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

Astronomy Education Review

2010, AER, 9, 010107-1, 10.3847/AER0009068

A National Study Assessing the Teaching and Learning of Introductory Astronomy Part II: The Connection between Student Demographics and Learning

Alexander L. Rudolph
Department of Physics, California State Polytechnic University, Pomona, California, 91768
Edward E. Prather
Department of Astronomy and Center for Astronomy Education, University of Arizona, Tucson, Arizona, 85719
Gina Brissenden
Department of Astronomy and Center for Astronomy Education, University of Arizona, Tucson, Arizona, 85719
David Consiglio
Bryn Mawr College, Bryn Mawr, Pennsylvania, 19010
Vicente Gonzaga
Department of Physics, California State Polytechnic University, Pomona, California, 91768
Received: 11/9/09, Revised: 02/3/10, Accepted: 02/3/10, Published: 04/19/10

Abstract

This is the second in a series of reports on a national study of the teaching and learning of astronomy in general education, nonscience major, introductory college astronomy courses (hereafter referred to as Astro 101). The analysis reported here was conducted using data from nearly 2000 students enrolled in 69 Astro 101 classes taught across the country. These students completed a 15-question demographic survey, in addition to completing the 26-question Light and Spectroscopy Concept Inventory (LSCI) pre- and post-instruction. The LSCI was used to determine students' learning via a normalized gain calculated for each student. A multivariate regression analysis was conducted to determine how ascribed characteristics (personal demographic and family characteristics), achieved characteristics (academic achievement and student major), and the use of interactive learning strategies are related to student learning in these classes. The results show dramatic improvement in student learning with increased use of interactive learning strategies even after controlling for individual characteristics. In addition, we find that the positive effects of interactive learning strategies apply equally to men and women, across ethnicities, for students with all levels of prior mathematical preparation and physical science course experience, independent of GPA, and regardless of primary language. These results powerfully illustrate that all categories of students can benefit from the effective implementation of interactive learning strategies.

1. INTRODUCTION

This article is the second in a series of articles describing the results of a national study of student learning in college level, general education, introductory astronomy courses (hereafter referred to as Astro 101). These courses enroll 250 000 students each year nationwide and are taken by 10% of all students at some time in their college careers, making it one of the most popular general education courses (Fraknoi 2001; Partridge and Greenstein 2003). This study was designed to investigate teaching and learning in these classes, with special emphasis placed on the effect of interactive learning strategies on student conceptual understanding. Considerable evidence from both physics and astronomy education research has shown that such strategies can improve student understanding of key concepts beyond what is achieved when more traditional lecture methods are used (Hake 1998; Crouch and Mazur 2001; Prather *et al.* 2004; Hudgins *et al.* 2006).

Students in the study were given the Light and Spectroscopy Concept Inventory (LSCI) (Bardar *et al.* 2005; Bardar *et al.* 2006) pre- and post-instruction in an effort to measure their gain in understanding of topics central to almost all Astro 101 classes (Slater *et al.* 2001; Zeilik and Morris-Dueer 2004). In addition to calculating the normalized gain[1] achieved by students from the LSCI, we also asked the instructors of each class to complete a survey known as the Interactivity Assessment Instrument (IAI). From the instructors' self-reported data, we were able to calculate the percent of total class time spent teaching with interactive learning strategies (which we called the Interactive Assessment Score or IAS). The first article in this series (hereafter referred to as Paper I, Prather *et al.* 2009) reports on the relationship between class-based pre-test scores, class-based normalized gains, institution type, class size, and level of interactivity in each classroom. In addition to the 26 astronomy questions contained in the LSCI, some students were also given a set of 15 demographic questions. In this article, we report on our analysis of the complete data set including the LSCI, IAI, and the demographic questions. We begin by briefly outlining our study methodology, and then report the frequencies of the various demographic categories as a snapshot of who is taking Astro 101 nationally. We then describe a series of multivariate regression models designed to determine how the various student characteristics and the interactivity level in the classrooms affect student learning.

2. STUDY METHODOLOGY

Paper I (Prather *et al.* 2009) outlined the study methodology in detail; however, we will provide the essentials here. A total of 3729 students took the LSCI pre-instruction (pre-test), and 2577 took it post-instruction (post-test) using a Scantron™ form. These students came from 31 institutions of all types (both 2 year and 4 year) and from classes ranging in size from fewer than 10 to 180 students. In addition, some of the students[2] were asked to answer 15 demographics questions, listed in Table A1 in the Appendix. These demographic questions inquired about ascribed characteristics (gender, native language, ethnic background, parents' education level, and parents' income level) and achieved characteristics (elementary and high school type, high school and college GPA, college major, and previous math and science courses taken). The Scantron™ forms were read, coded with a unique identifier for each student, and the data were entered into SPSS for analysis. The student identifier was then used to calculate a matched normalized gain score, g, for each of the 1970 students who took both the pre-test and post-test.

In order to assess the level of interactivity in each classroom, the instructors were asked to fill out a short survey (IAI)[3] detailing the frequency with which they used interactive learning strategies (e.g., Think-Pair-Share, Lecture Tutorials, and Ranking Tasks), from which we calculated an IAS for each instructor (Prather *et al.* 2009). This score was a number from 0 to 100% representing the approximate percentage of all possible instructional time each instructor spent using interactive learning strategies. Values for the IAS in this investigation ranged from 0 to 49%, indicating that this instrument was successful at distinguishing differing levels of interactivity in Astro 101 classrooms and that instructors were not inflating estimates of their classes' interactivity. If they had been, we would have expected to see many estimates of over 49% and none near 0%. Nonetheless, we note that the IAS is only an approximate measure of interactivity and provides no insight into the quality of implementation of these strategies.

3. DATA DESCRIPTION—WHO'S TAKING ASTRO 101?

Deming and Hufnagel (2001) reported on the demographic makeup of students taking Astro 101 based on questions provided with the Astronomy Diagnostic Test in an article entitled, "Who's Taking Astro 101?" Table 1 lists the frequencies of responses for each of the 15 questions we asked in our demographic survey. The number of respondents varied for each question for two reasons: 1) the first two questions were included on both forms of the LSCI (see Note 2) so the numbers are higher for those questions and 2) respondents were instructed that all answers were voluntary, so students did not always answer all 15 questions. For comparison, we also list in Table 1 the frequencies of similar questions from the Cooperative Institute Research Program (CIRP) survey entitled "The American Freshman: National Norms for 2008" (Pryor *et al.* 2009; http://

[1] The normalized gain is calculated as $g = (\text{post\%} - \text{pre\%})/(100 - \text{pre\%})$, where pre% and post% are the percent correct for each student on the LSCI before and after instruction respectively. The denominator removes bias introduced by different pre-instruction starting points for each student.

[2] The LSCI used in this study came in two forms, some with only two demographic questions: "What is your gender?" and "Have you previously taken an astronomy course?" others with all 15 demographics questions listed in the Appendix.

[3] The IAI can be downloaded at http://www.csupomona.edu/~alrudolph/professional/Interactivity_Assessment_Instrument.pdf

www.heri.ucla.edu/publications-brp.php). Although this CIRP data set contains only freshmen, it is still useful to compare our data to this snapshot of college students, especially since a significant fraction of the students in our data set are freshmen. Figure 1 shows pie charts comparing our data to the CIRP data set for three key demographic characteristics: gender, ethnic background, and family income (as a proxy for socioeconomic class).

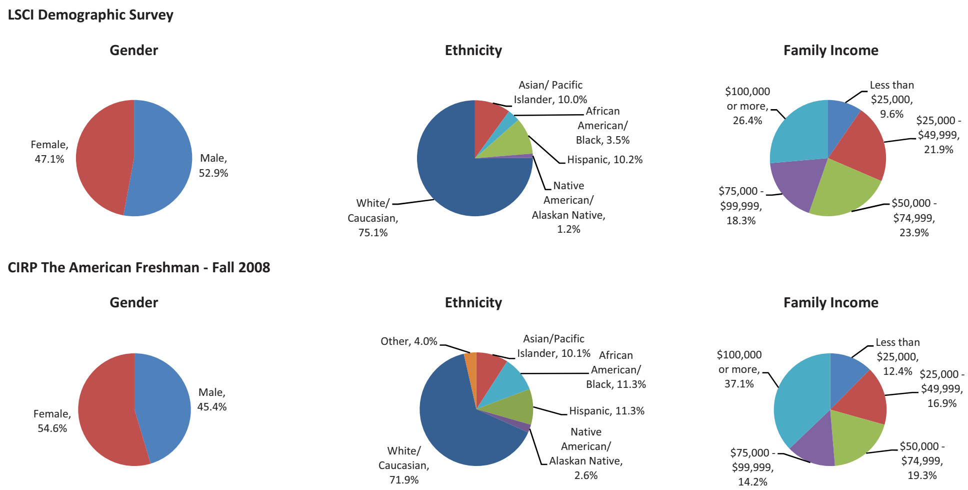

Figure 1. Pie charts showing a comparison of three demographic variables: gender, ethnicity, and parents' income between our national data set of students enrolled in Astro 101 classes (top row) and the national CIRP survey of college freshmen (bottom row).

Comparing all the questions in our data set that overlap with the CIRP data, we conclude that students taking Astro 101 reflect the national college student body as a whole (see Table 1).

Table 1. Demographic profile of study participants compared to national averages

Variables		LSCI Demographic Survey N	LSCI Demographic Survey Percentage	CIRP *The American Freshman* - Fall 2008 Percentage	Notes
Gender					
	Male	1014	52.9%	45.4%	
	Female	904	47.1%	54.6%	
	Total	*1918*			
Previous Course in Astronomy					
	Yes	216	11.3%		
	No	1696	88.7%		
	Total	*1912*			
Native English Speakers					
	Yes	1117	89.7%	91.2%	
	No	128	10.3%	8.8%	
	Total	*1245*			
Ethnicity/Race					
	Asian/Pacific Islander	125	10.0%	10.1%	CIRP's *The American Freshman* survey allows respondents to select more than one ethnicity/race. The LSCI Demographic survey version of the question allowed for only one response. The category "Other" was not available to the respondents of the LSCI Demographic survey.
	African American/Black	43	3.5%	11.3%	
	Hispanic	127	10.2%	11.3%	
	Native American/Alaskan Native	15	1.2%	2.6%	
	White/Caucasian	936	75.1%	71.9%	
	Other		NA	4.0%	
	Total	*1246*			
Mother's Education					
	Some High School	89	7.1%	7.6%	CIRP's *The American Freshman* survey in some cases offered significantly different categories for father's and mother's education. 'Associate's Degree' was not offered as a category. 'Some college' we recategorized as 'High School graduate', 'Postsecondary school other than college' we recategorized as 'Associate's degree', and 'College degree' and 'Some graduate school' we recategorized as 'Bachelor's degree.'
	High School graduate	452	36.2%	35.4%	
	Associate's Degree	211	16.9%	3.6%	
	Bachelor's Degree	339	27.2%	35.2%	
	Graduate Degree	157	12.6%	18.3%	
	Total	*1248*			

Table 1. (*Continued.*)

Father's Education				
Some High School	105	8.5%	9.2%	CIRP's *The American Freshman* survey in some cases offered significantly different categories for father's and mother's education. 'Associate's Degree' was not offered as a category. 'Some college' we recategorized as 'High School graduate', 'Postsecondary school other than college' we recategorized as 'Associate's degree', and 'College degree' and 'Some graduate school' we recategorized as 'Bachelor's degree.'
High School graduate	415	33.4%	34.7%	
Associate's Degree	162	13.1%	3.4%	
Bachelor's Degree	348	28.0%	29.6%	
Graduate Degree	211	17.0%	23.2%	
Total	*1241*			
Family Income				
Less than $25,000	115	9.6%	12.4%	
$25,000 - $49,999	264	21.9%	16.9%	
$50,000 - $74,999	287	23.9%	19.3%	
$75,000 - $99,999	220	18.3%	14.2%	
$100,000 or more	317	26.4%	37.1%	
Total	*1203*			
Elementary School Type				
Public (not charter or magnet)	1011	80.9%		
Public charter or magnet	43	3.4%		
Private religious/parochial	136	10.9%		
Private independent college prep	40	3.2%		
Home school	20	1.6%		
Total	*1250*			
High School Type				
Public (not charter or magnet)	1049	84.3%	77.8%	
Public charter or magnet	50	4.0%	5.3%	
Private religious/parochial	86	6.9%	10.5%	
Private independent college prep	40	3.2%	5.8%	
Home school	19	1.5%	0.6%	
Total	*1244*			
High School Grade Average: LSCI Demographic survey categories (CIRP categories)				
>3.5 (A-/A/A+)	557	45.1%	47.1%	CIRP's *The American Freshman* survey measures high school performance by asking for an average letter grade using the +/- system. The LSCI Demographic survey asks for the high school GPA in the ranges presented.
3.0-3.4 (B/B+)	412	33.4%	41.1%	
2.5-2.9 (B-)	203	16.5%	6.9%	
2.0-2.4 (C/C+)	46	3.7%	4.7%	
<2.0 (D)	16	1.3%	0.1%	
Total	*1234*			
Class Year in College				
Freshmen	457	37.0%	100.0%	
Sophomore	407	33.0%		
Junior	212	17.2%		
Senior	158	12.8%		
Total	*1234*			
Major/Area of interest				
Arts, Humanities, or Social Sciences	277	26.5%	26.0%	CIRP's *The American Freshman* survey is conducted in the fall of the freshman year. The majors reported by students are their anticipated majors.
Science, Engineering, or Architecture	374	35.8%	25.1%	
Education	257	24.6%	8.3%	
Professional (Business, Nursing)	95	9.1%	29.9%	
Other	41	3.9%	11.5%	
Total	*1044*			
College Grade Point Average				
>3.5	379	30.8%		
3.0-3.4	222	18.0%		
2.5-2.9	149	12.1%		
2.0-2.4	295	23.9%		
<2.0	187	15.2%		
Total	*1232*			
Last Math Course Taken				
Algebra	409	33.6%		
Geometry	74	6.1%		
Trigonometry	112	9.2%		
Pre-Calculus	244	20.1%		
Calculus	377	31.0%		
Total	*1216*			
Number of Previous Physical Science Courses				
0	87	7.4%		
1	254	21.6%		
2	312	26.5%		
3	236	20.1%		
4 or more	288	24.5%		
Total	*1177*			
College/University Type				
Research Institution	654	33.2%		CIRP's *The American Freshman* survey only includes 4 year degree granting institutions.
4 Year Masters/Baccalaureate University	737	37.4%		
4 Year Baccalaureate College	88	4.5%		
2 Year College	491	24.9%		
Total	*1970*			

Table 1. (*Continued.*)

Class Size			
	< 25 Students	132	6.7%
	25 - 49 Students	598	30.4%
	50 - 99 Students	380	19.3%
	100+ Students	860	43.7%
	Total	*1970*	
		N	**Mean**
Pre Percent Score		1970	24.5
Post Percent Score		1970	45.7
Interactivity Score		1970	32.3
Normalized Gain		1965	0.275

Thus, students taking Astro 101 are a representative cross section of current college students: men and women, all ethnicities, all socioeconomic backgrounds, and all majors.

Of the demographic categories for which it is possible to compare our data to the CIRP data, "major" is the only category for which there is statistically significant difference. We believe these differences in our results may stem from the wording of our question (see Table A1 in the Appendix), "In what field is your major (or current area of interest if undecided)?" Since the majority of Astro 101 students are early in their college careers (70% freshmen and sophomores) they may have been reporting on an area of topical interest rather than an actual intended area of future study or major. In particular, we find that, in our study, there are a disproportionately large number of students (36%) who responded to the choice of "science, engineering, and architecture." Analysis of a student roster from one class in our study showed that the number of actual declared science majors was closer to 5%, rather than the 35% reported by students for that class. The large "science" response may be due to the fact that the survey was administered in a science class, combined with the ambiguity of the question wording and the difficulty of asking students about their major so early in their college career. In addition, we find that students' responses provided in our study were higher in the category of "education" and lower in the category of "professional" as compared to the CIRP data and previous Astro 101 studies (Deming and Hufnagel 2001). Hence, all discussion of what we term "major" in our study must be understood as self-reported area of interest by each student. Nonetheless, it is interesting that such a large fraction of students in these Astro 101 classes indicated that science was their primary area of interest, in spite of the fact that most of these students are unlikely to pursue studies in science. We note that, in spite of these uncertainties around this question, clearly all majors/areas of interest are well represented in our Astro 101 classes.

The fact that the students taking Astro 101 are representative of college students as a whole, combined with the large numbers of students who take this class across the nation each year, underscores the critically important role that Astro 101 plays in developing scientific thinking and literacy in our college student population and by extension in our society as a whole. These students will become our future lawyers, physicians, business people, politicians, and teachers, and therefore the quality of their scientific education will have a large impact on how science is understood and perceived by the general population. Since Astro 101 is typically the last and only science course many of these students will take (Partridge and Greenstein 2003), it is especially important that we teach this course well. Equally importantly, we must design teaching and learning strategies that work well for all types of students in our classes.

The role of Astro 101 in training future teachers is especially worth noting. Lawrenz, Huffman, and Appeldoorn (2005) found that nearly 40% of students in introductory science classes plan to become licensed teachers. This is consistent with the 25% of students in our data set who chose "Education" as their major/area of interest.[4]

[4]Interestingly, this number is significantly higher than both the national data set (8%) and the Deming and Hufnagel (2001) data set (9%).

The large number of future teachers in our Astro 101 classes means that we are not just teaching future citizens, but we are also preparing the future teachers who will train the next generations of students, including those who will study the STEM[5] disciplines. In many ways, we can think of our Astro 101 courses as semester or quarter long professional development courses for future teachers. Thus, the impact of our teaching of Astro 101 will extend far into the future.

In addition, we have students in our data set with a wide range of abilities (as measured by college GPA) and science and math backgrounds. Thus, it is important that our teaching and learning strategies work well for students with a variety of academic abilities and backgrounds. This important topic is discussed in more detail in Sec. 4.

4. MULTIVARIATE MODELS OF NORMALIZED GAIN

To test the effect of various demographic characteristics on student learning, we constructed a series of multivariate regression models. For each model, we had a large number of independent variables, and one dependent variable: normalized gain. The resulting models indicate the degree to which various independent variables are related to the dependent variable. The results of the multivariate models do not imply that there must be a causal relationship between any given independent variable and normalized gain. However, we can clearly rule out causality in the cases where no statistically significant relationships are found in the data.

In addition to the demographic characteristics, we also include classroom interactivity (based on the IAS) as an independent variable in some models to measure the influence the use of interactive learning strategies has on normalized gain. The table in the Appendix (Table A1) lists all 15 variables representing the 15 questions we asked in the demographic survey, indicating which ones we included in the models, the naming convention and coding we used for the variables, and an explanation of why we excluded the variables we did.

Our first model, which we term "Model 0," contains 12 independent variables (ascribed and achieved characteristics) whose relationship to students' average learning gain is shown in Table 2. The results shown in Table 2 include both unstandardized coefficients, which allow us to interpret relationships in terms of actual changes in normalized gain with changes in each independent variable, and standardized coefficients, which allow us to compare the relative strengths among all the independent variables. Before describing the results of Model 0, we first provide a discussion of variable coding and how unstandardized coefficients are related to average changes in gain for three different variable types (see Table A1 in the Appendix for details of the coding of each variable). For example, the independent variable "Gender" (an ordinal value, i.e., either 0 or 1 in Table A1) was labeled "Male" to indicate that we coded "Male=1" and "Female=0." If the unstandardized coefficient for this variable is statistically significant, this means that the "1" group ("Male") performed significantly better (positive coefficient) or worse (negative coefficient) relative to the "0" group ("Female"). The value of the coefficient represents on average how much more or less gain is achieved by the "Male" group as compared to the "Female" group, holding all other variables constant. In the case where the coefficient is not statistically significant, the average difference in the gain between the two groups is shown to be insignificant and therefore treated as if there is no difference at all between the groups. In the case of the independent variable "Class year" (a ratio value in Table A1), the data is coded simply as "Freshman=1," "Sophomore=2," etc. If the unstandardized coefficient for this variable is statistically significant, the value of the coefficient represents the average change in gain for each additional year in college, e.g., "Freshman" to "Sophomore," "Sophomore" to "Junior," etc., again holding all other variables constant. Finally, for the independent variable "College GPA," we coded each range of GPA as the midpoint of that range (e.g., 3.0–3.5 becomes 3.25). If the unstandardized coefficient for this variable is statistically significant, the value of the coefficient represents the average change in gain for a one point change in GPA, holding all other variables constant.

[5] Science, technology, engineering, and math.

Table 2. Model 0

	Dependent variable = Normalized Gain	
	0	
Independent variable	Coefficients (standard error)	Standardized Coefficients
Constant	-0.057 (0.060)	
Male	0.081** (0.016)	0.160**
White	0.034 (0.020)	0.057
Native English speaker	0.015 (0.029)	0.030
Father with Bachelor's degree or higher	0.010 (0.017)	0.019
Natural log of Family Income	0.002 (0.010)	0.006
Class year	0.014 (0.008)	0.055
College GPA	0.043** (0.011)	0.124**
Sciences, Engineering, or Architecture	-0.019 -0.016	-0.037
Last math class taken	0.035** (0.005)	0.238**
Number of previous physical science courses	0.025** (0.007)	0.124**
Previous Astrophysics course	-0.032 (0.023)	-0.043
Pretest Percent Correct	-0.005** (0.001)	-0.222**
F Value	*15.1***	
N	*910*	
Adjusted R-Square	*0.157*	

***p < .05**
****p < .01**

Our first step in describing the results of Model 0 is to consider the adjusted R-squared of the model. Adjusted R-squared indicates the amount of the total variance in the dependent variable that can be explained by changes in the independent variables used. Model 0 has an adjusted R-squared of 0.157 meaning that this group of 12 independent variables accounts for 15.7% of all the variance in our normalized gain data. Of the ascribed characteristics, only being male has a statistically significant effect on changes in normalized gain with male students achieving an average gain of approximately 9 percentage points more than female students, which is consistent with other findings related to learning and gender in physical science (Hanson 1996; McCullough 2004). Ethnicity, language, father's education, and family income show no relationship to normalized gain in our data. By contrast, nearly all of the achieved characteristics show a statistically significant relationship to increased understanding of LSCI topics in the direction we would expect. Students with greater amounts of education in the form of class year, higher college GPA, highest level of math taken, and a greater number of physical science courses taken tend to have greater improvement in LSCI scores than other students. Surprisingly, students who have previously taken an astronomy course do not outperform their fellow students that are taking their first astronomy course. It may be that a previous astronomy class might simply give a student more factual knowledge yet not help them succeed in the class, whereas previous college experience, or a strong math and science background may give a student additional conceptual and reasoning tools to succeed in the class. Mathematics preparation, in particular, has the strongest positive effect on normalized gain of any of the achieved characteristics, suggesting that facility with math helps, even though the LSCI consists entirely of conceptual questions; that is, no calculations are required to answer LSCI questions correctly.

Pre-test percent correct has a statistically significant negative relationship to normalized gain. This result is not surprising and is caused by the well-known statistical effect of "regression to the mean."[6] The distribution of pre-test scores is well modeled by assuming that students guess on all 26 questions of the LSCI.[7] This assumption leads to a normalized distribution of pre-test scores with a mean and standard deviation of $25 \pm 10\%$, which is a good fit to our actual distribution. On the post-test, students have learned some fraction of the material, but will still guess on the remaining questions. Since the students who *by random chance* did better on the pre-test are not more likely to do better on the parts of the post-test where they are guessing (and vice versa), there is a tendency for the gain of high-pre-test students to be lower, and for the gain of low-pre-test students to be higher (regression to the mean), leading to the negative correlation seen in the model. Thus, this correlation is a statistical artifact that exists in *all data sets* of this type, and we include pre-test percent correct in the model to allow us to control and effectively remove this effect.

As noted above, the characteristic we describe as "major" may represent the actual major of some students, but for others is likely an expression of interest in a subject rather than the subject they will ultimately study. Despite this ambiguity, we were interested in exploring the relationship between these expressed interests and normalized gain. In Model 0, we coded the variable for major/area of interest of students into those that chose "Science, Engineering, and Architecture" (SEA) versus students who chose any other major/area of interest. We found that there was no statistical difference between the gains of these two groups after controlling for all other independent variables. That is, students who chose science, engineering and architecture as their major/area of interest did not achieve a higher average gain than the "nonscience" students. This demonstrates that "nonscience" students and students who self-reported science, engineering, or architecture as their major/area of interest *benefit equally* from the use of interactive learning strategies.

To further probe if any of these self-reported groups differed from the others, we conducted an Analysis of Variance (ANOVA) test of normalized gain by the major/area of interest options provided in the question and found a statistically significant difference ($p < 0.01$) between the mean gains of the majors. The "Arts, Humanities, and Social Sciences" (AHSS) group showed dramatically greater gains on average than the other groups. Accordingly, in our subsequent models, we recoded the "major" variable to focus on this group.

In order to probe more deeply into the relationships between the students' characteristics, interactivity in the classroom, and gain, we created a series of four models shown in Table 3. These models constitute a series in that each subsequent model builds on the previous model. The first model in this series (Model 1) is identical to Model 0 (described above), except that we have coded the major/area of interest independent variable with the students who responded with AHSS as the reference group (see Table A1 in the Appendix). Note that students who chose AHSS as their major/area of interest had, on average, approximately 10 percentage points greater gain than other majors. The adjusted R-squared of this model is 0.185, representing a significant increase over the R-squared from Model 0.

[6] See, for example, http://www.socialresearchmethods.net/kb/regrmean.php, for a more detailed explanation of this effect.

[7] We know from analysis of individual questions that, though students clearly guess on some questions, there are other questions where students do much better than guessing, and others where they do much *worse* than guessing due to attractive distractors in the question. However, the overall effect is a distribution which is the same as if the students guessed on every question.

Table 3. Models 1–4

	Dependent variable = Normalized Gain							
	1		2		3		4	
Independent Variable	Coefficients (standard error)	Standardized Coefficients	Coefficients (standard error)	Standardized Coefficients	Coefficients (standard error)	Standardized Coefficients	Coefficients (standard error)	Standardized Coefficients
Constant	-0.070		-0.235**		-0.266*		-0.208**	
	(0.059)		(0.060)		(0.120)		(0.061)	
Male	0.093**	0.183**	0.087**	0.170**	0.085*	0.167*	0.087**	0.171**
	(0.016)		(0.015)		(0.038)		(0.015)	
White	0.019	0.032	0.012	0.020	0.033	0.055	0.013	0.021
	(0.020)		(0.020)		(0.055)		(0.019)	
Native English speaker	0.019	0.022	0.013	0.015	-0.049	-0.057	0.011	0.013
	(0.029)		(0.028)		(0.080)		(0.028)	
Father with Bachelor's degree or higher	0.008	0.015	0.004	0.008	0.004	0.008	0.005	0.009
	(0.016)		(0.016)		(0.016)		(0.016)	
Natural log of Family Income	0.002	0.008	0.002	0.008	0.002	0.006	0.003	0.008
	(0.010)		(0.009)		(0.009)		(0.009)	
Class year	0.018*	0.071*	0.024**	0.092**	0.024**	0.093**	0.024**	0.093**
	(0.008)		(0.008)		(0.008)		(0.008)	
College GPA	0.036**	0.106**	0.037**	0.109**	0.067**	0.197**	0.036**	0.106**
	(0.010)		(0.010)		(0.026)		(0.010)	
Arts, Humanities, or Social Science	0.101**	0.176**	0.104**	0.181**	0.010	0.018	0.023	0.040
	(0.018)		(0.017)		(0.042)		(0.041)	
Last math class taken	0.031**	0.214**	0.034**	0.230**	0.040**	0.274**	0.034**	0.229**
	(0.005)		(0.005)		-0.011		(0.005)	
Number of previous physical science courses	0.024**	0.120**	0.024**	0.120**	0.021	0.105	0.023**	0.119**
	(0.006)		(0.006)		(0.015)		(0.006)	
Previous Astrophysics course	-0.029	-0.039	-0.028	-0.039	-0.031	-0.042	-0.030	-0.041
	(0.022)		(0.022)		(0.022)		(0.022)	
Pretest Percent Correct	-0.005**	-0.224**	-0.005**	-0.213**	-0.005**	-0.213**	-0.005**	-0.212**
	(0.001)		(0.001)		(0.001)		(0.001)	
Interactivity Score			0.0051**	0.258**	0.0062	0.314	0.0043**	0.217**
			(0.0006)		(0.0037)		(0.0007)	
Cross term: Interactivity score X Arts, Humanities, or Social Science					0.0032*	0.183*	0.0027*	0.158*
					(0.0013)		(0.0013)	
Cross term: Interactivity score X Male					0.0001	0.004		
					(0.0012)			
Cross term: Interactivity score X White					-0.0006	-0.044		
					(0.0018)			
Cross term: Interactivity score X Native English speaker					0.0022	0.129		
					(0.0027)			
Cross term: Interactivity score X College GPA					-0.0010	-0.182		
					(0.0008)			
Cross term: Interactivity score X Last math class taken					-0.0002	-0.057		
					(0.0004)			
Cross term: Interactivity score X Number of previous physical science courses					0.0001	0.016		
					(0.0005)			
F Value	*18.2***		*24.3***		*16.2***		*23.0***	
N	*910*		*910*		*910*		*910*	
Adjusted R-Square	*0.185*		*0.250*		*0.250*		*0.253*	

***p < .05**
****p < .01**

Model 2 adds each classroom's IAS to the analysis, as a measure of interactivity in each student's classroom. All the statistically significant variables in Model 1 remain statistically significant in the same direction in Model 2 with only marginal changes in the coefficients. However, we find that the added variable of interactivity shows a statistically significant and positive relationship to normalized gain. Looking at the unstandardized coefficient, we see that for every 10 percentage point increase in IAS there is, on average, a 5.0 percentage point increase in normalized gain. Taken to the extremes of the range of IAS, the model indicates that students in classes with a 45% IAS will, on average, outperform students in classes with a 5% IAS by *more than 20 percentage points in gain*, after controlling for the effects of the other variables in the model.

Since unstandardized coefficients are each measured in the units of the variable, they cannot be compared to each other directly. Standardized coefficients are measured in units of standard deviations, allowing direct comparison between variables: the larger the standardized coefficient, the greater the effect of a variable on normalized gain. By looking at the standardized coefficients in Model 2, we can see that *interactivity has the greatest effect on normalized gain of all of the independent variables in our model.* The change in adjusted R-squared from Model 1 to Model 2 indicates that the addition of interactivity to our model increases the model's explanatory power from 18.5 to 25.0% of the variance in normalized gain. Thus, while 12 variables were required to explain the first 18.5% of variance in gain, the addition of a single variable, interactivity, raised the R-squared by over 35%.[8]Such a large change in R-squared from the addition of a single variable is a striking result. This result confirms what we found in the first paper in the series (Prather *et al.* 2009), namely that *using interactive learning strategies in the Astro 101 classroom can have a strong positive impact on student learning,* even after controlling for all other variables.

In Model 3, we introduce seven cross (interaction) terms to test if the effects of classroom interactivity differ based on other key independent variables. These cross terms measure whether the increased learning gain attributed to interactivity is *different* for different members of the population as measured by the cross variable. For example, *if* the cross term "interactivity × Male" were statistically significant, that would suggest that though both groups (men and women) benefit from interactivity, one benefits more than the other; in fact, that particular cross term is *not* statistically significant, meaning that men and women benefit *equally* from interactive learning strategies.

Of the seven cross terms introduced, all but one are *not* statistically significant, meaning *the effects of interactivity on normalized gain are the same for males and females, whites and nonwhites, native English speakers and non-native English speakers, and regardless of college GPA, mathematical preparation, and the number of physical science courses previously taken*. These last two results show that even though most of the interactive learning strategies employed in Astro 101 classrooms are focused on conceptual learning, rather than calculation, students with a strong math and science background benefit from these activities as much as other students. It is also important to note that many instructors mistakenly believe that stronger students do not benefit from interactive learning strategies, a belief contradicted by our results, which show that students of *all* abilities benefit equally from these strategies. Together with the results of Paper I, which showed that normalized gain did not depend on institution type or class size, these results indicate that interactive learning strategies can be successfully applied to *any* Astro 101 classroom and will work equally well for nearly *all* the students in those classrooms.

The single cross term in Model 3 that is statistically significant is the one including academic major/area of interest AHSS, meaning that students who chose AHSS as their major/area of interest receive greater benefit from an increase in IAS than other majors, though *all* majors benefit. Model 4 includes only the "interactivity × AHSS" cross term; all other cross terms were removed because they were not statistically significant in Model 3. With only a single cross term, it is possible to have a relatively straightforward interpretation of the interactivity variable and the "interactivity × AHSS" cross term together. In Model 2, the coefficient of interactivity contained the average effect of interactivity on all majors combined. In Model 4, this coefficient has effectively been split in two: when the "AHSS major variable=0," the cross-term drops out; hence, the interactivity coefficient now measures the effect of interactivity *only* on the *non*-AHSS students, while the "interactivity × AHSS" term measures the effect of interactivity on the AHSS students *over and above the other students*. Hence, when the "AHSS major variable=1," the two coefficients added together measure the *total* effect of interactivity on the AHSS students. In Model 4, the coefficient for interactivity is still statistically significant, indicating that the non-AHSS students increase their normalized gain on average by 4.3 percentage points for every 10 percentage point increase in IAS. The coefficient for the "interactivity × AHSS" cross term measures the *additional* gain that AHSS students achieve as a result of a higher IAS. On average, for every 10 percentage point increase in IAS, AHSS students gain an *additional* 2.7 percentage points in normalized gain over non-AHSS students. Adding the coefficients for interactivity and the "interactivity × AHSS" cross term gives us the complete effect of interactivity for AHSS students. Taken together, on average AHSS students

[8]The R-squared value increased from 0.185 to 0.250, an increase of $100 \cdot (0.250-0.185)/0.185=35\%$.

achieve a total increase of 7.0 percentage points (4.3+2.7) in normalized gain for every 10 percentage point increase in IAS. Figure 2 shows the effect of IAS for students who chose the AHSS major/area of interest option (blue line) and all other students (red line), as well as indicating the overall effect of interactivity for all majors combined (black dotted line), from Model 2. Clearly, though all students benefit from higher interactivity, those who chose the AHSS major/area of interest benefit disproportionately. It is possible that these students are more comfortable with interactive learning strategies, and therefore benefit more from them.

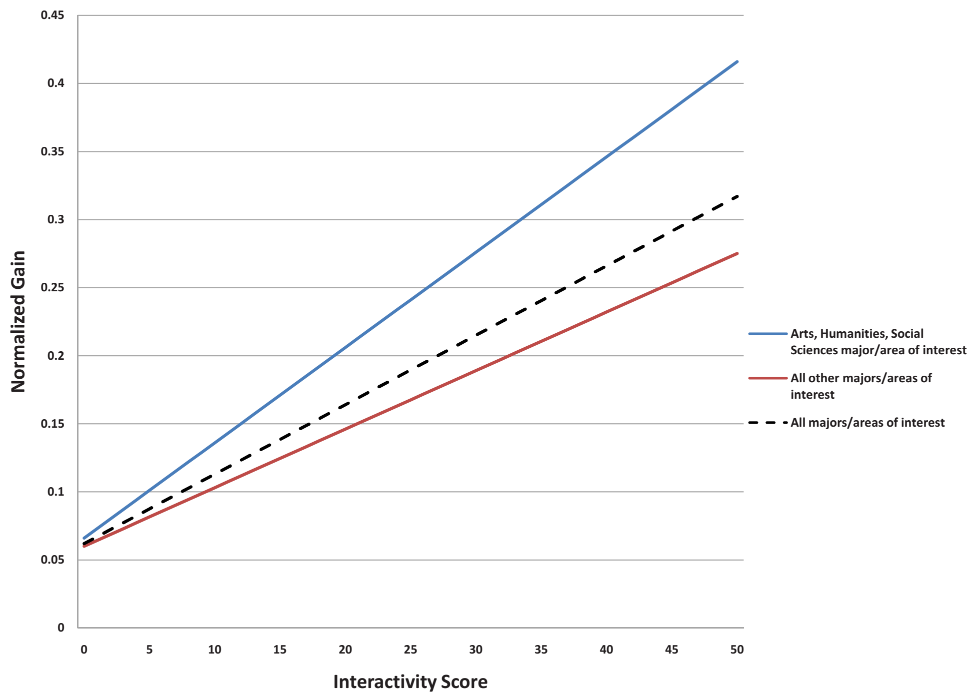

Figure 2. Graph of multivariate model results showing predicted gain by major (after controlling for all other variables) as a function of IAS for various students enrolled in Astro 101: all students (dashed black line); those who chose "Arts, Humanities, and Social Sciences" as their major/area of interest (blue line); and those who chose any other major/area of interest: "Science, Engineering, and Architecture," "Business or Professional," "Education," or "Other" (red line).

5. CONCLUSIONS

Our main conclusions are as follows.

1) The students in our study, who are enrolled in Astro 101 nationwide, are a representative cross section of the college student population as a whole. Thus, when we teach Astro 101, we are affecting the scientific literacy of all types of college students: men and women, native and non-native English speakers, all ethnicities, all majors, and students of all academic abilities. For many of these students, this is the last science course they will ever take.
2) 25% of the students in our sample of Astro 101 students are declared education majors or have expressed an interest in the study of education, much higher than the 8% reported in the CIRP national data set and the 9% previously reported by Deming and Hufnagel (2001). When we teach Astro 101, we are not only developing the scientific literacy of our future citizens, we are also training the future teachers of the next generation of students, including those who may choose careers in STEM disciplines.
3) A multivariate regression model of normalized gain with both ascribed and achieved student characteristics showed that students with a stronger academic background (more years in college, more math and science background), not surprisingly, had higher normalized gains, on average. However, none of the ascribed characteristics, other than gender, showed any statistically significant correlation with gain.
4) Adding the level of classroom interactivity (IAS) to the model increased the R-squared from 0.185 to 0.250, a very large (35%) increase in predictive power for a single variable. IAS was the variable with the highest standardized coefficient of any of the 13 independent variables in the model, indicating that *the use of interactive learning strategies has a stronger positive effect on student learning in Astro 101 classrooms than any other characteristic we measured*.
5) Adding cross terms between interactivity and other independent variables showed that *interactive*

learning strategies equally benefit men and women, students of all ethnicities, native and non-native English speakers, as well as students of all levels of academic ability, mathematical preparation, and previous physical science coursework.

6) Interactive learning strategies benefited students representing all majors/areas of interest. However, students who chose AHSS as their major/area of interest benefited by the greatest amount.

The results of both this and our prior study (Prather *et al.* 2009), taken together, emphasize that interactive learning strategies are capable of helping *every* student in our classroom. The differences we see in learning gains between classrooms are *not* due to the type of institution, *or* to the size of the class, *or* to the individual characteristics of the students in the class. This is a critical finding in our efforts to understand how we can help our students achieve the highest level of understanding possible in our classrooms. We have all experienced how much the learning can differ between individual students in our classes, and Prather *et al.* (2009) have shown how students' learning gains can differ widely between Astro 101 classrooms. We have identified the use of interactive learning strategies as a key factor that can help students learn in our classrooms, yet there is a great deal of spread in learning gains that remains to be understood. Clearly, not every instructor who uses such strategies succeeds in helping their students achieve high learning gains (Prather *et al.* 2009). Since we have ruled out the type of institution, the size of the class, and the particular students in our classes, we must look elsewhere for the answer. We believe that the instructor's effective implementation of interactive learning strategies is the crucial factor that allows some classes to achieve higher gains in student understanding than other classes, even when the same level of interactivity exists in both classes. Professional development for instructors is the best way to close this gap (Prather, Rudolph, and Brissenden 2009).

Why is professional development so critical to the improvement of teaching and learning in Astro 101 classes? The system we are investigating, the Astro 101 classroom, and its associated environment of curricular and professional development, is very complex. It involves the interplay between the instructor, the students, the subject matter, the classroom environment and the instructional strategies we employ. The depth of an instructor's understanding of each of these variables, and how they are related to each other, is referred to as Pedagogical Content Knowledge (PCK), and an instructor's understanding of their own PCK is directly related to their ability to *effectively* implement the research validated instructional strategies that have been shown to help students improve their conceptual understanding (Gess-Newsome and Lederman 1999; Prather and Brissenden 2008).

We know that well-crafted professional development opportunities, based on best practices, can be effective in improving an instructor's PCK (Gess-Newsome and Lederman 1999; Loucks-Horsley *et al.* 2003). In particular, an environment of peer review, in which participants offer suggestions and critiques of each other's implementation of interactive learning strategies, termed *Situated Apprenticeship*, helps instructors go beyond a merely intellectual awareness of which learning strategies are most effective in changing students' gains in understanding *to a real change in their practice* (Prather and Brissenden 2008). Hence, well-designed and well-executed professional development for Astro 101 instructors (future and current) is crucial if we expect to see real, sustained improvement in how much students benefit from this class.

The majority of Astro 101 instructors receive no formal training before stepping into the classroom to teach for the first time, and we applaud those instructors who voluntarily seek out their own professional development. However, we believe a lasting national impact will only come from the astronomical community adopting a more comprehensive commitment to invest in professional development. Our national societies (AAS, ASP, AIP, APS, AAPT, AGU, etc.), administrators, department chairs, and senior colleagues need to encourage or perhaps require, as well as provide resources for, professional development for all instructors. This is especially needed for new instructors and future instructors such as graduate students and postdocs. If we can reach these young colleagues early in their careers, and change the way they approach teaching and learning, we might see a real change in the quality of how Astro 101 is taught in the future.

Acknowledgments

We would like to thank all the members of the CAE community who cared enough about their students' learning to participate in this study and the thousands of Astro 101 students who freely gave of their time to participate. Your conscientious efforts, we believe, will help us all be better Astro 101 instructors. We could never have done this without you. More information about CAE can be found online at: http://astronomy101.jpl.nasa.gov/. This research was generously supported by Michael Greene and the Exoplanet Exploration Public Engagement Program (NASA JPL), Michelle Thaller, and the Spitzer Education and Public Outreach Program (CalTech), and by the National Science Foundation under Grants No. 0715517, a CCLI

Phase III Grant, and Grant No. AST-0847170, a PAARE Grant. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Appendix

Table A1. Demographic Survey Questions and Recodes

Survey Question	Variable	Survey Values	Recoded Values	Measurement Level	Notes
What is your gender?	Male	Male Female	1 0	Ordinal	
Have you previously taken an astronomy course? (Do NOT count this course in your response.)	Previous Astrophysics course	Yes No	1 0	Ordinal	
Is English your native language?	Native English speaker	Yes No	1 0	Ordinal	
What best describes your ethnic background (choose only one):	White	White/Caucasian Asian or Pacific Islander African American/Black Hispanic Native American (including Alaskan Native)	1 0 0 0 0	Ordinal	Over 75% of the respondents indicated "White/Caucasian" as their ethnicity. The relatively low number of respondents for some of the remaining groups prevented a deeper analysis of ethnicity.
What is the highest educational level attained by your mother?	N.A.	Some high school High school graduate Associates degree (2-year) Bachelor's degree (4-year) Graduate degree (e.g., MA, MS, MD, JD, PhD)			Mother's educational attainment is highly correlated with father's educational attainment and therefore was excluded from the models.
What is the highest educational level attained by your father?	Father with Bachelor's degree or higher	Some high school High school graduate Associates degree (2-year) Bachelor's degree (4-year) Graduate degree (e.g., MA, MS, MD, JD, PhD)	0 0 0 1 1	Ordinal	Father's education is recoded into two groups. There is no statistically significant difference in normalized gain within each of these groupings but there is a statistically significant difference in normalized gain between the two groups.
What is your best estimate of your parents' total income last year? Consider income from all sources before taxes.	Natural log of Family Income	Less than $25,000 $25,000-49,999 $50,000-74,999 $75,000-99,999 $100,000 or more	2.53 3.62 4.14 4.47 5.39	Ratio	The natural log of the midpoint of the income range is used. Logging the midpoint helps adjust for the right skewed nature of income. The midpoint of the final category ("$100,000 or more") was estimated using a Pareto interpolation.
What type of elementary school did you attend? (Mark one)	N.A.	Public school (not charter or magnet) Public charter or magnet school Private religious/parochial school Private independent school Home school			There was no correlation of this variable with normalized gain in our preliminary analysis and therefore we excluded it from the models.
From what type of high school did you graduate?	N.A.	Public school (not charter or magnet) Public charter or magnet school Private religious/parochial school Private independent college-prep school Home school			There was no correlation of this variable with normalized gain in our preliminary analysis and therefore we excluded it from the models.
What was your high school GPA?	N.A.	Above 3.5 3.0-3.4 2.5-2.9 2.0-2.4 Below 2.0			We chose college GPA as the measure of student ability since it is a more direct measure of student ability in the setting we are studying.
What is your class level in college?	Class Year	Freshman Sophomore Junior Senior	1 2 3 4	Ratio	Class year is recoded into a measure of college experience.
In what field is your major (or current area of interest if undecided)?	Science, Engineering, or Architecture	Arts, Humanities, or Social Sciences Science, Engineering, or Architecture Education Professional (e.g., Business, Nursing, etc.) Other	0 1 0 0 0	Ordinal	Used to test normalized gain differences between individuals choosing "Science, Engineering, or Architecture" and other individuals.
In what field is your major (or current area of interest if undecided)?	Arts, Humanities, or Social Science	Arts, Humanities, or Social Sciences Science, Engineering, or Architecture Education Professional (e.g., Business, Nursing, etc.) Other	1 0 0 0 0	Ordinal	Used to test normalized gain differences between individuals choosing "Arts, Humanities, or Social Science major" and other individuals.
What is your college GPA (if you have already completed at least one term)?	College GPA	Above 3.5 3.0-3.4 2.5-2.9 2.0-2.4 Below 2.0	3.75 3.25 2.75 2.25 1.75	Ratio	Respondents are recoded into the midpoint of the category range.

Table A1. (*Continued.*)

Survey Question	Variable	Survey Values	Recoded Values	Measurement Level	Notes
What was the last math class you completed prior to taking this course?	Last math class taken	Algebra	1	Interval	The math courses are recoded into the most common progression.
		Geometry	2		
		Trigonometry	3		
		Pre-calculus	4		
		Calculus	5		
How many physical science classes (e.g., astronomy, physics, chemistry) did you take in high school or college prior to this course?	Number of previous physical science courses	0	0	Ratio	
		1	1		
		2	2		
		3	3		
		More than 3	4		

References

Bardar, E., Prather, E., Brecher, K., and Slater, T. 2006, "Development and Validation of the Light and Spectroscopy Concept Inventory," *Astronomy Education Review*, 5, 103.

Bardar, E. M., Prather, E. E., Brecher, K., and Slater, T. F. 2005, "The Need for a Light and Spectroscopy Concept Inventory for Assessing Innovations in Introductory Astronomy Survey Courses," *Astronomy Education Review*, 4, 20.

Crouch, C. H. and Mazur, E. 2001, "Peer Instruction: Ten Years of Experience and Results," *American Journal of Physics*, 69, 970.

Deming, G. and Hufnagel, B. 2001, "Who's Taking ASTRO 101?" *Physics Teacher*, 39, 368.

Fraknoi, A. 2001, "Enrollments in Astronomy 101 Courses: An Update," *Astronomy Education Review*, 1, 121.

Gess-Newsome, J. and Lederman, N. G. 1999, *Examining Pedagogical Content Knowledge: The Construct and Its Implications for Science Education*, Norwell, MA: Kluwer Academic.

Hake, 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.

Hanson, S. L. 1996, *Lost Talent: Women in the Sciences*, Philadelphia, PA: Temple University Press.

Hudgins, D. W., Prather, E. E., Grayson, D. J., and Smits, D. P. 2006, "Effectiveness of Collaborative Ranking Tasks on Student Understanding of Key Astronomy Concepts," *Astronomy Education Review*, 5, 1.

Lawrenz, F., Huffman, D., and Appeldoorn, K. 2005, "Enhancing the Instructional Environment: Optimal Learning in Introductory Science Classes," *Journal of College Science Teaching*, 34, 40.

Loucks-Horsley, S., Love, N., Stiles, K. E., Mundry, S., and Hewson, P. W. 2003, *Designing Professional Development for Teachers of Science & Mathematics*, 2nd ed., Thousand Oaks, CA: Corwin Press.

McCullough, L. 2004, "Gender, Context, and Physics Assessment," *Journal of International Women's Studies*, 5, 20.

Partridge, B. and Greenstein, G. 2003, "Goals for 'Astro 101': Report on Workshops for Department Leaders," *Astronomy Education Review*, 2, 46.

Prather, E. E. and Brissenden, G. 2008, "Development and Application of a Situated Apprenticeship Approach to Professional Development of Astronomy Instructors," *Astronomy Education Review*, 7, 1.

Prather, E. E., Rudolph, A. L., and Brissenden, G. 2009, "Teaching and Learning Astronomy in the 21st Century," *Physics Today*, 62, 41.

Prather, E. E., Rudolph, A. L., Brissenden, G., and Schlingman, W. M. 2009, "A national study assessing the teaching and learning of introductory astronomy. Part I. The effect of interactive instruction," *American Journal of Physics*, 77, 320.

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

Pryor, J. H., Hurtado, S., DeAngelo, L., Sharkness, J., Romero, L. C., Korn, W. S., and Tran, S. 2009, *The American Freshman: National Norms for Fall 2008*, Los Angeles: Higher Education Research Institute, UCLA.

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

Zeilik, M. and Morris-Dueer, V. J. 2004, "What are essential concepts in 'Astronomy 101'? A new approach to find consensus from two different samples of instructors," *Astronomy Education Review*, 3, 61.

ÆR
010107-1–010107-15