Abstract
Through the use of open-ended surveys, students’ ideas about the properties and formation of stars were investigated. It was determined that, although students often know some basic properties of stars, their understanding lacks detail about fundamental issues such as nuclear fusion. These results led to the development and validation of the Star Properties Concept Inventory (SPCI), which can be used as a pretest/posttest to determine students’ growth in this content area. For participating ASTRO 101 courses, the average pretest score on the SPCI was about 31%, while the average posttest score was 51%.

1. INTRODUCTION
As cognitive science expands our understanding of how people learn, educators are placing more emphasis on identifying the preexisting student beliefs and reasoning difficulties that students bring into the learning environment. These ideas can strongly influence, in both positive and negative ways, the development of new knowledge. If a teacher has an idea of the preinstructional beliefs that are common among students, she may tailor her instruction to incorporate, for example, specific activities to promote meaningful conceptual change (Posner, Strike, Hewson, & Gertzog 1982).

When we look into the sky, we see the brightness of the Sun during the daytime; we see thousands of points of light—stars—in the Sun’s absence. For thousands of years, the stars played important roles in various religions and mythologies. In the 20th century, we determined that the Sun, the primary source of energy for the entire planet, is the same type of object as those tiny white dots. Given their importance in our cultural and biological history, it comes as no surprise that stars are considered one of the most important topics in astronomy. From a survey of U.S. college syllabi available on the Internet, Slater,
Adams, Brissenden, and Duncan (2001) report that stellar evolution ranked as one of the 10 topics most frequently covered in an undergraduate introductory astronomy course. Other topics commonly taught include the nature of light, cosmology, our Sun, lunar phases, and characteristics of the Milky Way galaxy.

This topic is also important for middle and high school students. *Evolution and equilibrium* is a unifying concept and process highlighted in the National Research Council’s *National Science Education Standards* (NSES; 1996) as an instructional topic for all students. Changes over time occur in a variety of natural processes and contexts, and the evolution of stars and their contribution to the evolution of the universe as a whole is a topic that can support this aspect of the NSES. In the earth science content standards, the history of Earth (Grades 5–8) and formation of our Solar System (Grades 9–12) are both recommended for study. There are several content standards in physical science (Grades 9–12 in particular) that the study of stars and stellar evolution supports; “interactions of energy and matter” is perhaps the best example (National Research Council 1996, 176). Adams and Slater (2000) list 11 NSES content strands in which astronomy is either explicitly or implicitly addressed, along with the relevant research to date focusing on alternative conceptions. Project 2061’s *Benchmarks for Science Literacy* (American Association for the Advancement of Science 1993) includes more detailed content benchmarks, some of which very specifically relate to stars. Slater (2000) lists 27 statements from the *Benchmarks*, organized by grade level, that contain astronomy content. If a student’s K–12 science experiences were based on either the NSES or the *Benchmarks*, students might reasonably be expected to understand that the Sun is a star, that the Sun formed from a cloud of gas and dust, that stars vary in properties such as temperature and mass, and that stars emit light because of nuclear fusion at their cores.

Unfortunately, there exists almost no research on students’ preexisting beliefs and reasoning difficulties about stars, star formation, or star death (Bailey & Slater 2003, 2005). This dearth of information about students’ understanding about stars, despite coverage in both high school standards and college astronomy courses, must be addressed in order to design effective learning events that explicitly elicit students’ naïve ideas and help them confront and resolve any scientifically inaccurate conceptions.

### 2. RESEARCH DESIGN

To provide effective instruction for our students, we must first be able to identify where their preexisting concepts about the topic stand to interfere with instruction. This was done through two phases of this study. In Phase I, students’ ideas about the properties and formation of stars were investigated using student-supplied response (hereafter SSR) surveys and semistructured interviews. Responses were analyzed and thematically grouped to develop the distractors for a concept inventory—a focused, multiple-choice instrument—that can be used to identify those understandings pre- and postinstruction. The development and validation of the resulting Star Properties Concept Inventory (SPCI) comprised Phase II of the study.

#### 2.1 Setting and Participants

This study was conducted at a large research university in the southwestern United States. The participants of Phase I were undergraduate students enrolled in an introductory astronomy course for non-science majors (hereafter ASTRO 101). In Phase II, new ASTRO 101 students comprised the primary group of participants. A comparison group of students in other introductory science courses (such as earth science or atmospheric science) also participated; hereafter, these courses will be labeled ES 101. Like ASTRO 101, these courses satisfied a general education requirement for all students. Most of the participants were
classified as freshmen or sophomores, and the gender and ethnic makeup of the courses was approximately representative of the university as a whole.

2.2 Phase I: Student-Supplied Response Surveys

Short (one to four questions) SSR surveys were administered to a total of 18 sections of the ASTRO 101 course during the summer 2003, fall 2003, spring 2004, and fall 2004 semesters. SSR surveys were administered on the first or second day of class in sections whose 10 different instructors volunteered access to their courses. To maximize the information gathered from the classes, different versions of the SSR survey were administered. A total of 13 different SSR survey forms, including 11 different main questions (plus variations on the wording of some questions), were administered to approximately 2,500 students. After being randomly assigned identification numbers, surveys were inductively analyzed and coded for themes. The codes were then tabulated and the frequencies of the themes calculated. A small number of semistructured interviews were conducted with student volunteers to confirm the results of the SSR surveys.

2.3 Phase II: Development and Validation of the Star Properties Concept Inventory

Using the results from Phase I, an instrument containing approximately 30 multiple-choice questions on the properties and formation of stars was created. Distractors were, whenever possible, formulated from students’ responses on SSR survey questions of a similar nature. Two early versions of the instrument were used during the spring 2005 semester (Version 1 pretest, Version 2 posttest). To test the reliability of the questions, Version 1 was created with three different types of questions: multiple choice only, multiple choice plus "explain your reasoning," and SSR using the stem of the multiple-choice question. All three versions were randomly distributed at the start of spring 2005 to approximately 1,000 students from six sections of ASTRO 101 and two sections of ES 101. The results of the different types of questions were compared to see if students responded in a similar manner.

Midway through the semester, interviews were conducted with a randomly selected subset of participants. During the interview, participants were asked to retake the concept inventory while talking through their thoughts and answers. After the completion of the instrument, students’ new responses were compared with their original answers and any differences discussed.

Based on the results of the Version 1 testing, some changes were made to the questions or to individual distractors as needed. For example, if it was found that participants in the SSR version (1c) provided more responses of a different type than one of the distractors (versions 1a and 1b), that distractor was changed. Version 2 was administered in two formats at the end of spring 2005 to approximately 500 participants. As before, the formats were randomly distributed and the results compared. Although the questions were not identical, they were close enough that a precourse/postcourse comparison was made.

A final round of changes was made prior to fall 2005. Questions or distractors that proved to be weak were changed or removed outright. The questions were also reviewed for content validity by more than 25 astronomy instructors from across the country. Version 3, containing 23 content items, was administered to approximately 1,100 participants in five sections of ASTRO 101 and three sections of ES 101 on the first day of class and again at the end of the semester. Additional details of the development process are
2.4 Description of the Star Properties Concept Inventory Version 3

The final version of the SPCI comprises 23 content items plus two background questions (see Note 1). Content items fall into three major categories: star properties (13 questions), fusion (5 questions), and star formation (5 questions). The star properties questions can be further subdivided into the categories of mass, temperature/color, luminosity, and mass–lifetime relationship, with two to four questions each. Each content area includes multiple questions to increase the likelihood that a student who correctly answers a question is doing so because she knows the content, as opposed to simply making a "lucky guess." The two background questions include gender and whether the participant has taken an astronomy class prior to the one in which the SPCI is being administered.

3. FINDINGS

3.1 Phase I: Student-Supplied Response Surveys

Results from the SSR surveys show that while participants understand some basic properties of stars and have ideas about how stars are formed, their ideas are typically incomplete and often scientifically inaccurate. For example, approximately 200 of the participants were asked to describe what a star is. Although over 40% of the participants described a star as a "ball of gas," a reasonable first start, only 4% included in their response nuclear fusion—the process of energy generation that astronomers believe to be the distinguishing feature of stars. Common alternative conceptions include the notion of stars being "on fire" or otherwise "burning," or that the energy was released by chemical reactions.

Of nearly 1,000 participants who were asked to describe the process of star formation, less than 1% were able to provide a complete and correct answer. Almost half of the respondents gave components of a correct answer but were either missing critical information or had elements that were incorrect. Most often missing from their explanations was the temperature increase that occurs during matter’s gravitational collapse, which eventually allows nuclear fusion to begin.

More than 1,000 participants were asked how light is created in stars; the desired answer here was through nuclear fusion. Nearly 90% of the responses to this question were incorrect, with common responses again including that the light we see from stars is created through burning or chemical reactions. Additional results from the SSR surveys are forthcoming (Bailey, Prather, Johnson, & Slater 2007b).

3.2 Phase II: Development and Validation of the Star Properties Concept Inventory

The distribution of responses was quite similar over the different formats for both Version 1 (pretest) and Version 2 (posttest). Midsemester interviews conducted with a subset of participants showed consistency between administrations; where answers had changed, the participants could identify recent instruction on the topic in their ASTRO 101 course. The mean scores (percentage) on the multiple-choice versions are shown in the Table 1. Note that the varying number of participants is a result of multiple factors, including elimination of the open-response version in the posttest (Version 2), absenteeism, and attrition. The drop in score in the ES 101 courses is as yet unexplained, but early thoughts center on the population of the
students who took both tests (i.e., students who scored high on the pretest either dropped the course or were absent from the posttest).

**Table 1. Mean Scores for Different Classes on Two Administrations of the SPCI**

<table>
<thead>
<tr>
<th>Class</th>
<th>Pretest (%)</th>
<th>Posttest (%)</th>
<th>Pretest (%)</th>
<th>Posttest (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASTRO 101</td>
<td>39 ± 18 (N = 396)</td>
<td>59 ± 21* (N = 469)</td>
<td>31 ± 12 (N = 586)</td>
<td>51 ± 17* (N = 417)</td>
</tr>
<tr>
<td>ES 101</td>
<td>44 ± 18 (N = 85)</td>
<td>32 ± 19* (N = 76)</td>
<td>31 ± 13 (N = 334)</td>
<td>30 ± 13 (N = 112)</td>
</tr>
</tbody>
</table>

*p < 0.05 between pretest and posttest scores, and between ASTRO 101 and ES 101 posttest scores in fall 2005*

During fall 2005, Version 3 was administered to 1,101 participants as a pretest (644 completed the posttest) in both ASTRO 101 and ES 101 classes. After removing responses that were incomplete or careless (e.g., choosing all Bs), means were calculated for both groups for the pretest and posttest administrations (Table 1). Using the subset of ASTRO 101 responses that could be matched between pretest and posttest, an effect size (Cohen’s $d$; Thalheimer & Cook 2002) of 1.35 was calculated. Additional details of the design and results of the SPCI are forthcoming (Bailey et al. 2007a).

**4. CONCLUSIONS**

In this research project, the range and frequency of preinstructional beliefs about star properties and formation held by students in an introductory college astronomy course for non–science majors has been determined through the use of SSR surveys, and this information has been used to inform the development of a concept inventory that can be used to measure instructional effectiveness on these topics.

Overall, this research study has provided much-needed information about what ASTRO 101 students understand about stars and star formation. The SPCI, a 23-item multiple-choice test, has been created and evaluated, and is a valid and reliable instrument that can measure conceptual change on the topics of star properties and formation over an instructional period. The astronomical community is poised to use inventories such as this to examine instructional effectiveness as it transitions to learner-centered instruction. With further research conducted using this instrument, instructors may develop instructional methods and strategies that will lead to a deeper and more profound understanding of the properties and formation of stars.

Several avenues of further research can be followed to build on these results. One important research project that should be conducted before the SPCI becomes widely adopted is the investigation of its generalizability across institutions (see Note 1). After having multiple instructors at a variety of different institutions administer the SPCI both pre- and postinstruction, an analysis of the data could help determine whether the students at the researcher’s home institution were, in fact, reasonably representative of other ASTRO 101 students both before and after instruction. Once this level of generalizability is determined,
data can be analyzed in a manner similar to that of Hake (1998) in investigating normalized gains of interactive engagement versus traditional lecture-based instruction. If it is determined that, like the traditional physics courses reported in Hake (1998), traditional astronomy courses are ineffective at promoting deep conceptual understanding about stars, then the development of additional interactive engagement methods or focused curricular materials designed to intellectually engage students could follow. Early reports using other assessment strategies have indicated that this is the case (Hudgins 2005; Prather et al. 2004).

Another possibility for further research is the expansion of this study to related topics. This research study deliberately focused on only a few basic properties of stars and the process of star formation. The breadth of student ideas about stellar evolution was not addressed. Data from the SSR surveys and interviews indicate that students have heard of other stages of stellar evolution—most commonly black hole, red giant, and white dwarf—and it was clear that these terms were not always used in a way that can be considered scientifically accurate. The investigation of these ideas could lead to the creation of an additional concept inventory on stellar evolution.

Note

Note 1. To protect the validity of the instrument, it is not provided here. If you are interested in using the SPCI in your classes, please contact the author, Janelle Bailey, via e-mail: janelle.bailey@unlv.edu.

Acknowledgments

This work was completed in partial fulfillment of the doctor of philosophy granted by the Graduate College of the University of Arizona. The dissertation committee was composed of Dr. Bruce Johnson (Department of Teaching and Teacher Education; codirector), Dr. Edward E. Prather (Steward Observatory; codirector), and Dr. Timothy F. Slater (Department of Astronomy).

This research project could not have been completed without the participation of more than 20 instructors and more than 4,000 students over three years. To them, and to my committee members, I am truly indebted. Thanks also go out to Robin Ward, Kevin Vinson, Doug Lombardi, and members of the CAPER Team at the University of Arizona for their input and assistance throughout the project.

References


ÆR
133 - 139