Klein 1982).

A item similar to this one was included in the 1969 National Assessment of Educational Progress of nine-year-old (third-grade) students. The percentage choosing each answer follows in parentheses:

One reason that there is day and night on Earth is that the

Sun turns. (8%) Moon turns. (4%) Earth turns. (81%)

Sun gets dark at night. (6%) I don't know. (1%)

This is a fine example of a question that is not designed to identify misconceptions. The high percentage of correct answers can be attributed to the lack of plausible distractors (Schoon 1988). Had the question included "the Earth goes around the Sun" and "the Sun goes around the Earth," the students' choices might have been quite different.

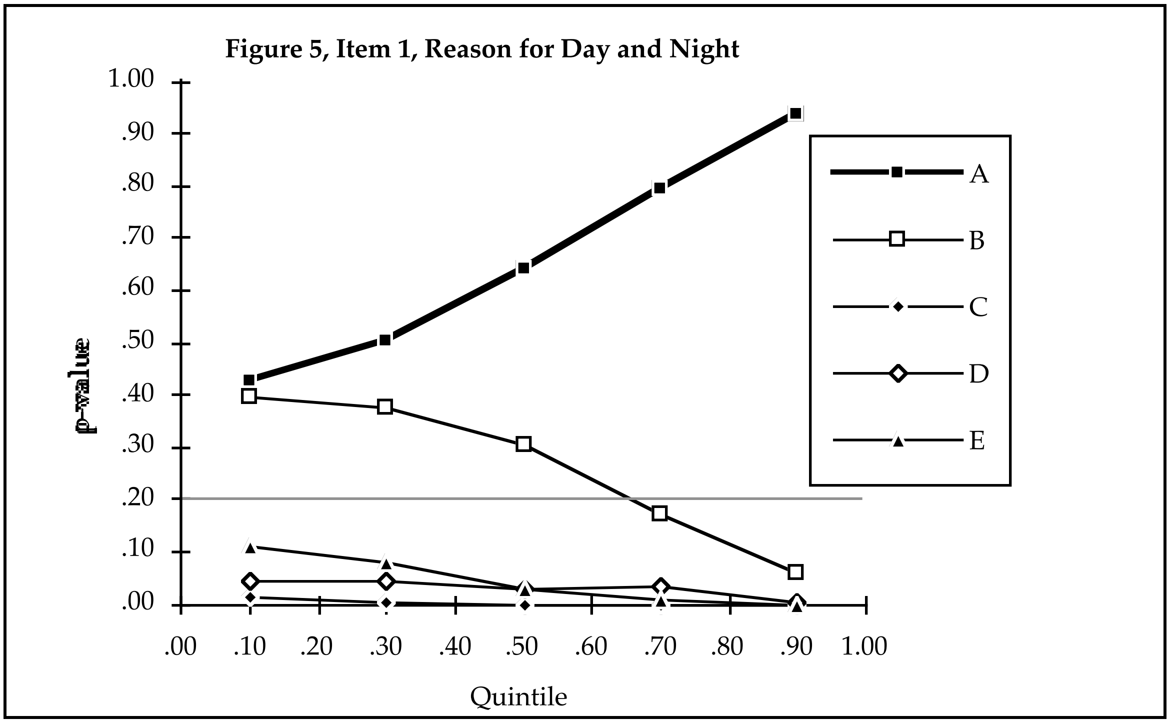
One researcher found that although college students prefer a heliocentric explanation of the solar system and reject a geocentric model, the majority could not give convincing arguments for their view when answering an exam question on the subject after an introductory astronomy course (Touger 1985). Justifications for heliocentrism took many surprising forms; two examples: "The Sun is the center... by observation we can see that the planets move around the Sun" (p. T-5) or "all our pictures and telescopes and space flights tell us that there is one big star, the Sun, with several smaller planets (moving) around it" (p. T-8). Touger argues that students' belief in heliocentrism is derived almost exclusively from secondary sources and lacks an empirical base. Many students believe that scientists have actually viewed the entire solar system from a vantage point in space.

I would argue that this acceptance of heliocentrism as dogma, without an ability to muster a shred of supporting evidence, makes this concept an attractive and almost universal answer to astronomical questions. Much as one researcher found when interviewing young children that God was invoked frequently to explain certain natural events (Za'rour 1976), the hard-learned belief in heliocentrism is called upon as justification for any astronomical problem for which the individual cannot give evidence supporting her or his view.

The follow table lists the P-values and D-values for each answer to Item 1. For all the following items a similar table is presented.

Item 1	A	B	C	D	E
P-value	.66	.26	.00	.03	.04
D-value	.39	-.29	-.06	-.09	-.19

Students do very well on this question, with 62 percent of them selecting the correct answer on the pre-test—that the reason for day and night is that the Earth spins on its axis. From this table, one can clearly see that answer B, day and night are caused by the Earth moving around the Sun, is preferred by .26 of the students in the survey. The following graph plots the P-values of each answer to Item 1 for each performance quintile of students.



From this graph, it appears that answer "B," the Earth circling the Sun causes day and night, is a major misconception for all but the best performing students in the test population. Note that the answer, "the Sun goes around the Earth," appears unattractive to students. It does not seem that they are confusing the Earth's rotation with their observations of the Sun apparently circling the Earth. They actually think that our orbiting the Sun causes day and night. Instruction does appear to help students with this misconception. Eleven percent of students changed their answer from a misconception to the correct answer in this study.

Item 5, When is the Sun Overhead?

How often is the Sun directly overhead at noon in your hometown?

A. Every day. D. Only for one day each year.

B. Only in the summer. E. Never.

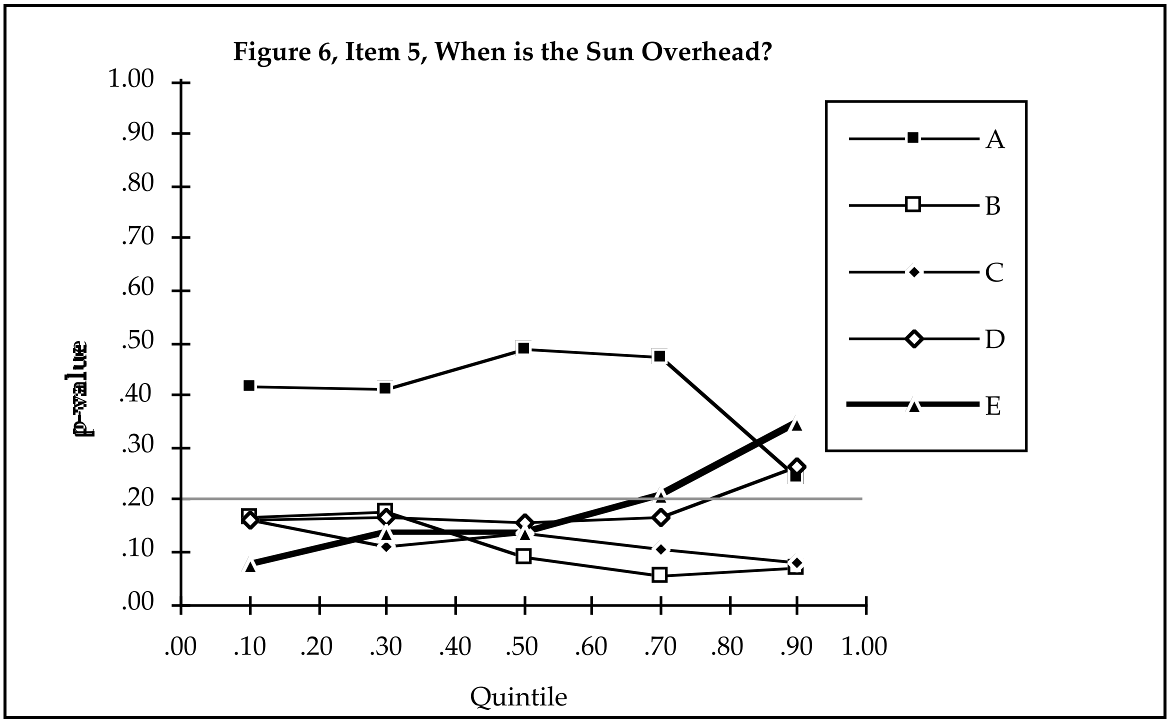
C. Only for the week of the summer solstice.

This test was given to students only in the continental United States. The Sun can only be seen directly overhead between the Tropics of Capricorn and Cancer (between 22.5°N and 22.5°S latitude). The Sun is never overhead in the continental United States. The correct answer is "E. Never." In Boston, the

Sun is only 25° above the horizon at noon on the winter solstice. On the first day of summer it is much higher, but still rises only to 71° altitude at its maximum.

Schoon found that 12 out of 13 participating teachers and 20 out of 32 student teachers believed that the Sun is always overhead at noon (Schoon 1988).

Item 5	A	B	C	D	E
P-value	.41	.11	.12	.18	.18
D-value	-.14	-.14	-.07	.09	.27

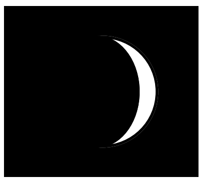


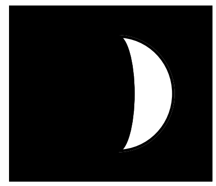
This appears to be a difficult question, with "A" being students' preferred answer. A plurality of students believe that the Sun is always directly overhead at noon. Teachers predicted that students would do poorly on this question, that only .27 would get this question right. Students do much worse than teachers predict. Not knowing that the Sun is lower in the sky in the winter precludes a proper understanding of the reason for seasons. The connection between the geocentric and heliocentric frames of reference is key to understanding this concept. Students do 11% worse on the post-test than on the pre-test. Perhaps with such a large number of students holding misconceptions, the effect on the class is regression rather than advancement.

The misconception that the Sun is overhead at noontime is the most common answer among all performance levels. These youngsters have not noticed how much longer their shadow is at noon in the winter than in the summer or how the Sun always seems to be in their eyes in the winter. Only students in the highest performance quintile show a substantial reduction in this misconception. However, many still cannot let go of the belief that the Sun is directly overhead at some time, choosing "D, only for one day each year," but with greater frequency than other students.

<u>Item 2, Reason for Moon Phases</u>

One night you looked A few days later you
at the Moon and saw this: looked again and saw this:





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.
<u>E. None of the above.</u>

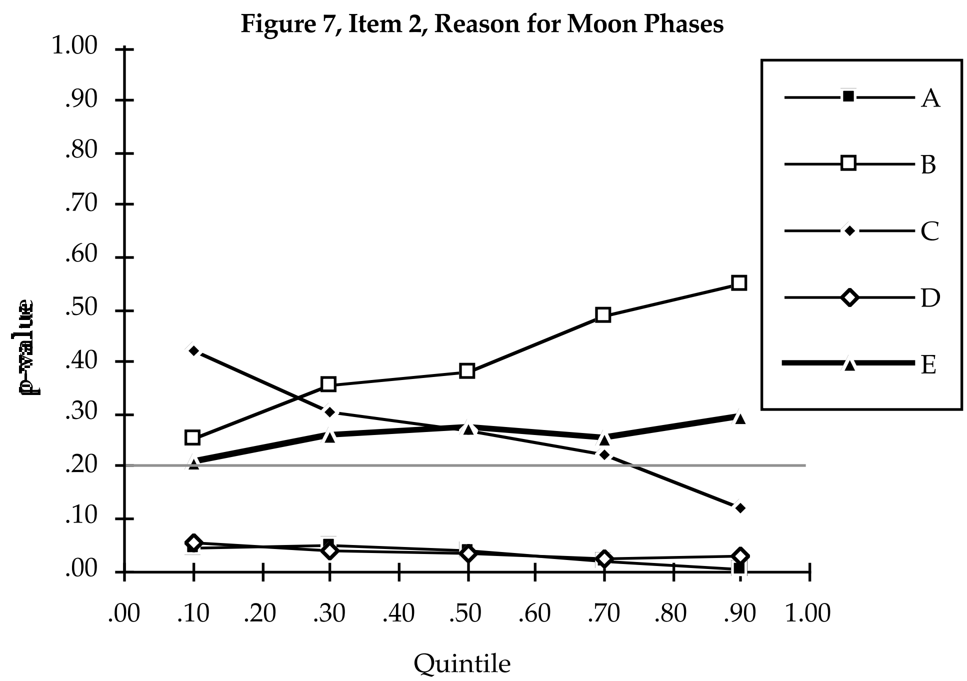
The Moon's phases are caused by the fact that our view of the lighted side of the Moon changes as the Moon orbits the Earth. The Moon has no light of its own and is illuminated by the Sun. This answer is not listed in the first four choices, so "E" is the correct answer.

Even teachers are confused by this concept. The Boston Curriculum Objectives (Marshall 1983) urge teachers to test their students on identifying Moon phases with a drawing of two "phases," a crescent Moon and a partial lunar eclipse. Clearly, whoever made up this guide would have answered Item 2 with "B" instead of the correct answer. In a study that interviewed fifty pre-service and in-service elementary school teachers, 74 percent of respondents were found to have incorrect concepts (Cohen 1982). In this study, eleven teachers thought that clouds, a planet, or a star blocked the Moon. Two thought that the Moon is black and white, and rotates, and twenty-four implicated the Earth or its shadow.

An early precursor to misconception studies examined the "sophisticated errors" of 100 recent high school graduates in 1963. Seventy percent believed that the Earth's shadow caused the phases of the Moon (Keuthe 1963).

Item 2	A	B	C	D	E
P-value	.03	.41	.27	.04	<u>.26</u>
D-value	-.09	.19	-.21	-.05	<u>.06</u>

This is a difficult question, especially because the correct reason for the phases is not listed in the answers, only "none of the above." However, teachers in our nationwide survey predicted that .34 of entering students would know the answer to this question and that the fraction who would learn it by the end of their course would rise to .72. Two distractors, "B" and "C," appear to be more popular than the correct answer. Students underwent virtually no change (-1%) during their courses.



This question reveals that most students have the wrong idea about the cause of the phases of the Moon. Higher-performing students appear much more likely than lower-performing students to think that the Moon's phases are caused by the Earth's shadow.

Item 7, The Reason for Seasons

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.

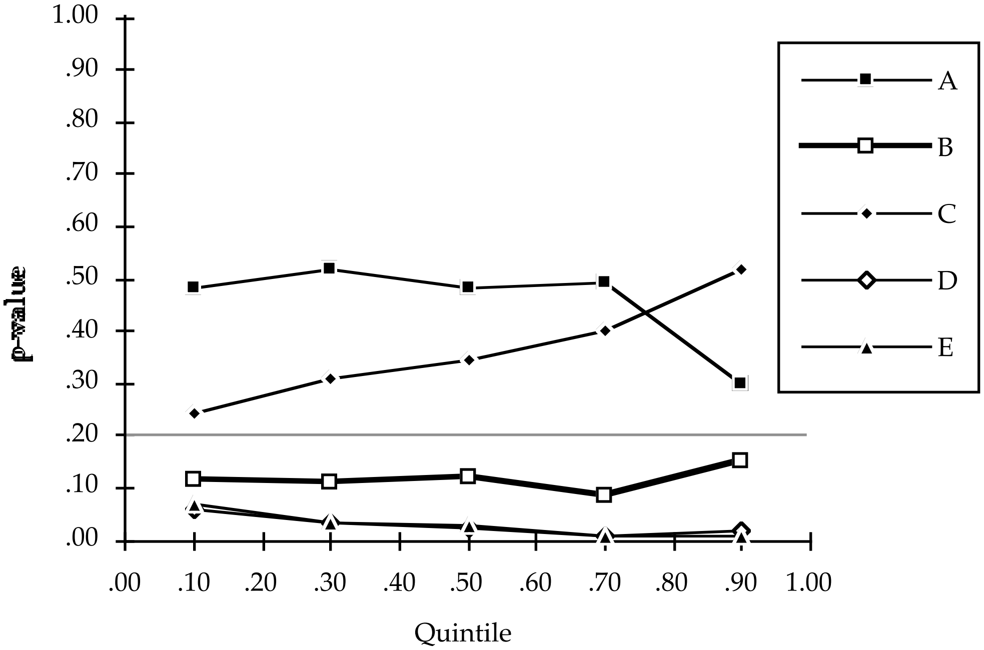
E. an increase in "greenhouse" gases.

The Sun is lower in the sky in the winter than in the summer. This change in altitude spreads the Sun's light over a much broader area on the Earth. The Boston Curriculum Objectives (Marshall and Lancaster 1983) for fifth grade explain correctly that the reason for winter is that "the Sun is lower in the sky," but then go on to qualify this reason with the incorrect statement, "its rays have to shine through more atmosphere before they reach us, losing heat energy in the process."

Item 7	A	B	C	D	E
P-value	.45	.12	.36	.03	.03
D-value	-.15	.04	.21	-.08	-.10

This is a question that is both extremely difficult and does not discriminate between students based upon overall performance. Teachers in the prediction survey thought students in introductory astronomy courses would score .29 before their courses and .76 after. The actual P-value is less than half of that score. Both answer "A" and answer "C" are far more popular than the scientifically correct response. The actual gain was only 6%.

Figure 8, Item 7, The Reason for Seasons



In answering this question, students appear to be torn between two distractors that mention changing distance. Many students believe that the Earth's orbit is highly eccentric so that the entire Earth is physically closer to the Sun in the summer than in the winter. A more "evolved" explanation is that the Earth leans toward the Sun in the summer and away from the Sun in the winter. This is consistent with many diagrams in textbooks that show the one pole proportionally much closer to the Sun in the summer. The correct answer, "B," appears to be avoided by most students at all performance levels. It is clear that students have not connected the Earth's tilt with the altitude of the Sun in the sky during different seasons.

This test was designed to uncover misconceptions. To that end, I have calculated the P-values of the eleven most popular distractors in Table I. Those that were chosen with a frequency greater than the correct answer for each problem are marked with an "X."

Table I. Ranking of Misconceptions by P-Value

Answer #	P-value	> Correct answer	Misconception exhibited
13C	.53	X	Colored filters mix like paints.
7A	.46	X	Changing distance is responsible for seasons.
2B	.41	X	The Earth's shadow makes Moon phases.
5A	.41	X	The Sun is overhead every day.
12A	.38	X	Astrology is a science.
9B	.37		The Moon orbits the Earth in a day.
3A	.33	X	Inability to use one axis.
6C	.32	X	The Earth is 10x larger than it is.
3B	.31	X	Inability to use one axis.
15C	.30		Light exists only where it can be seen.
2C	.27	X	Moon moves through the Sun's shadow.

It is not unusual for students to prefer a misconception to the correct answer on this test. The majority of the misconceptions listed in Table I have P-values greater than the P-values of the correct answer. Moreover, an attractive misconception has a powerful effect on the P-value of an item's correct answer. The more attractive the misconception is, the lower the P-value of the correct answer.

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 questions, 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.

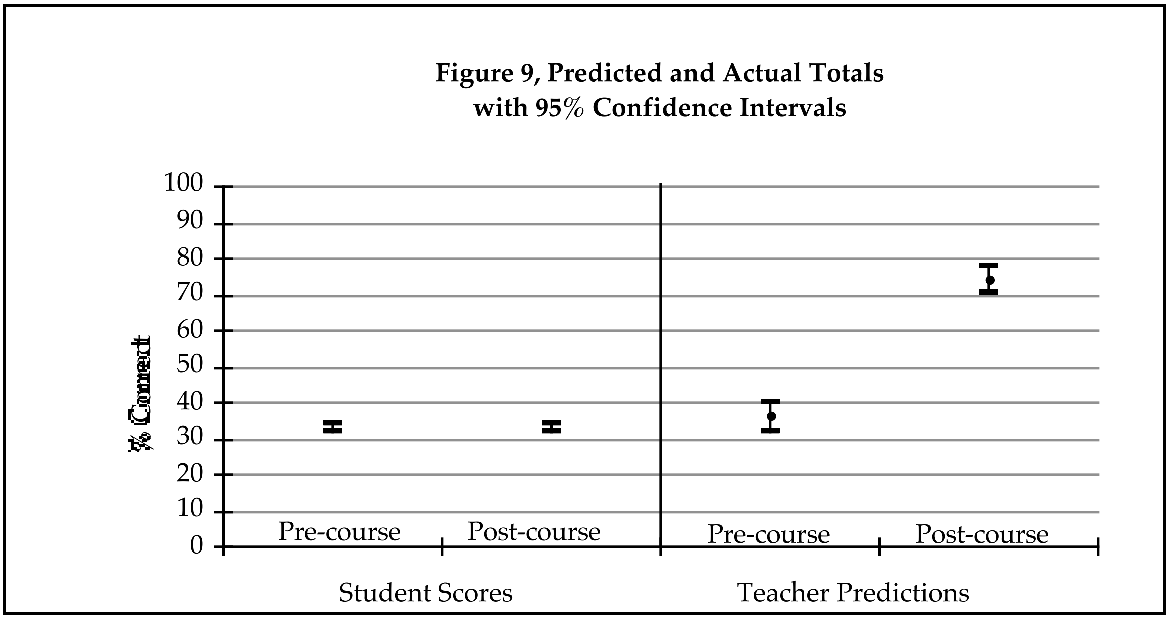


Figure 9, shows the overall findings of the study. The differences between teachers pre-course and post-course predictions were significant at the P=.05 level, while the differences in student performance were not. Teachers predicted large gains in student understanding of all the difficult concepts examined in this study. Although students did improve on some items, on average, there was no gain.

Interpretation of this study should be undertaken with the understanding that the results may be affected by other factors. In spite of our attempts to find a representative group of teachers and students, those used in this study may not be typical of the general population. Some teachers reviewing our work maintain that their students would do much better on this instrument than the students in this study. The few that were convinced to give the test to their student were uniformly surprised at how poorly they did on this test. However, since the students tested in this study were not the students of the teachers who predicted the scores, there is the possibility that these two populations were vastly different. One must wait for a study such as this one to be carried out having teachers predict the gains in the students that they teach. There is also the possibility that this instrument does not adequately measure the conceptual understanding of these basic astronomical concepts. We emphasize that our 16-question test is only a small subset of the concepts that astronomy teachers teach. No finite test could adequately reflect the material in a semester or year-long 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

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. This work has been supported by National Science Foundation (MDR 88-50424) and by the Smithsonian Institution.

IX. References

Anderson, B., and C. Karrqvist. "How Swedish pupils, aged 12-15 years, understand light and its properties." *European Journal of Science Education* 5 (4 1983a): 387-402.

Anderson, B., and C. Karrqvist. *Light and Its Properties*. Trans. by Gillian Thylander. Molndal, Sweden: University of Gothenburg, 1983b.

Camp, Carol Ann. "Problem solving patterns in science: Gender and spatial ability during early adolescence." Ed.D., University of Massachusetts, Amherst, 1981.

Cohen, Michael R. "How can the sunlight hit the Moon if we are in the dark? Teacher's concepts of phases of the Moon." Paper presented at *Henry Lester Smith Conference on Educational Research*, 1982.

Cohen, Michael R., and Martin H. Kagan. "Where does the old Moon go?" *The Science Teacher* 46 (1979): 22-23.

Duckworth, Eleanor. *"The Having of Wonderful Ideas" & Other Essays on Teaching and Learning*. New York: Teachers College Press, 1987.

Flick, L., "Preservice teachers conduct structured interviews with children to improve instructional methods," *2nd International Seminar on Misconception and Educational Strategies in Science and Mathematics in Ithaca, NY*, ed. by Joseph D. Novak, Cornell University Press, 1987: 113-117.

Guesne, Edith. "Light." In *Children's Ideas in Science*, ed. by Rosalind Driver, Edith Guesne, and Andrée Tiberhien. Philadelphia: Open University Press, 1985, pp. 10-32.

Gunstone, Richard F., and Richard T. White. "Understanding of gravity." *Science Education* 65 (3 1981): 291-299.

Hambleton, Ronald K., H. Swaminathan, and H. Jane Rogers. *Fundamentals of Item Response Theory.* Newbury Park, CA: Sage Publications, 1991.

Keuthe, James L. "Science concepts: A study of sophisticated errors." *Science Education* 47 (4 1963): 361-364.

Klein, Carol A. "Children's concepts of the Earth and the Sun: A cross cultural study." *Science Education* 65 (1 1982): 95-107.

Lightman, Alan, and Philip M. Sadler. "How can the Earth be round?" *Science and Children* (February 1986): 24-26.

Lightman, Alan, and Philip M. Sadler. "Teacher Predictions versus Actual Student Gains?" *The Physics Teacher* , 31:3 (March 1993): 162-167.

Mali, G., and A. Howe. "Development of earth and gravity concepts among Nepali children." *Science Education* 64 (2 1979): 213-221.

Marshall, Kim, and Oliver W. Lancaster. *Science: Elementary and Middle School Curriculum Objectives*. Boston Public Schools, 1983.

Narode, Ronald. "Standardized testing for misconceptions in basic mathematics." In *2nd International Seminar on Misconception and Educational Strategies in Science and Mathematics in Ithaca, NY*, ed. by Joseph D. Novak, Cornell University Press, 1987, pp. 222-333.

Novick, Shimshon, and Joseph Nussbaum. "Using interviews to probe understanding." *The Science Teacher* November (1978): 29-30.

Nussbaum, Joseph. "Childrens' conception of the Earth as a cosmic body: a cross age study." *Science Education* 63 (1 1979): 83-93.

Nussbaum, Joseph. "Students perception of astronomical concepts." In *GIREP - Cosmos - and Educational Challenge in Copenhagen*, ed. by J. Hunt. European Space Agency, 1986, pp. 87-97.

Nussbaum, Joseph. "The earth as a cosmic body." In *Children's Ideas in Science*, ed. by Rosalind Driver, Edith Guesne, and Andrée Tiberhien. Philadelphia: Open University Press, 1985, pp. 170-192.

Nussbaum, Joseph, and Joseph Novak. "An assessment of childrens' concepts of the earth utilizing structured interviews." *Science Education* 60 (4 1976): 535-550.

Osterlind, Steven J. *Constructing Test Items*. Boston: Kluwer Academic, 1989.

Rollins, M. M., J. J. Denton, and D. L. Janke. "Attainment of Selected Earth Science Concepts by Texas High School Seniors." *Journal of Educational Research* 77 (1983): 81-88.

Sadler, Philip M. "The Initial Knowledge State of High School Astronomy Students." Ed.D. dissertation, Harvard University, 1992.

Schoon, Kenneth J. "Misconceptions in Earth and Space Sciences, A Cross-Age Study." Ph.D. dissertation, Loyola University, 1988.

Sneider, Cary, and S. Pulos. "Childrens' cosmographies: understanding the Earth's shape and gravity." *Science Education* 67 (2 1983): 205-222.

Stead, B.F., K. E. and R. J. Osborne. "What is gravity? Some children's ideas." *New Zealand Science Teacher* 30 (1981): 5-12.

Targan, D. (1987). A Study of Conceptual Change in the Target Domain of the Lunar Phases. In J. D. Novak (Ed.), 2nd International Seminar on Misconception and Educational Strategies in Science and Mathematics, 2 (pp. 499-511). Ithaca, NY: Cornell University Press.

Touger, J. S. "Students' conceptions about planetary motion." Paper presented at *American Association of Physics Teachers in* 1985.

Treagust, David F., and Clifton L. Smith. "Secondary students understanding of the solar system: implications for curriculum revision." In *GIREP - Cosmos - and Educational Challenge in Copenhagen*, ed. by J. Hunt. European Space Agency, 1986, pp. 363-369.

Vosniadou, Stella, and William F. Brewer. "Theories of knowledge restructuring in development." *Review of Educational Research* 57 (1 1987): 51-67.

Wandersee, James H. "Can the history of science help science educators anticipate students' misconceptions?" *Journal of Research in Science Teaching* (1986).

Za'rour, George I. "Interpretation of natural phenomena by Lebanese school children." *Science Education* 60 (1976): 277-287.

Appendix 1, Table of Student Scores and Teacher Predictions

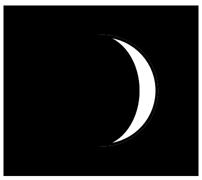
		adjusted teachers	Student scores N=236		Teacher Predictions N=203	
			Pre-course	Post-course	Pre-course	Post-course
1	c	Day/night	66%	77%	65%	89%
2	c	Phase change	25%	24%	34%	72%
3	m	HR Graph	30%	49%	47%	81%
4	c	Model of Sun & Earth	24%	29%	20%	61%
5	c	Sun overhead	18%	7%	27%	73%
6	f	Earth Diameter	29%	38%	36%	74%
7	c	Seasons	11%	17%	29%	76%
8	f	shuttle->planets->stars	44%	46%	49%	86%
9	f	Moon revolution	38%	41%	48%	85%
10	c	Moon revolution (Sun)	52%	50%	34%	73%
11	c	time zones	46%	35%	25%	63%
12	c	astrology	22%	11%	40%	80%
13	c	filters	15%	10%	23%	50%
14	c	light sources	39%	38%	35%	71%
15	c	light propagation - night	40%	42%	38%	73%
16	c	gravity	29%	18%	23%	65%
		facts	35%	34%	37%	75%
		concepts	31%	36%	39%	72%
		averages	33%	33%	36%	73%

Appendix 2, Project STAR Astronomy Prediction Questionnaire

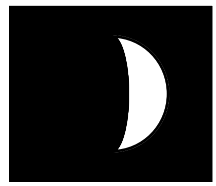
3/23/91

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:



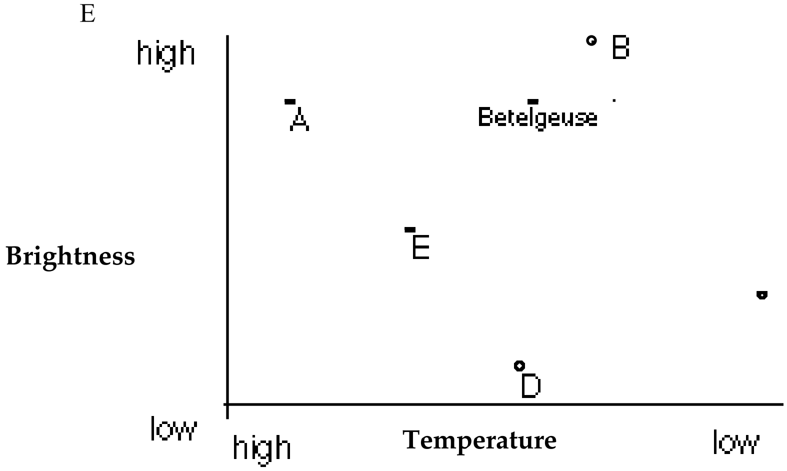
A few days later you looked again and saw this:



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. Using 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.

Give the best estimate of each quantity from among the numbers in the column on the right.

___ 6. Diameter of the Earth.

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 from the column on the right the letters that represent your best estimates of the times for the events listed in the left-hand column. Choices may be used more than once.

____9. The Moon to go around the Earth.
____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.

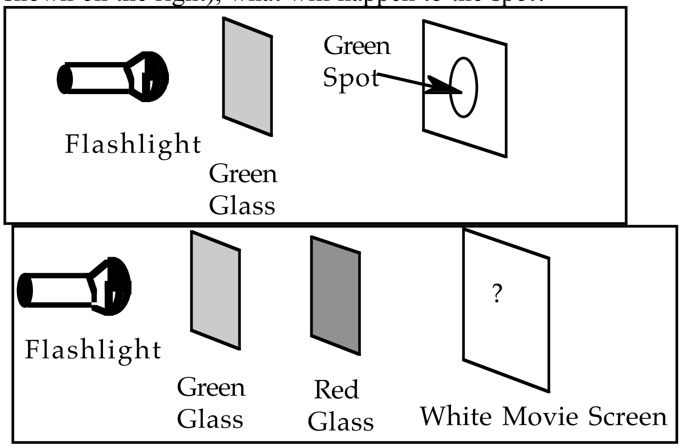
_ _

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.

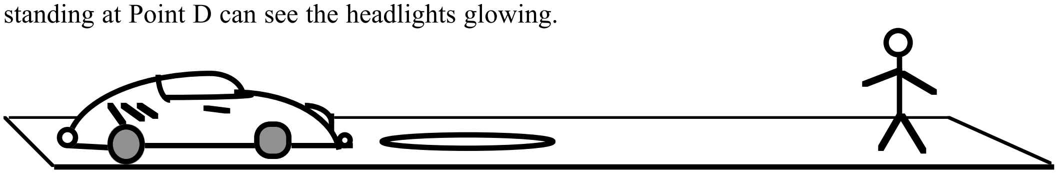
---

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.