A LIVELY ELECTRONIC COMPENDIUM OF RESEARCH, NEWS, RESOURCES, AND OPINION

Astronomy Education Review

Volume 4, Oct 2005 - Jul 2006
Issue 2

The Need for a Light and Spectroscopy Concept Inventory for Assessing Innovations in Introductory Astronomy Survey Courses

by **Erin M. (Weeks) Bardar**
Boston University
Edward E. Prather
University of Arizona
Kenneth Brecher
Boston University
Timothy F. Slater
University of Arizona
Received: 07/08/05, Revised: 12/07/05, Posted: 12/22/05

The Astronomy Education Review, Issue 2, Volume 4:20-27, 2006

Abstract

In this era of dramatically increased astronomy education research efforts, there is a growing need for standardized evaluation protocols and a strategy to assess both student comprehension of fundamental concepts and the success of innovative instructional interventions. Of the many topics that could be taught in an introductory astronomy course, the nature of light and the electromagnetic spectrum is by far the most universally covered topic. Yet, to the surprise and disappointment of instructors, many students struggle to understand underlying fundamental concepts related to light, such as blackbody radiation, Wien's law, the Stefan-Boltzmann law, and the nature and causes of emission and absorption line spectra. Motivated by predecessor instruments such as the Force Concept Inventory (FCI), the Astronomy Diagnostic Test (ADT), and the Lunar Phases Concept Inventory (LPCI), we call for, and are working on, the development and validation of a Light and Spectroscopy Concept Inventory. This assessment instrument should measure students' conceptual understanding of light and spectroscopy and gauge the effectiveness of classroom instruction in promoting student learning in the introductory astronomy survey course.

1. INTRODUCTION: THE NEED FOR A NEW ASSESSMENT INSTRUMENT

Although individual topics are included in many K–12 science education standards, astronomy is rarely offered as a stand-alone course in elementary, middle, or high school. Therefore, students' first formal introduction to many astronomy-related concepts in a cohesive manner is often at the college level. Research has demonstrated that many students enter these introductory college astronomy courses with tenaciously held prior knowledge and real-world experiences (Bailey & Slater 2003). When these preconceived commonsense theories are incorrect, they can be extremely difficult for students to overcome and can therefore negatively impact the effectiveness of instructional interventions (diSessa 1983; National Research Council 1999). Cognitive theory suggests that instructors need to be keenly aware of students' preexisting ideas prior to designing instruction in order to encourage the cultivation of more accurate scientific understanding (Nussbaum & Novak 1982). In particular, physics education research over the last 20 years has provided significant evidence that in order to construct informed instructional strategies that help students develop a rich conceptual understanding, it is essential that instructors be aware of the range and frequency of students' scientifically accurate and inaccurate ideas about the concepts being taught (Hestenes, Wells, & Swackhamer 1992).

In addition to calling attention to the need for identification of preinstruction misconceptions, the previously mentioned research in the closely related field of physics education has inspired tremendous changes in the way that some astronomy faculty think about the teaching and learning of science. This research has informed new and sometimes radical thinking about novel course designs (O'Donoghue 1998), innovative instructional approaches (Adams, Prather, & Slater 2003), and implementation of more authentic assessment strategies (Brissenden et al. 2002). However, despite these increased efforts from the astronomy education community, the introductory astronomy survey course is still taught primarily through only lectures that, unfortunately, have been shown to be largely ineffective in promoting significant changes in deep conceptual understanding (Prather et al. 2004). To enact widespread reform in the way that the introductory astronomy survey course is taught, compelling evidence must be provided to instructors to demonstrate which instructional interventions are most successful in producing meaningful learning gains. Consequently, an immediate need exists for a conceptual assessment tool that illustrates how students' naïve ideas evolve into more scientifically accurate interpretations as a result of the instructional intervention implemented in the astronomy classroom.

A viable solution to the need for an assessment instrument that is easy to administer and analyze is a research-based multiple-choice instrument sometimes referred to as a concept inventory. Such a test can provide insight into both student comprehension of fundamental concepts and the effectiveness of instructional strategies by probing students' understanding of key scientific concepts before and after formal instruction. Three relevant and successful examples from astronomy and physics are reviewed in the following section to establish the format and methodology associated with producing such an assessment instrument and to determine why there is need for a new instrument to serve as a standard for gauging the success of teaching and learning within the introductory college astronomy course.

2. PREVIOUS WORK

2.1 The Force Concept Inventory (FCI)

Cited more than any other work in physics education is the far-reaching work by Hestenes and his colleagues (1992), which focuses on the assessment of student conceptions of Newton's laws of motion. The Force Concept Inventory (FCI) was developed to serve as a diagnostic tool capable of assessing change in students' understanding of concepts related to force and motion, as a result of introductory university physics instruction. The FCI is a 29-item multiple-choice survey typically administered at the beginning and end of a course. The FCI requires no calculators and is written in the natural language of students, such that it is reliable and valid as a pretest when used prior to instruction. What is unique about this survey is that the items initially appear trivially easy to most faculty. However, research on how students actually respond to FCI questions reveals that the multiple-choice distracters are so enticing and attractive that students consistently and confidently select answers that are wholly incompatible with accurate scientific thinking. Even more revealing is the finding that traditional lecture-based instruction seems to have little, if any, impact on precourse/postcourse student gain scores. On the other hand, FCI results from classes that implemented innovative learner-centered instructional strategies informed by physics education research on student learning showed dramatic precourse/postcourse gains. In this way, the FCI has been able to successfully document that student achievement in traditional lecture-based physics courses consistently falls bellow the achievement of students who are taught using research-based instructional methods (Hake 1998).

The FCI has dramatically raised physics instructors' awareness of the shortcomings of their students' performance, even after instruction, on what they deemed relatively simple questions (Mazur 1997). As such, the FCI has been the catalyst for a paradigm shift in the way that many physicists view their instructional practices and has resulted in a number of extremely successful curriculum reform efforts. This one test has established itself as the standard by which different physics instructional methods are compared and deemed successful (or not). It is important to emphasize that the FCI does not try to cover all the topics covered in the conventional college physics course. Rather, it covers only the topics of Newton's laws of motion. Yet, the FCI is a commonly administered assessment instrument even when a course's curriculum addresses a series of topics far broader than Newton's laws.

The widespread adoption of the FCI can be attributed to three key factors. First, the test has a tight conceptual focus on a central theme, Newton's laws, which is common to all mechanics-based introductory physics courses. Second, the FCI generates consistent results across a broad range of courses, from high school through college level and across instructional styles, indicating the instrument's versatility within the field. Finally, FCI results help illustrate how instructors often overestimate the likely success of their students. An astronomy assessment instrument that mirrors these qualities would be extremely valuable for evaluating the effectiveness of instruction in the field.

2.2 The Astronomy Diagnostic Test (ADT)

The Astronomy Diagnostic Test (ADT) was a collaborative effort to create a single instrument that could assess student understanding across a broad array of astronomy topics. The ADT 2.0 is a 33-question multiple-choice survey consisting of 21 content questions and 12 demographic questions (Hufnagel 2002). Like the FCI, the ADT is a well-constructed test grounded in thorough qualitative research such that the wrong answers, or distracters, reflect ideas commonly held by the target student population. Each question is carefully structured to avoid scientific jargon and to focus on a single concept. The original authors of the ADT intended that the test be used to indicate the initial knowledge state of undergraduates taking the introductory astronomy course with regard to their understanding of topics that would commonly have been covered during a typical K–12 education experience. As such, the topics on the ADT are broad and reflect subject matter from the standards-based curriculum of K–12 instruction in the United States; consequently, they should be recognizable to most high school graduates. To this end, the ADT items focus on size and scale, night sky motion, lunar phases, and physical laws of motion. Notably absent are several topics commonly taught at the college level, such as spectroscopy; luminosity; the H-R diagram; cosmology and the expansion of the universe; formation, evolution, and properties of stars; and galaxies.

In recent years, assessment of introductory college astronomy survey courses, particularly those for non–science majors and preservice education majors, have shown that gains in students' understanding of many astronomy concepts, as determined by ADT scores, are far below the gains that most faculty predicted (Deming 2002). One reason for this might be that the topics on which the ADT is centered do not accurately represent the relative emphasis given to those topics in college-level courses, or perhaps there is insufficient clustering of related questions to allow each topic to be probed in enough depth to fully gauge students' conceptual understanding of the subject matter. From a design standpoint, we know that for a particular topic with multiple facets, several questions are required to sufficiently probe student understanding of each facet, and a large total number of questions are required to completely assess the entire concept domain (Frayer, Frederick, & Klausmeier 1969). The ADT was not designed to be such an instrument.

As an example of the ADT's lack of power to distinguish the relative effectiveness of various teaching interventions, we look at the research done by Alexander (2005), which showed no significant difference in ADT gain scores between a traditional lecture-based course and an equivalent course employing modified instruction in the form of *Lecture-Tutorials for Introductory Astronomy* (Adams et al. 2003). This finding suggests that the implementation of an interactive learner-centered intervention does not elevate student learning. This result stands in opposition to the findings of Hake (1998) regarding the FCI and the research on the efficacy of lecture-tutorials done by Prather et al. (2004). This forces another look at the Alexander study to investigate at what level there is sufficient overlap in topics chosen from the lecture-tutorials implemented and the questions on the ADT. Closer inspection of the study reveals that only 5 of the 21 content questions on the ADT correspond to the topics covered by the 10 lecture-tutorial activities employed in the modified course. As a result, it is not surprising that the ADT lacked the sensitivity to measure any statistically significant changes over such a small range of topic overlap.

Although it is tempting to distill the shortcomings of the ADT to the fact that numerous commonly taught topics are missing from the set of questions on the test, it would incorrectly suggest that we believe that more questions added to the ADT over a broader range of topics would make the test more suitable as a universal standard of astronomy teaching and learning. This solution path would simply lead to the creation of an assessment tool with far too many questions to be reliable or easily administered. What the astronomy education community needs is a new instrument that adequately covers a relatively narrow conceptual domain that is common to nearly all college astronomy courses rather than a broad survey that has disconnected questions representing the entire "universe of topics" that one might elect to teach in astronomy.

2.3 The Lunar Phases Concept Inventory (LPCI)

The Lunar Phases Concept Inventory (LPCI) is the astronomy community's closest counterpart to the FCI. The LPCI was fashioned to be a highly reliable and valid research-based concept inventory that targets lunar phases, a single coherent topic area commonly taught in the introductory astronomy course. This test is composed of 20 multiple-choice content questions and an additional eight demographic questions. The instrument is intended to assess individual students' understanding of lunar phases with respect to motion of the Moon; cause of lunar phases; period of lunar phases; period and direction of the Moon's orbit; and observational phenomena (Lindell 2001; Lindell & Olsen 2002; Lindell & Sommer 2003). The instrument was designed to map a student's mental model of lunar phases over a series of well-defined concept dimensions identified through in-depth qualitative interviews (Lindell et al. 2001). The research results from administering this test to a range of classrooms illustrate that it is possible to create a successful astronomy concept inventory around a single topic that can measure the effectiveness of instructional techniques and provide insight into individuals' conceptual frameworks (Lindell 2004). Although the LPCI more closely parallels the nature of the FCI than the ADT in its narrowness of conceptual focus, student understanding of the topic of lunar phases is not widely accepted as an overarching theme and core learning goal for the entire introductory astronomy survey course, but rather a topic to which instructors allocate perhaps a single class period. Therefore, student understanding of lunar phases, though arguably a desired course learning objective for many instructors, is not suitable as a "central topic" to use for assessing the teaching and learning of all introductory astronomy courses.

3. CHOOSING A STANDARD

The question now arises: What *is* the "central topic" that would serve as the best candidate for a single concept inventory to be universally adopted to assess teaching and learning within introductory astronomy survey courses? The overwhelming acceptance of the FCI encourages us to examine exactly why Newton's laws were the most fruitful choice for a topical focus in physics education. Although Newton's laws do not constitute the entire content of an introductory physics course, they are perhaps the most important concepts upon which the rest of mechanics is built and therefore provide a natural focus for the assessment of instructional success within the field (Halloun & Hestenes 1985). Introductory astronomy courses often cover a much broader range of topics in a single semester, making it more difficult to isolate a single fundamental concept.

However, a recent large-scale survey of college and university faculty who teach introductory astronomy courses cited the nature of light and the electromagnetic spectrum as the most frequently taught topics (Slater et al. 2001). Moreover, a sample of 44 participants in an American Association of Physics Teachers workshop entitled Teaching Astronomy Conceptually ranked the electromagnetic spectrum as the most important "Astronomy 101" topic out of a list of 200 choices (Zeilik & Morris-Dueer 2005), yet evidence shows that it is also a topic area with which students struggle (Brecher 1991). We believe that to satisfy the critical need within the astronomy education community for a valid, reliable, and easily administered assessment tool that measures the effectiveness of instructional interventions in the introductory astronomy survey course, we should concentrate on light and the electromagnetic spectrum.

Light is the fundamental carrier of astronomical information. Spectral features serve as the "fingerprints" of the universe, revealing temperature, elemental compositions, and relative motion, along with many other important properties of objects in the cosmos. Therefore, in much the same way that Newton's laws of motion have become the central topic for which the FCI serves as a standard litmus test for physics teaching and learning, we posit that the nature of light, the electromagnetic spectrum, and spectroscopy are the best choices for the central topics on which to focus the development of a concept inventory that will allow the community of astronomy teachers to compare the success of teaching and learning that take place in their introductory astronomy courses.

Such a Light and Spectroscopy Concept Inventory will be most useful if it has a tight topic focus with sets of clustered questions on discrete subtopics within the broader concept domain, as exemplified by the FCI and LPCI. The concept domain of this inventory must directly reflect the most commonly taught concepts addressed by the majority of courses within the astronomical community, such as properties of the electromagnetic spectrum; Doppler shift; Wien's law; the Stefan-Boltzmann law; and Kirchhoff's laws of spectral analysis (Slater & Adams 2003; Slater et al. 2001; Zeilik & Morris-Dueer 2005). To earn the respect and acceptance of the astronomy education community, the inventory must abide by the same rigorous research-based development procedures followed by the ADT, FCI, and LPCI. Extensive qualitative research and field testing must be conducted to ensure the production of a reliable and valid instrument, with content approved by professionals in the field and presented in the natural language of students.

3.1 Progress

We have been developing such an inventory in an effort to address the need identified in this article for a single conceptual assessment tool that is easy to administer. We are now in our third year of researching students' understanding of light and spectroscopy-related phenomena within the context of introductory college astronomy. Through a systematic investigation that has included multiple rounds of clinical interviews, open-ended written surveys, and multiple-choice testing, we have developed the Light and Spectroscopy Concept Inventory (LSCI). The LSCI is currently being tested at multiple institutions to establish reliability and validity. Following this current path of iterative testing, we plan to have a completed inventory within the next year.

4. SUMMARY

In this article, we have endeavored to illustrate that there is an immediate need for a conceptual assessment tool that is able to evaluate how students' naïve ideas evolve into more scientifically accurate interpretations as a result of the instructional intervention implemented in the astronomy classroom. This assessment tool must be focused on a central topic common to introductory astronomy instruction. We believe that selecting a combination of light and spectroscopy is the natural choice because they constitute the foundation for nearly all astronomical research, and moreover, they are the most commonly taught topics in introductory college astronomy. Following the rigorous research and development standards set by the FCI, ADT, and LPCI, this new assessment tool must be informed by known students' beliefs and reasoning difficulties. Most important, this assessment tool should have the ability to gauge the effectiveness of various instructional strategies and materials. If these criteria are met, we are hopeful that such an inventory will, within the astronomy community, parallel the widespread acceptance and utility that the FCI has enjoyed within the physics community. In an effort to attain this goal, we are currently in the late stages of developing and validating the Light and Spectroscopy Concept Inventory. It is our hope that once the LSCI is completed, instructors of introductory astronomy who teach in all classroom environments—from small classes to large lecture courses, community colleges to universities—will systematically implement this assessment tool in their courses to evaluate the effectiveness of their teaching. Further, we encourage these instructors to share their results with the entire astronomy education community so that together we might advance our understanding of the extent to which our instructional strategies are actually helping students learn. We believe that an effort on behalf of the astronomy education community as described here may bring about much-needed reform in the way that introductory college astronomy courses are taught nationwide.

Acknowledgment

Funding for this project is provided in part by NSF Grant No. DUE-0125992 and NASA GSRP Grant No. NGT5-60482.

References

Adams, J. P., Prather, E. E., & Slater, T. F. 2003, *Lecture-Tutorials for Introductory Astronomy*, Upper Saddle River, NJ: Prentice Hall.

Alexander, W. R. 2005, Assessment of Teaching Approaches in an Introductory Astronomy College Classroom, *Astronomy Education Review*, 2(3), 178.

Bailey, J., & Slater, T. F. 2003, A Review of Astronomy Education Research, *Astronomy Education Review*, 2(2), 20.

Brecher, K. 1991, Do Atoms Really ‘Emit’ Absorption Lines?, *The Physics Teacher*, 29, 454.

Brissenden, G., Slater, T. F., Mathieu, R. D., & NISE College Level-One Team. 2002, The Role of Assessment in the Development of the Introductory Astronomy Course, *Astronomy Education Review*, 1(1), 1.

Deming, G. L. 2002, Results from the Astronomy Diagnostic Test National Project, *Astronomy Education Review*, 1(1), 52.

diSessa, A. 1983, Phenomenology and the Evolution of Intuition, in *Mental Models*, Albert L. Stevens (Editor), Hillsdale, NJ: Lawrence Erlbaum Associates, 15.

Frayer, D. A., Frederick, W. C., & Klausmeier, H. G. 1969, *A Schema for Testing the Level of Concept Mastery*, Technical Report No. 16, Madison: University of Wisconsin, Wisconsin Center for Education Research.

Hake, R. R. 1998, Interactive-Engagement versus Traditional Methods: A Six-Thousand-Student Survey of Mechanics Test Data for Introductory Physics Courses, *American Journal of Physics*, 66, 64.

Halloun, I., & Hestenes, D. 1985, The Initial Knowledge State of College Physics Students, *American Journal of Physics*, 53, 1043.

Hestenes, D., Wells, M., & Swackhamer, G. 1992, Force Concept Inventory, *The Physics Teacher*, 30, 141.

Hufnagel, B. 2002, Development of the Astronomy Diagnostic Test, *Astronomy Education Review*, 1(1), 47.

Lindell, R. 2001, *Enhancing College Students' Understanding of Lunar Phases*, Unpublished Dissertation. University of Nebraska-Lincoln.

Lindell, R. 2004, Measuring Conceptual Change in College Students' Understanding of Lunar Phases, *American Institute of Physics Conference Proceedings*, 790, 53.

Lindell, R., & Olsen, J. 2002, Developing the Lunar Phases Concept Inventory, in *Proceedings of the 2002 Physics Education Research Conference*, S. Franklin, J. Marx, & K. Cummings (Editors), New York: PERC Publishing.

Lindell, R., & Sommer, S. 2003, Using the Lunar Phases Concept Inventory to Investigate College Students' Pre-instructional Mental Models of Lunar Phases, *American Institute of Physics Conference Proceedings*, 720, 73.

Lindell, R., Bean, M., Meltzer, D., & Fanetti, T. 2001, *Developing a Concept Inventory to Assess Student Models of Lunar Phases*, Contributed poster, American Association of Physics Teachers national meeting, San Diego, CA.

Mazur, E. 1997, *Peer Instruction*, Upper Saddle River, NJ: Prentice Hall.

National Research Council. 1999, *How People Learn: Brain, Mind, Experience, and School*, Washington, DC: National Academy of Sciences.

Nussbaum, J., & Novak, S. 1982, Alternative Frameworks, Conceptual Conflict and Accommodation: Toward a Principled Teaching Strategy, *Instructional Science*, 11, 183.

O'Donoghue, A. A. 1998, Studio Astronomy, *Bulletin of the American Astronomical Society*, 30, 1335.

Prather, E. E., Slater, T. F., Adams, J. P., Bailey, J. M., Jones, L. V., & Dostal, J. A. 2004, Research on a Lecture-Tutorial Approach to Teaching Introductory Astronomy for NonScience Majors, *Astronomy Education Review*, 2(3), 122.

Slater, T. F., & Adams, J. P. 2003, *Learner-Centered Astronomy Teaching: Strategies for ASTRO 101*, Upper Saddle River, NJ: Prentice Hall.

Slater, T. F., Adams, J. P., Brissenden, G., & Duncan, D. 2001, What Topics Are Taught in Introductory Astronomy Courses?, *The Physics Teacher*, 39, 52.

Zeilik, M., & Morris-Dueer, V. J. 2005, What Are Essential Concepts in 'Astronomy 101'? A New Approach to Find Consensus from Two Different Samples of Instructors, *Astronomy Education Review*, 2(3), 61.