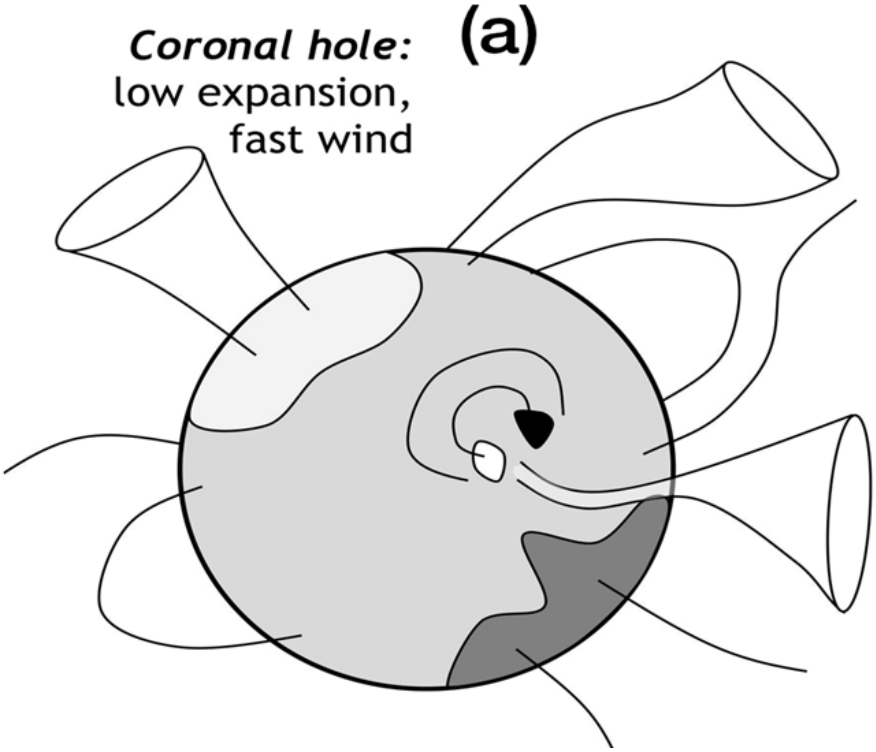
3

April 2015
Volume 1 Number 3
ISSN 2372-8388

Journal and Review of Astronomy Education and Outreach

A Scholarly Publication of Research, News and Commentary



In This Issue: TEMPEST, Validations, and 3-D Images

Journal and Review of Astronomy Education and Outreach (*JRAEO*)

Editor: Dr. Lawrence Krumenaker, Journalist in Residence, HITS, Germany, editor@jraeo.com

Assistant Editor: Dr. Kristine Larsen, Professor, Central Connecticut State University, assteditor@jraeo.com

Staff Statistician: Dr. Seock-Ho Kim, Professor, University of Georgia, shkim@uga.edu

JRAEO Editorial / Advisory Board

JRAEO Scope

Articles on any topic relating to educating students or the general public about astronomy are of interest. This can include areas of study often designated as astronomy education or space science education, as well as topics from allied fields such as physics, geoscience, math or history education, as long as the topic is found in an astronomy course or used to excite and inform people about astronomy.

The subject areas include all grade levels–primary, elementary, middle school, and secondary level (in the US, the K-12 system)–undergraduate education (both astronomy majors and general education courses or other domains where astronomy can be included, such as exobiology or cosmochemistry), teacher education, graduate education, and informal or free-choice education involving the general public, including studies of the media (newspapers, social media, etc.) and dedicated facilities such as science centers and planetariums.

Contact

Address: 3605 Sandy Plains Road, Suite 240-203, Marietta, GA 30066

Main phone: +1-404-702-8147 Fax: +1-404-393-7037 Website: www.jraeo.com

Journal and Review of Astronomy Education and Outreach

April 2015 Issue 3
Volume 1 Number 3

Section A – Research

Section B – General Articles

Hermograph Press LLC, Marietta, GA 30066 USA

www.jraeo.com

The *Journal and Review of Astronomy Education and Outreach (JRAEO)* [ISSN 2372-8388, online only] is published up to three times a year, as one volume of three issues. Beginning in 2015 it will be published only during the academic year, without any summer issue. It is published by Hermograph Press LLC (the LLC), 3605 Sandy Plains Road, Suite 240-203, Marietta, GA 30066. http://www.hermograph.com

Subscription price: $45.00 per year (three consecutively numbered issues). Prices are exclusive of tax. The price includes a reduction in page charge/open access. As of January 2015, the current issue is available only to subscribers but then all past issues are free to all readers. Note that articles are accepted from any person, but non-subscribers have higher page fees. Subscribers prior to January will still have no page charges until their subscription expires.

Subscriptions may be ordered and paid for online, via fax, or through postal mail with a check in US dollars drawn on a US bank, by wire transfer or international Money Orders.

Advertising Sales: Inquiries concerning advertising should be forwarded to the Publisher, at the contact information above. Advertising will be accepted as long as it moves forward the astronomy education mission of the publication and does not compete against any specific Hermograph Press product. Acceptance of ads does not constitute an endorsement of the product by the journal or the LLC; placement of the advertising will not be allowed to dictate any editorial control by the advertiser.

Advertising may include professional opportunities.

The LLC reserves the right to refuse any advertisement for any reason.

On the Web: www.jraeo.com (journal)
www.hermograph.com (the LLC)
On Twitter: @JnlRevAstroEdO
On Facebook: Hermograph Press

Larry Krumenaker
Editor

When last we published, Hermograph and *JRAEO* were in the early stages of being uprooted from its nominal Georgia home base. Now, as a digital publication, Home is where the Ethernet cable gets plugged in. Now it is on temporary assignment in Heidelberg, Germany, where yours truly is ensconced as a Journalist in Residence at the Heidelberg Institute for Theoretical Sciences (HITS). It is a really nice way to be immersed in science; freedom to roam the institute, talk to anybody about their work, no deadline pressures, take a jaunt over to another institute – it is said there are more astronomers per square kilometer here than in any other part of Europe – and get way over-carbed on German lunches.

I do have to wonder about astronomy education here; is it most abundant here rather than anywhere else in Germany or Europe? I intend to find out. Certainly there is even some astronomy education in Heidelberg, and in HITS. See the current issue of *JRAEO*'s sister publication, *The Classroom Astronomer*, for examples of astronomy education here.

But of concern to *JRAEO*'s audience, I have not yet found out how much astronomy education research is going on here. Astronomy is in many of the curriculae of the German states (Heidelberg is in Baden-Württemberg). Does that mean that there is also research going? I don't know. But Kristine Larsen and I here at *JRAEO* do wish to make this less an American journal and more a global one, and I intend to find out here if we can do that.

I'd like to take this opportunity to remind all of you that we encourage the submission of articles on any aspect of astronomy education or outreach from anyone involved with these fields, as well as those who work with astronomical content in formal and informal science education more broadly. We also want to more of our colleagues to volunteer to become peer reviewers. It is a very rewarding professional experience. Contact either Kristine or myself at jraeo@toteachthestars.net if either of the above are something that you fit.

Sincerely,

Lawrence E. Krumenaker

Kristine Larsen
Assistant Editor

Besides our Board Members and our Staff Statistician, we could not get by WITHOUT our pool of reviewers. While we are not publishing a list of their names (in the spirit of double-blind peer review), we very enthusiastically thank each and every one of them for their timely work and attention to detail, and hope to keep them occupied with additional interesting research to review in the near future. (And, you get a free copy of the issue in which you reviewed an article; what a perk!) Both Larry and I have done our share of peer reviewing for journals, and understand the time and effort that goes into a thoughtful manuscript review. This journal could not exist without your hard work.

For the rest of you, I would like to reiterate what Larry said in his column - reviewing is an interesting and worthwhile experience, and for those of you in academia, remember that peer reviewing counts towards promotion and tenure in most places (usually under professional activity – consult your local contract for specifics). Those of you who would be interested in doing peer review should contact either him or me, or get the application form on the website (jraeo.com) and send it to us in email. We look forward to you joining the *JRAEO* team.

Collegially yours,

J Rev Astron Educ Outreach Vol. 1 No. 3 pp. A3-30 (2015)

A DIRECT EXAMINATION OF COLLEGE STUDENT MISCONCEPTIONS IN ASTRONOMY.
II. VALIDITY OF THE ASTRONOMY BELIEFS INVENTORY

Andrej Favia[1], Neil F. Comins[1], and Geoffrey L. Thorpe[2], *University of Maine*

Received January 22, 2015; Revised March 12, 2015; Accepted April 10, 2015

Abstract: This is the second in a series of papers in which we examine the persistence of 215 common misconceptions using a new instrument, the Astronomy Beliefs Inventory. In the first paper, we showed that the data in the ABI meet the statistical criteria for an internal validity analysis. In this paper, we examine the concurrent validity of the ABI by comparing student responses from the instrument with data from four instructional instruments, including multiple-choice exams and free-response questions, administered in their introductory astronomy course during the Fall 2013 semester. Two of the instruments were administered prior to instruction on the associated topics. We show that the validity of the ABI is supported through significant correlations with all four of the instruments under consideration.

Keywords: Students - Non-science Majors - General Astronomy - Astro 101 - Misconceptions - Undergraduate Education

INTRODUCTION

Misconceptions about astronomy are acquired at all ages through a variety of incorrect thought processes and incorrect information provided by the media and other sources (e.g., Comins, 2001). Children, who lack background information and experience thinking logically and scientifically, are especially at risk of accepting misconceptions as fact. In their study, Smith III, diSessa, and Roschelle (1993) observed that misconceptions persist even after instruction, and that misconceptions “have a strong influence on how student learning is currently evaluated." That is to say, misconceptions (which comprise deeply held incorrect beliefs) interfere with learning. The persistence of misconceptions about astronomy are so widespread that many researchers are still publishing studies that recognize the prevalence of misconceptions and their influences on student learning (Bailey & Slater, 2004; Bailey, Prather, Johnson, & Slater, 2009; Comins, 2001; Lombardi, Sinatrab, & Nussbaum, 2013, and references therein).

Various researchers tend to interpret misconceptions from the standpoint of how they think students learn. For example, Vosniadou (1994) suggests that learning involves the strategic refinement of one's own mental models. In their study, Bailey et al. (2009) refer to “the constructivist movement of the late 20th century" (p. 2), a paradigm in which effective instruction requires building upon pre-instructional student ideas. Teaching from a constructivist framework supports the use of inquiry-based tutorials and interactive clicker questions, with class-wide emphasis on group interactions (Slater & Adams, 2002). As Smith III et al. (1993) state, learning “involves the acquisition of expert concepts and the dispelling of misconceptions” (p. 122). Bransford (2000) emphasizes that the constructive process may enhance the students’ post instruction understanding of astronomical concepts. Constructivism

---

[1] Department of Physics and Astronomy
[2] Department of Psychology

thus takes place at the intersection of what students may already know and how students can construct more sophisticated models.

Multiple researchers have alternate, often loosely-worded, definitions of the term "misconception." As examples, Sadler (1998) defines misconceptions as "conceptions that are quite different from their teachers' or those of scientists." Vosniadou (1994) defines misconceptions as "attempts to reconcile scientific information with false information regarding how one observes the world." Comins (2001) presents a list of definitions for the term "misconception," which he developed after asking researchers at a conference on misconceptions in science and mathematics in 1994. Comins presents a list of thirteen such definitions (p. 55), which include: "any belief that is untrue," "a mistaken idea," "an incorrect mental construct," "a misunderstanding," "only incorrect beliefs that are deep-seated in our minds," "a conception or belief that produces a systematic pattern of errors," and others. For our study, we adopt the definition of misconception used by Comins: "any deeply held belief that is inconsistent with currently accepted scientific concepts. Deeply held beliefs are distinctly different from superficial ones, such as details we might memorize for an exam and then promptly forget" (p. 56). This definition is consistent with the premise that misconceptions are deep seated and avoids grounding the term in a particular learning framework. The interpretation of misconceptions being *deep-seated* is also supported by the results of other studies on misconceptions in astronomy (Sadler, 1998; Bailey et al., 2009).

The analysis of the data for this paper concerns the college-level astronomy course for non-majors, as taught at the University of Maine during the Fall 2013 semester. The class is taught in the style of a modified lecture. Common misinformation held by students about a relevant topic is presented in class, after which an explanation is provided to demonstrate why this information is, in fact, wrong. Students are also asked to respond to misconception-based, free-response "attendance questions" on a topic to be addressed in the subsequent class. For example, if the question is "How many zodiac constellations are there?" then the subsequent lecture would include a discussion about zodiac constellations. The answer is 13, not 12, as is often commonly thought by students in the course, in part because of the common association of "12 months for the 12 zodiac constellations." The subsequent lecture would have extra instruction on (i) the duration of each zodiac, which is not 30 days, and (ii) why there are *not* 12 zodiac constellations. Extra credit is also awarded to students for identifying their own misconceptions, as well as for identifying misinformation presented in other current sources (e.g., in a television show). We note that the course, as taught during the Fall 2013 semester, did not utilize inquiry-based tutorials. About half of the students in the course are non-science majors, and given the discussion by Clark et al. (2012), perhaps only the very top students in the class would have benefited more in learning the same scope of topics taught in the class from a constructivist-oriented learning environment.

We introduced the Astronomy Beliefs Inventory (ABI) in 2014 as a new instrument to explore the persistence of misconceptions (Favia, Comins, Thorpe, & Batuski, 2014, hereafter, Paper I). The ABI is a comprehensive inventory of 215 misconceptions held by college students taking an introductory-level course in astronomy. According to Strike and Posner (1992), conceptual change requires that the learner become dissatisfied with current concepts, and that learning occurs when a new conception seems to solve the current problem and encourages new ways of looking at natural phenomena. In other words, students have to be presented with strong evidence to contradict their preconceived notion, and then believe the new information. Because the ABI asks students to report what they believe for a variety of topics, the ABI can cover a broad range of misconceptions very efficiently.

The ABI is unique in that it avoids the conventional pre/post-test approach in learning assessment by including a retrospective component. The design of retrospective studies, however, is subject to some issues regarding reliability. Briefly, memory is a reconstructive process (Olson & Cal, 1984). As Henry, Moffitt, Caspi, Langley, and Silva (1994) note, a retrospective approach may be of questionable validity in some contexts, notably in the recall of personally-significant emotional and psychosocial material. The authors suggest that "the use of retrospective reports should be limited to testing hypotheses about the relative standing of individuals in a distribution" (p. 92). A comprehensive review by Brewin, Andrews, and Gotlib (1993), however, "suggests that claims concerning the general unreliability of retrospective reports are exaggerated" (p. 82). The likelihood of inaccurate responses may be significantly reduced by asking subjects to provide reports on a timeline for abandoning misconceptions. Additionally, the ABI consists of very brief statements (e.g., "Saturn's rings are solid") each of which enhances the recall of one's own past beliefs, compared to longer statements with the same content. These latter versions are more likely to invoke additional memories that might obfuscate the recollection.

The ABI probes whether or not students believe the presented misconceptions. As with all tests, responses are affected by how accurately students fill in the correct bubbles on bubble sheets and how they interpret the statements in the inventory. About 10 departmental faculty with training in physics education research evaluated a pool of 235 items, and after receiving feedback, 20 items were discarded due to issues with clarity in the wording, leaving the 215 statements that we chose to constitute the ABI. In Paper I, we showed that students provided the correct bubble and statement interpretation over 90% of the time for a random subset of five of the ABI statements. Compared to traditional surveys, the ABI has the additional aspect of asking students when (before or during college) they learned a misconception. The ABI measures what students believe after instruction in class (anywhere between a week and several months, depending on the topic). The administration of a delayed post-test has been employed elsewhere in the literature to study student preconceptions in astronomy (Lombardi et al., 2013; Prather et al., 2004).

The one-semester, survey astronomy course upon which our results are based was taught at the University of Maine using the latest editions of the textbooks *Discovering the Universe* (Comins & Kaufmann III, 2012) or *Discovering the Essential Universe* (Comins, 2013). The ABI was administered to students between the last lecture and the final exam. Statements in the ABI were labeled, e.g., sA46, for statement number 46. (We omitted the labels when we administered the ABI in class.) Previously, the ABI was written and administered during different semesters in either of two formats: (i) as all inaccurate statements, from Fall 2009 to Fall 2012, or (ii) as a random distribution of scientifically correct and inaccurate statements, during the Spring 2013 and Fall 2013 semesters. The labels of those statements that were rephrased from incorrect to correct inherited an "n" at the end. For example, sA46 is "Jovian planets (Jupiter, Saturn, Uranus, Neptune) have solid surfaces," whereas sA46n, in Appendix A, is "Jovian planets (Jupiter, Saturn, Uranus, Neptune) have gaseous or icy surfaces."

For either version, the students were asked to indicate when they first heard of each statement, if ever, and indicate if they used to believe them only as a child, adolescent, or as a result of taking the course. See Paper I for details. We found that the ABI is not biased in terms of the order of statement presentation or its relatively long length (215 statements). However, we found that the different formats of the inventory (all incorrect statements vs. a mixture of correct and incorrect statements) created a statistically significant difference in the fraction of misconceptions that students believed by the end of the course. This means that we would get a more accurate representation of misconception endorsement using correct/incorrect questions, which we did. As we will show in a future paper, however, *correlations* between misconceptions (the tendency that if you believe X, you also believe Y, and vice versa) are relatively unaffected by the different formats.

METHOD

The ABI was administered in six semesters at the University of Maine from Fall 2009 through Fall 2013, with a total of 639 subjects in the overall study. In this paper, we focus exclusively on data from the Fall 2013 semester (*N* = 110, of which 67 students are male, and 43 students are female). The overall age of the sample is 19.6 years, with a median of 19, a mode of 18, and a range from 18 to 41. The subjects' ethnicities are 82.7% Caucasian, 2.73% Hispanic, 1.82% Asian, and 12.7% other/unspecified.

The version of the inventory used in the Fall 2013 semester consisted of both true and false statements. The course was taught by Neil F. Comins (hereafter NFC). As noted above, when taking the ABI, students were asked to indicate if they ever believed each of the statements. Students were also given the option to report if they first believed a statement prior to instruction in college. In this paper, we examined the validity of the data provided by the students (i.e., are they trying to recall their beliefs or are they providing incorrect reports) by comparing the ABI responses with data from four other instruments administered that semester: (i) a cumulative set of three prelims (standard course exams) in the course each covering different topics that span the material presented in the course, with each prelim administered prior to the ABI, (ii) the final exam, administered a few days after the ABI, (iii) a special pretest on black holes and galaxies, administered prior to the presentation of this material in class, and (iv) written responses to select "attendance questions" asked by NFC in class, as described in the introduction. Attendance questions were administered at the end of each class (except during a test day). Table 1 presents the chronology of test administration during the Fall 2013 semester, with the exception of the attendance questions.

Table 1
Chronology of test administration during the Fall 2013 semester

Date	*Test*
October 3	Prelim 1 (standard course exam 1)
October 31	Prelim 2 (standard course exam 2)
October 31	BHGP (administered after Prelim 2)
December 5	Prelim 3 (standard course exam 3)
December 13	ABI
December 17	Final Exam

The framework in which we test the validity of the ABI is the method of concurrent validity. As described by Zumbo and Rupp (2004), "one of the current themes in validity theory is that construct validity is the totality of validity theory and that its discussion is comprehensive, integrative, and evidence based" (p. 84). Those authors suggest that there have been few advances in validity theory in recent years, and the emphasis is now on the consequences of testing, essentially meaning criterion-referenced validity. Kline (2005) adds that "one common way to assess the utility of test scores is to use them to predict other variables of interest" (p. 203). We test for concurrent validity in the ABI, using the approach suggested by Kline, by measuring the correlation between the data in the ABI and the data from each of the other four instruments, one at a time. The statistic that we use to determine the measure of the correlation is the Pearson product-moment coefficient (Onwuegbuzie, Daniel, & Leech, 2007), which we define as the correlation coefficient *r*. A test of the correlation coefficient (i) measures the extent that two variables are linearly related, and (ii) calculates the significance of the correlation by reporting the p -value. The p -value is the probability that a result at least as extreme as observed happens by chance from a random selection of the available data. A correlation coefficient of $r = 1$ represents a perfect linear relationship, and values between $0 < r < 1$ indicate a positive correlation. We define a correlation as statistically significant if $p < .05$. For variables with dichotomous scoring (e.g., correct vs. incorrect), this correlation is called a phi coefficient.

The internal consistency of a set of data is reported by coefficient alpha (α), , sometimes called Cronbach's alpha, which is a measure between 0 and 1 indicating how closely related a set of variables are in a group (Schmitt, 1996). The value $\alpha \approx 0$ indicates dissimilar items, whereas the value $\alpha \approx 1$ indicates very similar items, with $\alpha = 1$ referring to two or more identical data sets. Values of $\alpha \geq .70$ represent a group of items with "adequate" internal consistency. If the group consists of items pooled together from different topics, then it is possible for the pooled data to exhibit lower consistency than when comparing the data within only one topic. The effect is that α could be lower for the pooled group than for the data in the sets taken individually.

As described in Paper I, a master code was developed for all responses to the ABI, to indicate relative degrees of misconception retainment. We define retainment as "the tendency for students to hold on to a misconception from either their childhood or during some point in the course" (Favia et al., 2014). Three of the four instruments (prelims, final exam, and the pretest on black holes and galaxies) are scored dichotomously, with "1" for a correct answer, and "0" for an incorrect answer. As we discuss in the section on attendance, originally three categories were established for each attendance question. For the purpose of assessing the concurrent validity of the ABI, we reduced the number of attendance question categories down to two, based simply on "correct" and "incorrect" answers. Effectively, all four instruments are scored dichotomously.

Table 2
Codes for comparing the data between the ABI and each instrument

Prelims	0: disbelieved correct statement or endorsed inaccurate statement 1: endorsed correct statement or disbelieved inaccurate statement
Final Exams	0: disbelieved correct statement or endorsed inaccurate statement 1: endorsed correct statement or disbelieved inaccurate statement
Black Holes and Galaxies Pretest	0: disbelieved correct statement or endorsed inaccurate statement, prior to college 1: endorsed correct statement or disbelieved inaccurate statement, prior to college
Attendance Questions	0: disbelieved correct statement or endorsed inaccurate statement, prior to college 1: endorsed correct statement or disbelieved inaccurate statement, prior to college

To make comparisons between each of the four instruments and the ABI, we also coded the data in the ABI into two categories. The ABI was administered after instruction in the course, as were the prelims (one for each set of topics) and the final exam. To make post-instruction comparisons between the ABI and either of these tests, we recoded the ABI data into "1" if the student endorsed a scientifically-correct statement or disbelieved an inaccurate statement, and "0" if the student disbelieved a scientifically-correct statement or endorsed an inaccurate statement. Using the data from all students, we then correlated the scores (0 or 1) for each item on the test with the scores for the associated item on the ABI. On the other hand, the pretest and the attendance questions were administered prior to instruction on their respective topics. To make pre-instruction comparisons between the ABI, which was administered at the end of the course, and these two pre-instructional tests, we recoded the ABI data into "1" if the student claimed to believe a scientifically-correct statement or disbelieve an inaccurate statement at a time prior to college, and "0" if the student claimed to disbelieve a scientifically-correct statement or believe an inaccurate statement at a time prior to college. Table 2 presents a summary of the codings used for comparing the data between the ABI and each instrument.

RESULTS

Course prelims

Table 3
Correlations between prelim questions and their associated ABI statements

Prelim	*Question*	*ABI Statement*	*r*	*p*
1	14	sA50n	-.041	.670
1	38	sA80	.255	.007*
2	2	sA164n	.187	.051
2	25	sA50n	-.032	.740
2	30	sA172	.069	.474
2	39	sA76	.252	.008*
2	41	sA46n	-.040	.678
3	7	sA150	-.084	.384
3	7	sA222n	-.037	.703
All questions		All statements	.239	.012*

Three prelims (standard course exams, 50 multiple-choice questions each with five answers) were administered in AST 109. For our purposes, the prelims serve as a check on the consistency of one's recollections, because some prelim questions are based on the common misconceptions explored in the ABI. Our expectation is that students who answer a prelim question correctly ought to endorse the correct science for the associated ABI statement. The prelim questions are scored as "correct" and "incorrect." We initially identified 12 prelim questions associated with one or more associated ABI statements. Cronbach's alpha for the prelim questions was 0.48. Because the prelims are just course exams, they are not optimally designed as research instruments. To improve the reliability of the test, we removed four questions (one of which was associated with two ABI statements) to bring alpha up to 0.52. As discussed earlier, we expect a low value for alpha, because the prelim questions under consideration are pooled together from a broad variety of topics, as would be expected for a multiple-choice exam, and so are not generally expected to correlate internally with each other. In total, we made nine comparisons between prelim questions and their associated ABI statements. A list of ABI statements and their associated prelim questions is presented in Appendix A.

Table 3 presents the correlations between prelim questions and their associated ABI statements, using the method described. The first and second columns, respectively designated "Prelim" and "Question," identify the prelim (which of the first three exams) and the question number from that exam. The third column, designated "ABI Statement," identifies the statement from the ABI that is associated with the prelim question. The correlation coefficient, r, and the p-value for each comparison are reported in the fourth and fifth columns, respectively. Statistically significant p-values are less than .05 and are marked with an asterisk. The final row compares the mean score on all prelim questions under consideration to the mean score on all related ABI statements. We remind the reader that statement labels ending in "n" are phrased as scientifically correct statements, and we have incorporated this into our coding scheme for consistency.

Table 3 shows that there is an overall significant correlation between prelim scores and ABI scores (p = .012). When comparing one question at a time, weak correlations are to be expected, since the data are scored dichotomously. However, when comparing the mean score on all prelim questions under consideration to the mean score on all related ABI statements, we observe a significant correlation, which suggests that student responses to the ABI partially reflect how students perform on the prelims. This result provides support for the validity of the ABI.

Final Exam

The final exam (100 multiple-choice questions each with five possible answers) serves as an additional check for the consistency of one's recollections, because some misconceptions are also shared between the final exam and the ABI. Our expectation is that students who endorse a misconception on the ABI would select an incorrect answer on the associated final exam question. The final exam questions are scored as "correct" and "incorrect." We initially identified 17 final exam questions associated with one or more associated ABI statements. Cronbach's alpha for the prelim questions was .50. To improve the reliability of the test, we removed five questions (two of which were associated with two ABI statements) to bring alpha up to .57. Again, alpha is low because of the variety of topics covered on the final exam. In total, we made 13 comparisons between final exam questions and their associated ABI statements. A list of ABI statements and their associated final exam questions is presented in Appendix B.

Table 4 presents the correlations between final exam questions and their associated ABI statements, using the method described. The first column, designated "Question," identifies the question number from the final exam. The second through fourth columns follow the same format as the third through fifth columns in Table 3. Table 4 shows that there is an overall significant correlation between final exam scores and ABI scores (p = .034), which suggests that student responses to the ABI partially reflect how students perform on the final exam. This result provides further support for the validity of the ABI.

In order to develop more detailed insights into certain pre-instructional beliefs, we next examine student perceptions of black holes and galaxies.

Table 4
Correlations between final exam questions and their associated ABI statements

Question	*ABI Statement*	*r*	*p*
3	sA111n	.225	.020*
9	sA2	.152	.116
10	sA185	.135	.165
21	sA46n	-.065	.504
22	sA172	.178	.065
23	sA171	.275	.004*
24	sA8	.336	<.001*
40	sA248	.186	.054
44	sA108n	.214	.027*
44	sA170	.188	.053
51	sA208	.097	.317
67	sA50n	.032	.747
98	sA262	-.037	.702
All questions	All statements	.204	.034*

Black Holes and Galaxies Pretest

Motivated by *a priori* knowledge of student misconceptions about black holes and galaxies, we developed an 18 multiple-choice question test, called the Black Holes and Galaxies Pretest (BHGP), specifically for the Fall 2013 semester. The BHGP was administered without penalty of course credit lost for incorrect responses. The design of the BHGP followed the "distractor driven" model by Sadler (1998), in the sense that we designed the incorrect answers to distract students from endorsing the correct information by prompting memory recall of their prior misconceptions. The questions and response options to the BHGP are reported in Appendix C. No posttest based on these questions was administered; instead, the ABI, administered at the end of the semester, served as the posttest, since responses to the associated statements about black holes and galaxies can be used to check the reliability of self-reports. The percent of correct responses for each question of the BHGP are presented in Table 5. The associated statements on the ABI to each question are also presented. The overall fraction of questions on the BHGP that students answered correctly was 44%.

To test for the concurrent validity between the BHGP and the ABI, we created two separate sets of questions, one set for the black holes questions, and one set for the galaxies questions. For each set of questions independently, we determined Cronbach's alpha, then improved the reliability of each set by removing selected questions. We then correlated the data from the questions in each reduced set with their associated ABI statements. Because the BHGP is a pre-instructional instrument, we recoded the data in the ABI in the manner presented in Table 2. Specifically, we coded the data to reflect whether or not students reported that they believed a misconception prior to instruction in college.

For the set of black hole questions, Cronbach's alpha was .249. To improve the reliability of the test prior to correlating the data between the black hole questions and the ABI, we removed three questions (one of which was associated with two ABI statements) to bring alpha up to .408, which is still low. We attribute the low value of alpha to (i) the short length of the test and (ii) the different sets of black hole qualities, as presented within the test. For instance, questions 1, 3, and 7 pertain to the creation, composition, and detection of black holes, respectively, whereas questions 6 and 8 ask students to consider how black holes affect objects in their vicinity. These sets of qualities convey different aspects about black holes, and by asking about them together in the same test, the internal consistency of the test was lowered. In total, we made seven comparisons between black hole questions and their associated ABI statements.

Table 5
Black Holes and Galaxies Pretest questions, the percent of students who answered correctly, and the associated statement(s) in the ABI

ABI Statement(s)	*% Correct*	*Question*
sA232	40%	1. How are black holes created today?
sA233n, sA245, sA247	30%	2. What is the fate of black holes?
sA235, sA237n	73%	3. Black holes consist of:
sA246	14%	4. As detectable from our universe, what shape does a black hole have?
sA238	9%	5. A spacecraft is put into circular orbit around a distant black hole. Which one of the following outcomes will happen to the spacecraft?
sA240n	67%	6. What would happen to an asteroid that passed into a black hole?
sA242n	70%	7. How do astronomers detect black holes?
sA243n, sA244	65%	8. If we boarded a spaceship to journey into a black hole, what would happen to us?
sA218n, sA225n	74%	9. How many galaxies exist in the universe today?
sA219n	12%	10. Relative to the Milky Way galaxy, the solar system is located:
sA220	42%	11. How many distinct shapes do galaxies have?
sA221, sA224	53%	12. Where in the universe is the Milky Way located today?
sA222n	18%	13. Where is the Sun located relative to the Milky Way?
sA226	20%	14. How are galaxies distributed throughout the universe?
sA227n	19%	15. Which statement about observing stars in the Milky Way is most accurate?
sA228n	80%	16. Which one statement about galaxy properties is correct?
sA230n	82%	17. Which statement most accurately describes the contents of the Milky Way?
sA231n	34%	18. Which one of the following is true about the Milky Way?

Table 6 presents the correlations between the black hole questions and their associated ABI statements, using the method described earlier. The first column, designated "Question," identifies the question number from the BHGP. The second through fourth columns follow the same format as those in Table 4.

Table 6
Correlations between black hole questions and their associated ABI statements

Question	*ABI Statement*	*r*	*p*
1	sA232	.275	.004*
3	sA235	.139	.152
3	sA237n	.142	.143
6	sA240n	.209	.030*
7	sA242n	.124	.202
8	sA243n	.106	.278
8	sA244	.078	.424
All questions	All statements	.348	<.001*

Table 6 shows that there is a significant correlation between black hole question scores and ABI scores ($p < .001$), which suggests that student responses to the ABI partially reflect students' preconceived notions about black holes.

We then proceeded to analyze the galaxy questions in the same manner. For the set of galaxy questions, Cronbach's alpha was .435. To improve the reliability of the test prior to correlating the data between the galaxy questions and the ABI, we removed five questions (one of which was associated with two ABI statements) to bring alpha up to .534. In total, we made six comparisons between galaxy questions and their associated ABI statements. Table 7 presents the correlations between the galaxy questions and their associated ABI statements, using the method described. The first column, designated "Question," identifies the question number from the BHGP. The second through fourth columns follow the same format as those in Table 4.

Table 7
Correlations between galaxy questions and their associated ABI statements

Question	*ABI Statement*	*r*	*p*
11	sA220	.270	.005*
12	sA221	.246	.010*
12	sA224	.213	.027*
16	sA228n	.114	.241
17	sA230n	.100	.307
18	sA231n	.195	.044*
All questions	All statements	.343	<.001*

Table 7 shows that there is a significant correlation between galaxy question scores and ABI scores ($p < .001$), which suggests that student responses to the ABI partially reflect students' preconceived notions about galaxies. This result, combined with the result of Table 6, holds the promise that the ABI could be administered as a pre-instructional evaluation of student beliefs on black holes and galaxies.

Attendance Questions

As discussed in the introduction, at the end of each class, NFC asks a misconception-based question (for attendance credit) about a topic to be lectured in the subsequent class day. A select set of these "attendance questions" serves as yet another pre-instruction probe for checking the consistency of student recollections of their past misconceptions. Appendix D presents the attendance questions under consideration, their specific categories, their associated ABI statements, the number of student responses to each attendance question, and the fraction of student responses to each category of answers. Included with the set of attendance questions in Appendix D is Fleiss' kappa (κ) for each question. A brief discussion of the meaning of Fleiss' κ follows shortly. Occasionally two ABI statements would overlap with one attendance question. In total, we made 12 comparisons between attendance questions and their associated ABI statements. For each attendance question, we determined categories that best capture the correctness of the responses, as described in Table 8.

Table 8
Attendance question response categories

Category	*Representation*
1	a typical correct response by a student
2	a typical incorrect response that is associated with one or more common misconceptions, such as those already present in the ABI
3	all other "incorrect" responses

We checked the consistency of the categories by performing an analysis of inter-rater reliability (Elder Jr., Pavalko, & Clipp, 1993, pp. 42–44) on four of the nine attendance questions selected at random. For a particular attendance question, one of us scored a sample of written responses, then presented the same sample to an independent rater who did the same scoring without seeing the results of the first rater. For each analysis, a subset of 25 responses was selected at random. The agreement between the two raters, represented by Fleiss' κ, is given by

$$\kappa = \frac{\text{overall agreement} - \text{chance agreement}}{1 - \ \text{chance agreement}}$$

where κ > .60 represents at least substantial agreement (Landis & Koch, 1977). For our analysis, κ ranged from 0.61–1.00.

Three categories were established for each attendance question. For the purpose of assessing the concurrent validity of the ABI, we now reduce the number of attendance question categories down to two, in which we coded score "1" for the correct answer and score "0" for an incorrect answer. We then correlated the scores on the attendance questions with the scores on the associated ABI statements. Because the attendance questions serve as a pre-instructional instrument, we recoded the data in the ABI in the manner presented in Table 2. Specifically, we coded the data to reflect whether or not students reported that they believed a misconception prior to instruction in college. Table 9 presents the correlations between attendance questions and their associated ABI statements, using the method described earlier. The first column, designated “Question,” identifies the question number from the list of attendance questions in Appendix D.

Table 9
Correlations between attendance questions and their associated ABI statements

Question	*ABI Statement*	*r*	*p*
1	sA2	.122	.219
2	sA112	.207	.041*
2	sA148	.277	.006*
3	sA50n	.118	.236
4	sA161	-.067	.508
5	sA46n	.066	.508
6	sA172	.342	.001*
7	sA189	.371	.001*
7	sA190n	.189	.086
8	sA248	-.051	.636
9	sA150	.265	.009*
9	sA222n	.289	.004*
All questions	All statements	.343	.001*

Table 9 shows that there is a significant correlation between attendance question scores and ABI scores (p = .001), which suggests that student responses to the ABI partially reflect the reports that students give on the attendance questions. This result not only provides further support for the validity of the ABI, but also holds the promise that the ABI could be administered as a pre-instructional evaluation of student beliefs in a college-level setting.

DISCUSSION

Comparison of Instruments

We have compared the ABI data to data from the course prelims, the final exam, a pretest on black holes and galaxies, and written responses to attendance questions. Our findings regarding the concurrent validity of the ABI in relation to these instruments are summarized in Table 10, which additionally reports the sample size n, the mean score M, and the standard deviation σ for these instruments, except for the attendance questions, which were administered weekly throughout the semester and whose responses were coded into categories. We remind the reader that for the BHGP, one test for each of the two topics was conducted, giving a total of five tests in four instruments.

Table 10
Correlations between the ABI and each of the four instruments. Summary statistics for the prelims are reported in the text

Instrument	*n*	*M*	σ	*r*	*p*
Prelims				.239	.012*
Prelim 1	167	66.0	16.0		
Prelim 2	165	58.8	16.4		
Prelim 3	161	61.2	17.6		
Final exam	161	60.1	12.9	.204	.034*
Black Holes and Galaxies Pretest					
Black holes	148	46.0	17.4	.348	<.001*
Galaxies	148	43.4	17.2	.343	<.001*
Attendance questions				.343	.001*

All four instruments provided support for the validity of the ABI based on statistical significance ($p < .05$). The prelims and final exams were administered after instruction, while the BHGP and each attendance question were administered prior to instruction on the relevant topics. Table 10 shows a significant correlation between data from the conventional exams (prelims and the final exam) and the data from the ABI. A significant correlation also exists between the data from free responses and the data from the ABI. However, there is only a weak correlation in the data between the BHGP and the ABI. We are thus confident that the validity of the ABI is supported through significant correlations with conventional tests given after instruction and free-response questions asked prior to instruction. We briefly discuss possible influences on the correlations reported in Table 10.

One influence on the nature of our observed correlations is the scoring of the data. The prelims, the final exam, and the BHGP are multiple-choice exams with all questions containing five options. The attendance questions, however, are free-response questions. The written responses must be categorized and rated. The response format of the ABI more closely mimics a multiple-choice exam than an essay-based exam, since a student completes the ABI by filling in a bubble sheet much like on a multiple-choice exam. One would thus initially expect that correlations may be stronger for multiple-choice exams than for the attendance questions. However, the BHGP is also “distractor driven” (Sadler, 1998), in that the questions were intentionally designed to prompt memory recall of misconceptions that students may harbor. We find that all four instruments produce significant correlations with the ABI, which supports our belief that tests of the concurrent validity of the ABI are not particularly sensitive to the nature of the questions on the instruments listed in Table 10.

An additional influence on the nature of the correlations is that a student who answers a question incorrectly on a prelim could learn the correct related science after taking a prelim and before taking the ABI. For instance, a student has anywhere between a few weeks and a few months (depending on the prelim) to learn the correct science before taking the ABI. The data would have been coded “0” on the prelim but coded “1” on the ABI, which would have caused a reduction in the observed correlation between the prelims and the ABI. Such a delay in administering the ABI would also make the observed correlation less significant. That we observed a significant correlation between the prelims and the ABI ($p = .012$) in our analysis already is promising, to say the least.

Our results motivate us to consider additional plausible influences on the reliability of one's own recollections. In particular, we considered two influences: (i) the student's childhood interest in astronomy, prior to instruction in the course; and (ii) the student's level of stress in the semester. We designed two surveys for the Fall 2013 semester to assess childhood interest in astronomy and stress. We emphasize that the purpose of these surveys is not to test the validity of the ABI, but rather to examine childhood interest and current stress level as *possible influences* on how students report their beliefs on the ABI at the end of one semester of college-level astronomy.

Childhood Interest in Astronomy

The Childhood Interest in Astronomy Survey (CIAS, see Appendix E) asks eight questions of the students to indicate their childhood interest in various activities related to astronomy. A total of 80 students taking AST 109 in the Fall 2013 semester completed the survey, which we posted online and made available only for the first two weeks of the course. For each question, students were asked to report their interest level for each activity, with the options for each question being: "never," "occasionally," and "very often." For example, in Question 1 of the CIAS, if a student has read astronomy books very often during one's own free time, the student responded to Question 1 with "very often." The responses were coded on a kind of Likert metric with "never" scored as 0, "occasionally" scored as 1, and "very often" scored as 2. The codes indicate relative measures of childhood interest. Total scores for each student fell within the range of 0 to 11, with a maximum possible score of 16. The mean and standard deviation of scores on the CIAS were 4.57 and 2.56, respectively, which suggests that the majority of students who took the CIAS had little more than an occasional interest in activities related to astronomy as a child.

One concern about the CIAS is that telescopes, planetaria, and observatories may not have been available to some students, who may have been genuinely interested in astronomy as children. For example, one's interest in astronomy may be influenced if a student has access to these instruments while still a child. We found that for the CIAS, $\alpha = .72$, which is considered adequate (Schmitt, 1996). The internal consistency of the responses would decrease if any one of the eight questions was removed from the survey, except for Question 3, which asks about how often the student used binoculars or a telescope, in which case α increases only marginally to .73. Our results support the hypothesis that we can use data from all eight questions without sacrificing the reliability of the data.

To examine if childhood interest in astronomy influences a student's endorsement of misconceptions after college instruction in astronomy, we compared the childhood interest in astronomy scores for all students with their scores for the entire ABI (that is, the mean of all 215 statements). We again recoded the ABI data into "1" if the student endorsed a scientifically-correct statement or disbelieved an inaccurate statement, and "0" if the student disbelieved a scientifically-correct statement or endorsed an inaccurate statement. The correlation between childhood interest in astronomy and mean ABI score is .265 ($p = .046$), which is statistically significant, because $p < .05$. The fact that the correlation, .265, is positive, is consistent with the notion that higher childhood interest in astronomy is associated with believing more of the correct science and fewer misconceptions.

While we observe a statistically-significant correlation between childhood interest in astronomy, as measured in the CIAS, and overall score on the ABI, we cannot determine if increased childhood interest in astronomy *causes* students to believe fewer misconceptions. To affirm the extent of causality, we would need to measure, in tandem, both the beliefs and interests in astronomy, of subjects still in their childhood, and we do not have these paired data. We thus cannot firmly conclude anything about the nature of causality between these two instruments. However, our result is at least consistent with the hypothesis that the more interested children are in astronomy, the less likely they are to endorse misconceptions after college-level instruction in astronomy.

Stress

The results of the above section are consistent with our expectations that students who take an early-in-life interest in astronomy are more likely to recall their memories (which may or may not include misinformation) of what they learned compared to those students who were not interested in the subject material. On the other hand, memory recall is a reconstructive process, so one might expect that high levels of stress in one's life may interfere with the reconstructive process, thus causing inconsistencies in one's own past recall of their own misconceptions (Kuhlmann, Piel, & Wolf, 2005). In response to these considerations, we developed an instrument to evaluate the stress experienced by students at the time of completing the ABI. One such instrument, originally designed by Cohen, Kamarck, and Mermelstein (1983), asks students to report on their own measure of stress in general. For each question in the survey, subjects indicate the extent to which a particular facet represents them (e.g., how often one has felt overwhelmed). Scores for each question range from 0 to 4, where 0 means the subject has never experienced the facet, and 4 means the subject has very often experienced the facet. A related survey designed to assess one's own stress would be most appropriate to administer simultaneously with the administration of the ABI, which we did.

From the original instrument, we created a shortened survey, consisting of four particular questions that were determined by Geoffrey L. Thorpe as being most inter-correlated and comprising a sufficiently large proportion of the variance. (It would be unfair to the students to ask too many questions about their stress level on the same day as the ABI administration and only a few days before their final exams.) The set of these four questions comprises the Stress Survey (see Appendix F). A total of 108 students completed the Stress Survey. In Questions 1 and 4, higher scores indicate higher stress. In Questions 2 and 3, the metric is reversed, so that lower scores indicate higher stress. We found that for the Stress Survey, $\alpha = .79$. To quantify the responses on a kind of Likert metric, we considered the direction of the scores for all four questions, and ultimately we preserved the codes as originally presented to the students. The total stress score per student is $8 + Q1 + Q4 - Q2 - Q3$, where Q1, Q2, Q3, and Q4 are the scores for each of questions 1-4, respectively. (Note that the number of questions has nothing to do with the metric of the response codes.) Scores fell within the range of 0 to 16. The mean and standard deviation of scores on the Stress Survey were 6.63 and 3.47, respectively, indicating that at the time of administration of the ABI (only days before the final exam), the majority of students who took the Stress Survey reported moderate stress in their lives.

To examine if stress influences a student's endorsement of misconceptions after college instruction in astronomy, we compared the stress scores for all students with the *entire* ABI (that is, the mean of all 215 statements). The correlation between stress and mean ABI score is .049 ($p = .651$), which is not statistically significant. The point of the stress survey was to evaluate whether or not a student's stress level, as measured at the end of the course, has a significant influence on how students report their beliefs on the ABI. Our result does not support this hypothesis, and thus we conclude that there is no relationship between a student's stress level and their reports on the ABI.

CONCLUSIONS

In this paper we have tested the validity of the ABI by comparing the results we get from it with results from other testing instruments. These include prelims, the final exam, a pretest on black holes and galaxies, and attendance questions. We correlated the results from one instrument at a time to those of the ABI. We have found that the validity of the ABI is supported through significant correlations with exams given after instruction, a distractor-driven pretest on black holes and galaxies, and free-response questions asked prior to instruction. We are confident that the ABI could be administered as a pre-instructional evaluation of student beliefs in a college-level setting. We have also discovered a statistically significant correlation between childhood interest in astronomy and the extent to which a student believes more of the correct science and fewer misconceptions after college instruction. We also found that there is no relationship between a student's stress level and their reports on the ABI.

In future papers, we will present the theoretical background for principal components analysis, a technique for identifying groups of correlated misconceptions, as the technique applies to our overall project. We will clarify the extent to which semester-to-semester variability in misconception endorsement influences *correlations* between misconceptions, and we will address the concern that the correlations are not significantly affected by the per-semester variability in misconception endorsement. Further papers will explore item response theory and, using it, we will propose an optimal sequence to teach concepts within individual topics in astronomy.

The authors would like to acknowledge Jessica W. Clark, graduate student at the University of Maine, for volunteering her time as the second rater in our inter-rater analysis.

REFERENCES

Bailey, J. M., Prather, E. E., Johnson, B., & Slater, T. F. (2009). College students' preinstructional ideas about stars and star formation. *Astronomy Education Review, 8,* 010110.

Bailey, J. M., & Slater, T. F. (2004). A review of astronomy education research. *Astronomy Education Review, 2*(2), 20-45.

Bransford, J. D. (2000). *How people learn: Brain, mind, experience, and school.* Washington, DC: National Academy Press.

Brewin, C. R., Andrews, B., & Gotlib, I. H. (1993). Psychopathology and early experience: A reappraisal of retrospective reports. *Psychological Bulletin, 113*(1), 82-98.

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior, 24,* 385-396.

Comins, N. F. (2001). *Heavenly errors: Misconceptions about the real nature of the universe.* New York, NY: Columbia University Press.

Comins, N. F. (2013). *Discovering the essential universe (5th ed.).* New York, NY: W. H. Freeman & Company.

Comins, N. F., & Kaufmann III, W. J. (2012). *Discovering the universe (9th ed.).* New York, NY: W. H. Freeman.

Elder Jr., G. H., Pavalko, E. K., & Clipp, E. C. (1993). *Working with archival data: Studying lives*. Newbury Park, CA: Sage Publication.

Favia, A., Comins, N. F., Thorpe, G. L., & Batuski, D. J. (2014). A direct examination of college student misconceptions in astronomy: A new instrument. *Journal and Review of Astronomy Education and Outreach, 1*, A21-A39.

Henry, B., Moffitt, T. E., Caspi, A., Langley, J., & Silva, P. A. (1994). On the 'remembrance of things past': A longitudinal evaluation of the retrospective method. *Psychological Assessment, 6*, 92-101.

Kline, T. J. B. (2005). *Psychological testing: A practical approach to design and evaluation.* Thousand Oaks, CA: Sage Publications.

Kuhlmann, S., Piel, M., & Wolf, O. T. (2005). Impaired memory retrieval after psychosocial stress in healthy young men. *Journal of Neuroscience, 25*(11), 2977-2982.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159-174.

Lombardi, D., Sinatrab, G. M., & Nussbaum, E. M. (2013). Plausibility reappraisals and shifts in middle school students' climate change conceptions. *Learning and Instruction, 27,* 50-62.

Olson, J. M., & Cal, A. V. (1984). Source credibility, attitudes, and the recall of past behaviors. *European Journal of Social Psychology, 14*, 203-210.

Onwuegbuzie, A. J., Daniel, L., & Leech, N. L. (2007). Pearson product-moment correlation coefficient. In N. J. Salkind & K. Rasmussen (Eds.*), Encyclopedia of measurement and statistics* (pp. 751-756). Thousand Oaks, CA: Sage Publications.

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 non-science majors. *Astronomy Education Review, 3,* 122-136.

Sadler, P. M. (1998). Psychometric models of student conceptions in science: Reconciling qualitative studies and distractor-driven assessment instruments. *Journal of Research in Science Teaching, 35*(3), 265-296.

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment, 8*(4), 350-353.

Slater, T. F., & Adams, J. P. (2002). *Learner-centered astronomy teaching: Strategies for astro 101*. Boston, MA:Addison-Wesley.

Smith III, J. P., diSessa, A. A., & Roschelle, J. (1993). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *The Journal of the Learning Sciences, 3*(2), 115-163.

Strike, K. A., & Posner, G. J. (1992). A revisionist theory of conceptual change. In R. Duschl & R. Hamilton (Eds.), *Philosophy of science, cognitive science, and educational theory and practice* (pp. 147-176). Albany, NY: State University of New York Press.

Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction, 4*, 45-69.

Zumbo, B. D., & Rupp, A. A. (2004). Responsible modeling of measurement data for appropriate inferences: Important advances in reliability and validity theory. In D. Kaplan (Ed.), *The sage handbook of quantitative methodology for the social sciences* (pp. 73-92). Thousand Oaks, CA: Sage Publications.

APPENDIX A

PRELIM QUESTIONS AND ASSOCIATED STATEMENTS FROM THE ASTRONOMY BELIEFS INVENTORY

sA50n: few objects orbit the Sun in circular orbits
Prelim 1 Question 14: Which of the following most accurately describes the shape of Neptune's orbit around the Sun?

sA80: craters are volcanic in origin
Prelim 1 Question 38: Which statement about the surface of the Moon is correct?

sA164n: Mars currently lacks running water on its surface
Prelim 2 Question 2: Which statement about water on Mars is correct?

sA50n: few objects orbit the Sun in circular orbits
Prelim 2 Question 25: What shape is the orbit of Halley's comet?

sA172: Saturn's rings are solid
Prelim 2 Question 30: Saturn's rings are best described by which of the following?

sA76: Jupiter's Great Red Spot is a volcano erupting
Prelim 2 Question 39: Jupiter's Great Red Spot is most accurately called a(n)?

sA46n: Jovian planets (Jupiter, Saturn, Uranus, Neptune) have gaseous or icy surfaces
Prelim 2 Question 41: Which statement most accurately describes descending into Jupiter?

sA150: the Earth is in the middle of the Milky Way galaxy
sA222n: the Sun is far away from the center of the Milky Way galaxy
Prelim 3 Question 7: What is located at the center of the Milky Way galaxy?

APPENDIX B

FINAL EXAM QUESTIONS AND ASSOCIATED INVENTORY STATEMENTS

sA111n: Earth's axis is tilted compared to the ecliptic
Final Exam Question 3: Where is the ecliptic as seen from the Earth?

sA2: there are 12 zodiac constellations
Final Exam Question 9: How many *zodiac* constellations are there?

sA185: the Sunspot cycle is 11 years long
Final Exam Question 10: What is the length of the Sun's entire magnetic cycle?

sA46n: Jovian planets (Jupiter, Saturn, Uranus, Neptune) have gaseous or icy surfaces
Final Exam Question 21: Which of the following best describes the surface of Jupiter?

sA172: Saturn's rings are solid
Final Exam Question 22: Which statement about Saturn's rings is most accurate?

sA171: Saturn is the only planet with rings
Final Exam Question 23: Which of the following planets does not have rings?

sA8: the north star is the brightest star in the sky
Final Exam Question 24: What is the brightest star in the night sky?

sA248: cosmic rays are light rays
Final Exam Question 40: Cosmic rays are best described by which of the following?

sA108n: Venus is similar to earth in size
sA170: Mars is the sister planet to earth in physical properties and dimensions
Final Exam Question 44: Which planet is most similar in physical dimensions and composition to the Earth?

sA208: the Sun will explode as a nova
Final Exam Question 51: The Sun will end its "life" as a(n):

sA50n: few objects orbit the Sun in circular orbits
Final Exam Question 67: Which of the following most accurately describes the shape of Venus's orbit around the Sun?

sA262n: the universe as a whole is changing
Final Exam Question 98: Which of the following best describes the overall motion of the universe?

APPENDIX C

BLACK HOLES AND GALAXIES PRETEST

Question 1: How are black holes created today?

(A) They form spontaneously (that is, without the need for any external matter or energy) in empty space
(B) While black holes once formed, they do not form today
(C) They form deep inside some old stars
(D) They form as the result of collisions between high-energy particles (atoms or molecules) in space
(E) There is no evidence that black holes exist, hence I don't believe they form today.

Question 2: What is the fate of black holes?

(A) They last forever, each with a fixed mass
(B) They slowly lose mass, that is, they evaporate
(C) They continue to gain mass indefinitely
(D) They continue to grow, but one is growing faster than the others. It will eventually swallow the universe
(E) Black holes don't exist, hence the question is irrelevant

Question 3: Black holes consist of:

(A) nothing - they are just holes in space
(B) a uniform "sea of energy"
(C) ultra-condensed concentrations of matter and energy
(D) ultra-condensed concentrations of matter, only
(E) misleading question since black holes don't really exist

Question 4: As detectable from our universe, what shape does a black hole have?

(A) a pinhole in space
(B) a sphere
(C) a wormhole
(D) many different possible shapes, including spherical, disc, and donut shapes
(E) misleading question since black holes don't really exist

Question 5: A spacecraft is put into circular orbit around a distant black hole. Which one of the following outcomes will happen to the spacecraft?

(A) It will be sucked straight into the black hole, like dust into a vacuum cleaner
(B) It will spiral inward, eventually falling into the black hole
(C) It will stop and hover above the black hole
(D) It will orbit around the black hole forever in the same orbit
(E) The question assumes black holes exist, which I believe they don't.

Question 6: What would happen to an asteroid that passed into a black hole?

(A) It would be crushed and would remain in the black hole
(B) It would enter another dimension
(C) It would tunnel through a wormhole back into our universe
(D) It would remain intact in the black hole
(E) The question assumes black holes exist, which I believe they don't.

Question 7: How do astronomers detect black holes?

(A) Astronomers detect black holes by observing their glow
(B) Astronomers detect black holes by seeing matter being ejected from them
(C) Astronomers detect black holes by putting spacecraft in orbit around them
(D) Astronomers detect black holes by their effects on nearby gas and stars
(E) No one has actually found a black hole

Question 8: If we boarded a spaceship to journey into a black hole, what would happen to us?

(A) We would begin to spin faster and faster
(B) We would get stuck inside the black hole
(C) We would have the ability to travel freely in time
(D) We would teleport to another dimension
(E) We would never find a black hole since they don't exist

Question 9: How many galaxies exist in the universe today?

(A) There are no galaxies
(B) There is just our galaxy
(C) There are a few galaxies
(D) There are several thousand
(E) There are billions of galaxies

Question 10: Relative to the Milky Way galaxy, the solar system is located:

(A) above the Milky Way
(B) between two spiral arms
(C) in a spiral arm
(D) in or near its center
(E) far away from Milky Way

Question 11: How many distinct shapes do galaxies have?

(A) All galaxies are shaped like spirals
(B) Galaxies can be either spiral or elliptical
(C) Galaxies can have any shape other than spiral or elliptical
(D) Galaxies can be spiral, elliptical, or irregularly shaped
(E) Galaxies can have any shape

Question 12: Where in the universe is the Milky Way located today?

(A) The Milky Way has no special location in the universe
(B) The Milky Way is located at or near the center of the universe
(C) The Milky Way is the universe
(D) The Milky Way is located outside the universe
(E) The Milky Way is located inside the solar system

Question 13: Where is the Sun located relative to the Milky Way?

(A) The Sun is located above the plane of the Milky Way
(B) The Sun is located between two spiral arms of the Milky Way
(C) The Sun is located at the Milky Way's center
(D) The Sun is located in a spiral arm of the Milky Way
(E) The Sun is located far outside of the Milky Way

Question 14: How are galaxies distributed throughout the universe?

(A) The galaxies are distributed randomly
(B) The galaxies are grouped together by size
(C) The galaxies are in large groups separated by voids
(D) The galaxies are grouped together by type or design
(E) There is only one galaxy in the universe

Question 15: Which statement about observing stars in the Milky Way is most accurate?

(A) Astronomers can observe up to a few hundred stars in the Milky Way
(B) Astronomers can observe those stars not obscured by galactic dust
(C) Astronomers can observe hundreds of thousands of stars in the Milky Way
(D) Astronomers can observe all of the stars in the Milky Way
(E) Astronomers can observe only the brightest stars (O, B, A)

Question 16: Which one statement about galaxy properties is correct?

(A) All galaxies are held together by gravity
(B) All galaxies have the same shape
(C) All galaxies have the same size
(D) All galaxies have the same color
(E) All galaxies rotate at the same speed

Question 17: Which statement most accurately describes the contents of the Milky Way?

(A) The Milky Way lacks gas, but contains dust and stars
(B) The Milky Way lacks dust, but contains gas and stars
(C) The Milky Way lacks gas and dust, but contains stars
(D) The Milky Way contains gas, dust, and stars
(E) The Milky Way lacks stars, gas, and dust altogether

Question 18: Which one of the following is true about the Milky Way?

(A) The Milky Way is still producing both stars and planets
(B) The Milky Way is still producing stars, but not planets
(C) The Milky Way is still producing planets, but not stars
(D) The Milky Way is no longer producing stars or planets
(E) The Milky Way never created its own stars or planets

APPENDIX D

ATTENDANCE QUESTIONS

Attendance Question 1

n = 171, no inter-rater analysis conducted

sA2. there are 12 zodiac constellations

September 5, 2013 (immediately at the start of class, prior to lecture). How many zodiac constellations are there?

1. (36%) 13
2. (34%) 12
3. (35%) Other, vague, unsure

Attendance Question 2

n = 162, $\kappa = 0.74$

sA112. summer is warmer because we are closer to the sun during the summertime

sA148. seasons are caused by speeding up and slowing down of Earth's rotation

September 5, 2013 (at the end of class, as normal). What causes the seasons?

1. (38%) Tilt of the Earth
2. (17%) Changing distance between the Earth and the Sun
3. (46%) All other (less logical) (e.g. "the Earth's rotation")

Attendance Question 3

n = 166, no inter-rater analysis conducted

sA50n. few objects orbit the Sun in circular orbits

September 10, 2013. What is the shape of the Earth's orbit around the sun?

1. (86%) elliptical or oval
2. (13%) circular or sphere
3. (2%) other

Attendance Question 4

n = 154, $\kappa = 0.62$

sA161. Mars is green (from plant life)

September 30, 2013. Have astronomers found life on Mars?

1. (38%) No, (5%) Only as fossils (traces)
2. (18%) Other (e.g., found water/ice, didn't answer question)
3. (40%) Yes

Attendance Question 5

n = 150, no inter-rater analysis conducted

sA46n. Jovian planets (Jupiter, Saturn, Uranus, Neptune) have gaseous or icy surfaces

October 3, 2013. Describe the surface of Jupiter

1. (37%) Gas, liquid (but not water), or "no surface" - but not solid
2. (35%) Solid - not gas (e.g., "cratered")
3. (29%) All other, including mixed surface types (e.g., gas and rocky)

Attendance Question 6

n = 138, κ = 0.71

sA172. Saturn's rings are solid

October 10, 2013. Briefly describe the rings of Saturn

1. (78%) Particles, ice, debris, rocks, asteroids, metals, "like Jupiter's"
2. (14%) Solid continuous matter, inc. gravity-held "material" or "rocky elements"
3. (8%) Gas only, clouds, or liquid

Attendance Question 7

n = 120, no inter-rater analysis conducted

sA189. the Sun is the brightest star in universe

sA190n. some objects in the universe are brighter than the Sun

November 7, 2013. How bright is the sun compared to other stars?

1. (40%) average or about average
2. (37%) dimmer, or (22%) brighter
3. (2%) the Sun is the brightest star

Attendance Question 8

n = 128, κ = 1.00

sA248. cosmic rays are light rays

November 12, 2013. What are cosmic rays?

1. (1%) High-energy particles from space, (2%) High-energy particles (not from space or unspecified)
2. (21%) Light or light rays
3. (77%) All other (inc. vague, unspecified, multiple responses, e.g., "rays coming from stars," "radiation," "shock waves from supernovas")

Attendance Question 9

n = 125, no inter-rater analysis conducted

sA150. the Earth is in the middle of the Milky Way galaxy

sA222n. the Sun is far away from the center of the Milky Way galaxy

November 21, 2013. Where in the Milky Way are we?

1. (2%) Between two spiral arms, (30%) On or in a spiral arm
2. (30%) In or near the center
3. (37%) All other or too vague to tell (e.g., "upper left," "off to the side," "on the edge")

APPENDIX E

THE CHILDHOOD INTEREST IN ASTRONOMY SURVEY

Possible response options for each question: “never,” “occasionally,” “very often”

1. Prior to college, how often did you read astronomy books on your own?
2. Prior to college, how often did you watch astronomy programs on your own?
3. Prior to college, how often did you use binoculars or a telescope to view the night sky?
4. Prior to college, how often did you participate in an astronomy club out of pure interest?
5. Prior to college, how often did you go to planetarium shows out of pure interest?
6. Prior to college, how often did you go to an observatory out of pure interest?
7. Prior to college, how often did you keep up with news stories related to astronomical events?
8. Prior to college, how often did you choose to talk to others about astronomy, out of pure interest?

APPENDIX F

STRESS SURVEY

[As discussed in the Section on Stress, these questions were adopted from Cohen et al. (1983).]

The questions in this scale ask you about your feelings and thoughts during the last month. In each case, you will be asked to indicate how often you felt or thought a certain way. Although some of the questions are similar, there are differences between them, so you should treat each question separately. The best approach is to answer each question fairly quickly. That is, don’t try to count up the number of times you felt a particular way; instead, indicate the alternative that seems like a reasonable estimate.

For each question, choose from the following alternatives:

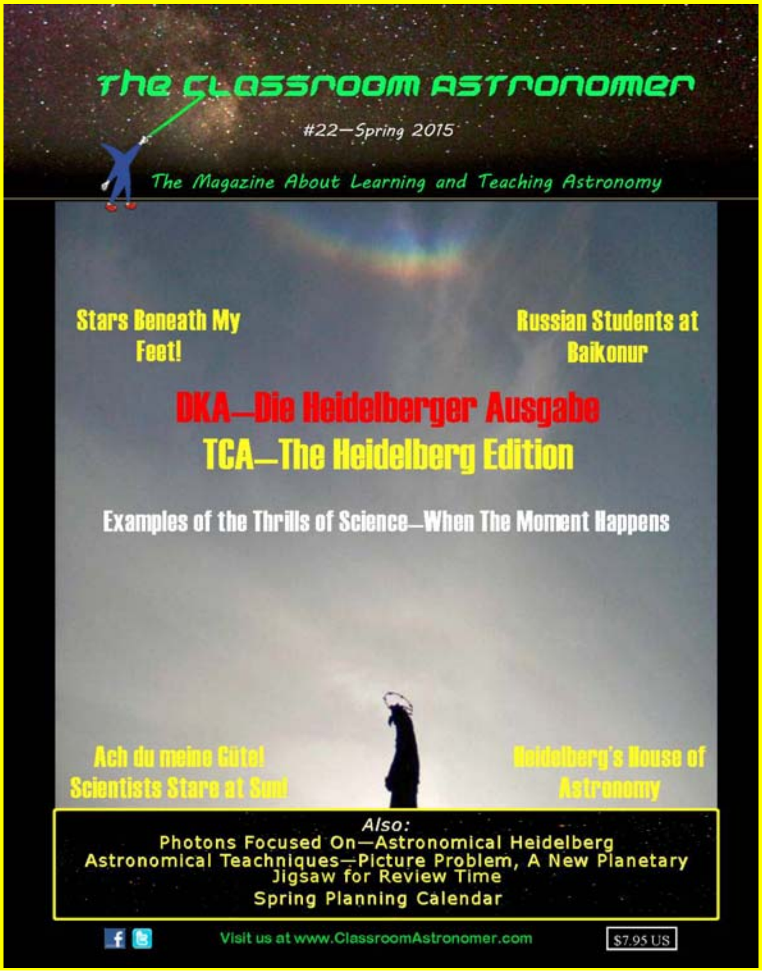
0	1	2	3	4
Never	Almost Never	Sometimes	Fairly Often	Very Often

1. In the last month, how often have you felt that you were unable to control the important things in your life?
2. In the last month, how often have you felt confident about your ability to handle your personal problems?
3. In the last month, how often have you felt that things were going your way?
4. In the last month, how often have you felt difficulties were piling up so high that you could not overcome them?

The Classroom Astronomer

http://www.classroomastronomer.com
3605 Sandy Plains Road, Suite 240-203
Marietta, GA 30066-3068
Tel: 1-404-702-8147 TOLL
Fax: 1-404-393-7037
Email: TCA@toteachthestars.net

The practitioner magazine for astronomy teachers.



The Classroom Astronomer
#22—Spring 2015
The Magazine About Learning and Teaching Astronomy

Stars Beneath My Feet!

Russian Students at Baikonur

DKA—Die Heidelberger Ausgabe
TCA—The Heidelberg Edition

Examples of the Thrills of Science—When The Moment Happens

Ach du meine Güte! Scientists Stare at Sun!

Heidelberg's House of Astronomy

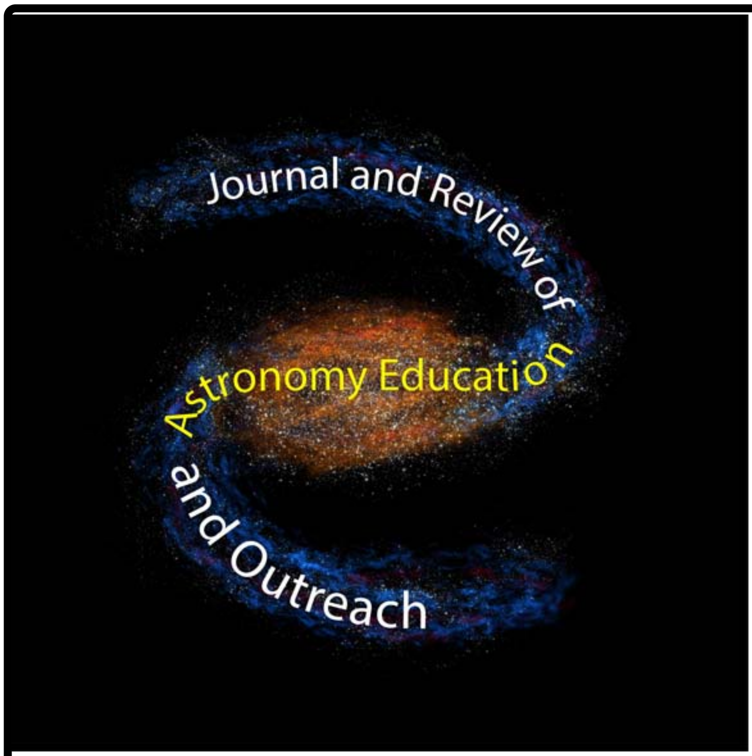
Also:
Photons Focused On—Astronomical Heidelberg
Astronomical Teachniques—Picture Problem, A New Planetary Jigsaw for Review Time
Spring Planning Calendar

Visit us at www.ClassroomAstronomer.com
$7.95 US

An impressive mass of interesting material!
—Guy Ottewell, Author *Astronomical Calendar*

I just spent my lunch hour reading through it, and it's a lot of fun with some good classroom ideas ready to go.
—Colin Jagoe, Kawartha Pine Ridge, District School Board, Canada

Your magazine will be a good resource both for me, and for the B. Ed. students who are taking the elementary science course...Well-done!
—Terry Bridger, Queen's University, Canada



A Scholarly Publication of Research, News and Commentary

From Hermograph Press

The *Journal and Review of Astronomy Education and Outreach* (*JRAEO*) is accepting submissions of articles on any topic that relates to educating students or the general public about astronomy. *JRAEO* is open access, free to read!

JRAEO contains peer-reviewed research articles and regular articles on news of the field, pertinent stories, and commentary.

As long as the topic is found in an astronomy course or used to excite and inform people about astronomy, articles will be welcomed as submissions from writers in:

- astronomy education,
- space science education,
- allied fields such as geoscience, physics, math or history education,
- informal education and public outreach as well as teacher or student education.

All education levels from K to college and beyond will be covered, and submissions are welcome from all countries around the globe.

Digitally published.
Three issues per year.
Subscribers will have reduced page charges waived when they submit an article.
For subscription and submission information, see www.JRAEO.com .

YOU CAN TOUCH THESE! CREATING 3D TACTILE REPRESENTATIONS OF HUBBLE SPACE TELESCOPE IMAGES

Carol Christian, Antonella Nota[1], Perry Greenfield, *Space Telescope Science Institute (STScI)*, Baltimore, MD
Noreen Grice, *You Can Do Astronomy LLC*, New Britain, CT
Natalie Shaheen, *National Federation of the Blind (NFB)*, Baltimore, MD

Received August 11, 2014; Accepted January 18, 2015

Abstract: Astronomical imagery is engaging, inspiring, stimulates public interest, and has proven successful in advancing Science Technology Engineering and Mathematics education. Hubble Space Telescope (HST) discoveries are publicly disseminated routinely through text, images, graphics, visualization, multimedia, social networking and other online mechanisms. To extend these resources to visually impaired individuals and other individuals who could benefit from hands-on physical materials, we created prototype tactile renditions of stunning HST astronomical images on a 3D printer. This paper describes briefly the translation of scientific data from the analysis of HST observations into a format appropriate for 3D printing. Then we describe the resulting tactile 3D prints, outfitted with textures associated with features in a celestial object, specifically using the star-forming region NGC 602 as a prototype. We outline the various textures we adopted and how they were evaluated by several focus groups. We converged on a production and print method and a robust set of textures for blind users and other learners who can benefit from tactile materials. Ultimately a library of these tactile 3D print files can be integrated with a suite of HST educational resources available and distributed through the internet.

Keywords: specialized (visually impaired) students - all levels – astronomy – technology in education - 3D printing - tactile materials

INTRODUCTION

Astronomical images provide a principal source of information that forms public conceptions about space (Snyder, 2011) and can contribute to science, technology, engineering, and mathematics (STEM) education. Historically astronomy originated as a visual science, that

---

[1] Also European Space Agency, c/o STScI

is, observation of phenomena emitted in the visible part of the spectrum that could be perceived by the human eye. Even today, the visual presentation of data from all kinds of astronomical observations (not restricted to the visible regime) is an essential part of scientific research. Visual presentations in the form of imagery from data, models, and artist's conceptions serve as communication tools for conveying astrophysics information to the public (Frankel, 2004). Astronomy is an engaging subject with broad public appeal, stimulating curiosity about ourselves and the universe.

STEM education is considered critical for U.S. competitiveness and in a very broad sense can contribute to sustainability. Astronomy is relevant for K-12 STEM education and contributes to public science literacy as a whole, although it must be cleverly integrated into the standard U.S. curriculum because it is not always a standard subject in the local classroom. Due to the discipline's public appeal, it makes sense to invoke astronomy to contribute to science literacy (National Science Foundation 2010, International Astronomical Union 2012, and documents from Astronomy U.S. Committee on Science, Technology, Engineering, and Math Education[2]).

Important astronomical discoveries, insights, and breakthroughs are often a direct result of perception and human cognition of the visual presentation of data. However, like all science, the real core of astronomical research is a combination of human characteristics including imagination, perception, inquisitiveness, creativity, perspective, and many other capabilities. Of course, these abilities are not restricted to sighted individuals or persons with particular learning styles. Building skill and expertise in science depends not only on the individual's traits but also on the accessibility of tools, appropriate education, access to current data, research, and training, so the more inclusive science is, the greater is the return to individuals and society. A conscious effort to design, create, and adapt educational resources to meet diverse audience learning needs, researching and implementing innovative techniques, and applying appropriate technologies (c.f., Smith, et al., 2014, Kessler, 2007) can broaden the public access to astronomy. The reward will be engaging the previously untapped intellect of individuals who otherwise routinely encounter barriers in participating in science (Diaz, 2014; Grice, 2012b; Grice, 2010; Beck-Winchatz and Riccobono, 2008).

The authors of this paper, which we call the 3D Astronomy Team, know that astronomy contributes greatly to public understanding, engagement, and education in science. Each member of our team has contributed to a variety of education and outreach programs aimed at diverse target audiences. In particular, for the 3D Astronomy Project described here, we are interested in reaching the visually impaired who are seriously underrepresented in science but have a keen interest in astronomy (Christian et al., 2014; Grice, 2012a; Grice, 2010; Grice et al., 2004, and *c.f.*, http://astrokit.uv.es). Our 3D Astronomy Team wished to create an authentic, innovative, education/ outreach tactile resource based on real data and printed on a 3D printer. Therefore, we chose to use our own data obtained from the Hubble Space Telescope (HST) on a celestial object we are researching. Since HST imagery has high public appeal and recognition (National Research Council, 2010; NASA, 2014) the experimental 3D printing project described here also can serve as a gateway for producing a library of 3D tactile materials from any HST image when suitably analyzed as we describe in this paper.

In this paper, we specifically describe the first steps of an experimental project to produce 3D tactile renditions of celestial objects for visually impaired and multi-sensory learners. Our aim is to make science more accessible to those audiences using our innovative tactile approach to data representation. Eventually, we desire to distribute our 3D print files through the robust infrastructure provided by STScI. We will use the STScI framework because exemplary HST resources (found online at http://hubblesite.org) are available and understandable through textual materials, imagery, and other elements such as graphics, videos,

---

[2] http://www.whitehouse.gov/administration/eop/ostp/nstc/committees/costem

and computer visualization of astronomical objects and have been distributed for over 20 years to the public (c.f., NASA, 2014) and http://outreachoffice.stsci.edu). However before we can consider introducing such new resources as regular components to be distributed through the extensive public HST HubbleSite infrastructure, we desired to test the suitability and usability of the 3D materials through focus groups. This paper outlines the creation and iterative testing of the tactile 3D printouts and the potential for utility and distribution of the materials in the future.

UNDERLYING SCIENCE DATA

In all HST education and outreach programs, the essential astrophysics forms the core content for materials made available to diverse audiences and especially in support of STEM education. The scientific expertise of our 3D Astronomy team is centered on the studies of star clusters, that is, groupings of stars that form together and often later dissipate. This topic served as the springboard for the science content for the tactile materials.

Our prototype cluster for this project is named NGC 602 and is shown in a color composite image (Figure 1). This cluster is located in the wing of a companion galaxy to our own, called the Small Magellanic Cloud and was formed about 5 million years ago (Cignoni, M., 2009). The distance to the object (about 200,000 light years from Earth) is far enough so that the HST Advanced Camera for Surveys (ACS) can observe the entire cluster and surroundings, but close enough that HST can see and measure the brightness of the individual stars. The deep optical observations were taken in 2004 to study the cluster's complex morphology. It has a bubble of gas and dust out of which the cluster was formed, and features in the object, such as stars, gas, and dust can be identified.

Figure 1. A color composite image of the NGC 602 star-forming region created from science observations taken with Hubble Space Telescope's Advanced Camera for Surveys (ACS). North is up and East is to the left. (Courtesy of NASA, ESA. Hubble Heritage and STScI/AURA: http://hubblesite.org/newscenter/archive/releases/2007/04/)

As seen in Figure 1, NGC 602 is an ideal target for testing our 3D printing process because it is rich in the following features as noted by Carlson et al (2007):

> *The morphology of the region is reminiscent of a partial ring because the visual image appears as a two-dimensional projection of a ruptured bubble. Two ridges of dust and gaseous filaments outline the nebular shape towards the southeast (lower left in the image) and to the northwest (upper right) and are highlighted by magnificent "elephant trunk" structures. The primary stellar cluster shines in the middle of the broken ring, slightly closer to the southeast ridge. This is where the brightest stars are concentrated. Their winds appear to have swept out the gas and the dust from the center, creating the inner cavity of the bubble. Many faint stars are visible on, or in close proximity to, the two ridges, indicating that star formation is still active there.*

All of these attributes can be appreciated in the HST image through visual examination. Our challenge was to represent this same information in a tactile format.

In addition to graphically documenting the appearance of the composite NGC 602 HST image, we measured the individual HST observations to determine the brightness of each feature seen in the image. Measurement of the individual HST frames allows us to calculate the flux in each wavelength, whereas the intensities in the one composite image are enhanced during image processing to make a better color image presentation, and therefore do not represent an absolute scaling of the physical radiation output from each feature. The outer contour of each individual feature was determined, and then the area within the contour was measured for intensity and used to populate a data cube. The components of the data cube are: an X-Y position, a feature type (stars, gas, dust, and filaments), a measured intensity, and a Z-distance from the observer. As a result, the cluster's physical attributes were all codified--the brightness and the position of all the stars, and the relative intensity and the position of the dust and gaseous regions. The HST observations were augmented with information from the literature, thus being able to further infer the locations where dust and gas are primarily concentrated. We measured the size of the surrounding bubble and estimated its thickness. Ultimately, for the purpose of transforming the scientific measurements into information appropriate for 3D printing, we consolidated all data into three key concepts to convey:

- The overall structure of NGC602 is spherical.
- There is a very bright stellar cluster near the center.
- There are extended regions of gas and dust, for which we had measured the relative intensities.

TRANSFORMATION OF SCIENCE DATA TO 3D PRINT FORMAT

The data cube described above provided the locations and intensities of all the features in the image. This data had to be transformed into a format suitable for 3D printing. Two versions of the data cube were created: first, a file representing the feature locations of the nebula, each with a unique texture (we called this the composite "*texture map*", c.f., Figures 2 and 3a) and second, a file with the measured light intensities for each feature presented as variations or undulations in the surface height or "elevation" of the texture map (called the composite "*elevation map*" *c.f.*, Figure 3b). Therefore the 3D prints represent each type of feature in the HST image in Figure 1 with a unique texture, providing a tactile analog to the visual image that shows NGC 602 as a composite of unique colors and intensities. The *texture* and *elevation maps* are the first steps towards our longer-term objectives which include representing the full 3D structure of the object in a variety of ways – a series of 3D prints organized in individual planes representing different distances, a 3D printout that can be assembled easily into a 3D object, or digital representations integrated with other methods and technologies we are considering which are beyond the scope of this paper.

The transformation of scientific data to a format for 3D printing is not trivial, and the process we developed was a significant part of the 3D Astronomy project. Once the scientific analysis and measurement of the NGC 602 HST observations was complete, we re-sampled, smoothed, and interpolated the data to obtain a reasonable portrayal of NGC 602 that had to be converted into the file format appropriate for 3D printers. Specifically, 3D printers expect the STereoLithography (STL) file format[3], commonly used in Computer Aided Design (CAD) programs and 3-D animation rendering programs, such as Blender. Unfortunately, none of

---

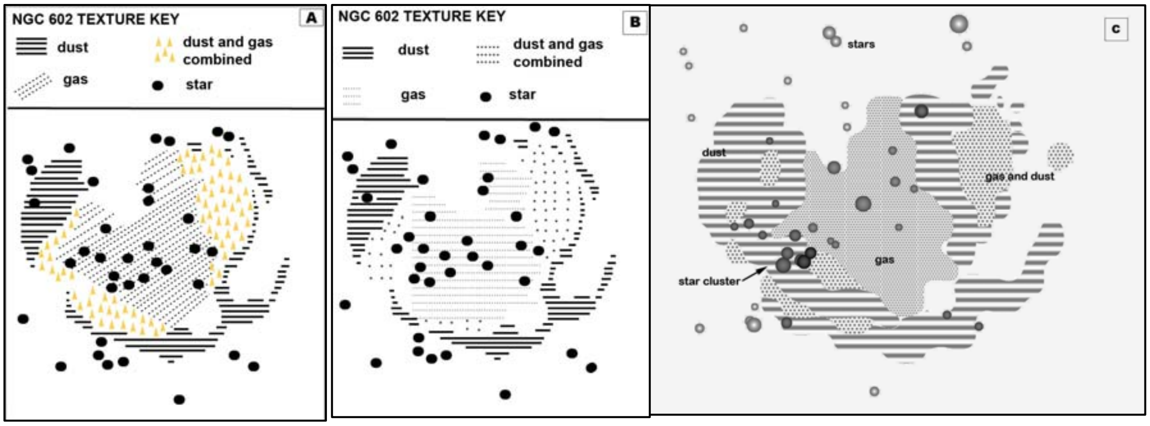
[3] http://www.3dsystems.com/quickparts/learning-center/what-is-stl-file

those current types of software handle astronomical data gracefully. Also, the programs are not able to ingest cubes of data from any numerical scientific analysis or technical measurements. Therefore we processed the smoothed data using in-house developed scripts and programs to produce a format that the commercial software packages mentioned above could handle. Many commercial packages are able to export the data in the STL format if the input data is in the form of grids, meshes, or CAD-like files. The advantage of using STL files is to enable the printing of the tactile materials by anyone with access to low-end affordable printers as well as to high-end exhibit quality devices. Thus an important outcome of this project is that we cannot use a hand-tuned, human-intensive process, but rather, we will need to create a pipeline to allow us to annotate and measure an image, convert the data to suitable file formats, and then transform those files appropriately, in order to process numerous HST images and create a library of STL files for the public.

TEXTURES

Considerable experience has been gained in representing astronomical data with textures to convey information to the visually impaired, c.f., Grice (2012). Before we decided on specific textures to use for 3D printing, we also investigated the literature regarding different learning styles, primarily focusing on sighted and blind individuals (e.g., Dion, Hoffman and Matter (2000); Braille Authority of North America (2011)). Other journals we found useful that contained articles as well as other online reference materials include the *Journal of Visual Impairment and Blindness* (http://www.afb.org/info/publications/jvib/), the *Journal of Blindness Innovation and Research*, and the *A to Z of Brain, Mind and Learning.*

Our testing sessions, described below, were aimed at evaluating the textures and the usability of the printed products. Over time we have learned, through observation of individuals examining displays of HST images, that when some sighted individuals look at an image, they might take in the whole image first, and then examine the finer detail, especially if prompted by textual, hyper-text, or audio guidance. In contrast, some blind and visually impaired learners examine details and then use their “mind’s eye” to combine the parts, creating the whole image.



Figures 2a,b,c. Textures used for testing. a) Initial textures used for Swell Form and small format 3D prints, tested at the NFB 2013 Orlando conference and the STEMX conference. The slanted lines used for gas were less identifiable than other textures on the 3D prints. b) Revised textures for the Connecticut NFB Chapter and Maryland NFB State Convention tests. The gas pattern, although somewhat distinguishable, was thought to be “non intuitive” by some testers. c) Current set of textures being used in informal testing scenarios (not formalized meetings or conventions). The textures for the gas and the gas + dust combination are similar stipple patterns, but the gas texture is much smoother and “softer”, while the gas=dust pattern is rougher and more evocative of that feature.

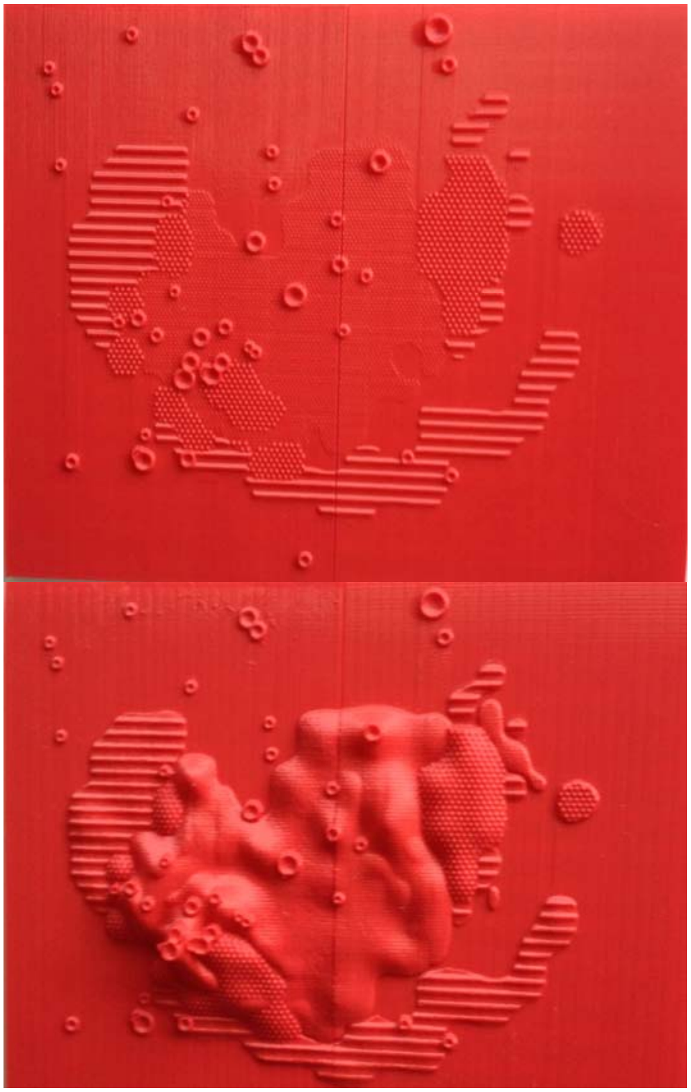


Figure 3. Texture and elevation 3D print maps (large format) of the star cluster NGC 602 from HST data. These maps were tested at the CT and MD NFB State Conventions in November 2013.

We had to accommodate both approaches with our materials by providing tactile legends for the textures, and aural guides to the nature of the astronomical object being depicted. We found that assisting the ability of individuals to distinguish small details can be critical in understanding complex situations depicted in images, so this had to be a key component of our design.

Before we started using the 3D printer, we created template textures on Swell Form prints (thermal expansion paper where darkened areas of the paper “puff” and become tactile when heated in a Swell Form machine, see Figure 2a,b,c,; note that NFB means National Federation of the Blind). This step in our project demonstrated that the Swell Form paper technology was useful in introducing various textures to visually impaired individuals, especially those familiar with Braille, and aided them in distinguishing different patterns, in preparation for examining 3D solid prints. Also the Swell Form prints are much faster to produce, so iteration of textures could be accomplished quickly.

After lengthy examination and tuning of the print process, as well as matching the resolution, file size, and other characteristics to our specific printer, we printed 3D composite *texture* and *elevation* maps for NGC 602, using the textures shown in Figure 2a, then revised as based on testing to those in Figure 2b, and finally those in 2c.

EVALUATION PLAN

Before our 3D materials can be integrated into any robust education or outreach program, we desired to investigate if our development process produced 3D materials that were actually usable and understandable. We developed a method for evaluating the materials and engaged a series of focus groups to assess the materials. The feedback from the groups provided input for modifying both our process and the materials each step of the way.

Evaluation Plan

Our procedure for the testing of materials was organized by asking participants to:

- Examine a visual image, if the user had partial vision,
- Listen to a description read aloud before experiencing the textured materials,
- Examine the Swell Form paper textured image and the legend to introduce the textures (note that after the first two focus groups, this step was recommended as unnecessary by the subsequent focus groups),
- Examine the 3D prints and compare with a single composite Swell Form illustration,
- Compare *texture* and *elevation* maps,
- Respond to a survey to assess user responses to the textures and understanding of the features of the NGC 602 image.

Instrument

We designed a suite of survey questions to probe the participants' specific understanding of the textures and preliminary familiarization with the printout formats, and to obtain some free form feedback (Appendix). Basically, participants were asked:

- If they could identify each of the features (stars, gas, dust, combination) on both the Swell Form (eliminated after the second focus group) and the 3D printouts,
- If they began to understand the Nature of NGC 602 from the 3D printouts by, for example, verbally describing the structure and the spatial relationship between the star cluster, the gas, the dust and the overall extension of the cluster system.
- How they would improve both the Swell Form and the 3D printouts.

FOCUS GROUP RESULTS

Table 1

Summary of testing and surveys of NGC 602 Tactile Materials

Venue	Testers characteristics	Materials	Survey Responses	Open Feedback
NFB Convention Orlando, FL Early July, 2013	15 individuals, 20-68 years of age. HS - PhD. One college student--Physics and Astronomy; One PhD in Astronomy. ~50% with prior tactile/3D materials experience.	Textual description (read aloud). Visual image (for sighted participants). Swell Form composite (8.5 x 11 inches). *Texture* map (small format). *Elevation* map (small format).	Swell Form Textures identifiable by 100%. 3D map textures: stars identifiable by 100%. Other features identifiable by half the participants, -gas+dust most difficult. Gas (dotted slant lines). Dust (horizontal lines). Gas+Dust (triangles). Swell Form helped understand 3D print.	3D map textures harder to discern. Regions should have clear boundaries. 3D map size does not match Swell Form and is too small. Color for individuals with partial vision would be helpful.
NFB STEM-X Towson, MD late July 2013	22 individuals. 12 boys, 10 girls, 14-18+ years of age. High school. At least 13 with tactile experience and 3D, 5 with neither.	Textual description (read aloud). Visual image (for sighted participants). Individual Swell Forms for each feature (gas, dust, stars, etc.). Swell Form composite (8.5 x 11 inches). *Texture* map (small format). *Elevation* map (small format).	Stars (circles) easiest to identify. Dust (horizontal lines) by half the participants. Gas (dotted slant lines) and gas+dust combination (triangles) least identifiable. Composite and individual Swell Forms helped understand concepts and 3D print textures.	3D maps too small. Regions need to have clear boundaries or edges. Put a texture key near or on the 3D printout. Use color if possible.
NFB State Convention Connecticut Early November 2013	11 individuals. Adults.	Textual description (read aloud). Visual image (for sighted participants). Textures 3D legend. *Texture* 3D map (large format). *Elevation* 3D map (large format).	Stars (circles) easiest to identify on both maps. Dust (horizontal lines) identifiable on both maps by 100%. Gas (soft stipple pattern) identifiable on both maps by 9 individuals. Gas+dust combination (wide stipple) identifiable by 100%.	Features “easier to distinguish” – some preferred *elevation* map, some *texture* map, some had no preference. 100% could identify brightest features on elevation model Some liked both *texture* and *elevation* to build cluster mental model .
NFB State Convention Maryland Late November 2013	40 individuals. 10-81 years of age. Grade school through PhD. Some with science backgrounds. Some teachers. Widely varied experience with Braille. Some with 3D tactile experience.	Textual description (read aloud). Visual image (for sighted participants). Textures 3D legend. Texture 3D map (large format). Elevation 3D map (large format).	Stars (circles) easiest to identify on both maps. Dust (horizontal lines) identifiable on both maps by 100%. Gas (soft stipple pattern) identifiable on both maps by most. Gas+dust combination (wide stipple) identifiable by all but 2 individuals.	Features: Easy to distinguish for most, preferred legend *on* prints, not separate. Dust horizontal-line texture was “non-intuitive” but distinguishable. 100% could identify brightest features on elevation model. Prior Braille or tactile experience: most felt comfortable using legend and elevations. Less Braille or tactile experience: Some liked both texture and elevation.

First Iteration - Focus Groups

Testing results are summarized in Table 1. For the very first set of prints, we used a half-sized print on the 3D printer because 3D printing is very time consuming (several hours), and we had early opportunities to have our materials examined by focus groups. The first two groups of testers used materials with the textures demonstrated in Figure 2a. One group attended the 2013 National Federation of the Blind Convention (NFB) in Orlando, and the second group attended the 2013 NFB STEM-X Resource Fair in Maryland aimed at high school students.

At the NFB convention, 15 blind and visually impaired individuals with a range of ages (20-68 yrs) who were participants at a Science and Technology Division meeting volunteered as testers. The background and professions of the individuals in the division made it easy for them to understand the project idea and basic concepts for the testing. Several individuals were fami--liar with astronomy. Many but not all of the participants had experience with tactile materials. Generally, participants could identify the circular symbols depicting stars on all materials. The dust and gas features, represented respectively by thick horizontal lines and slanted dotted lines, were identifiable on the Swell Form, but less than half of the participants perceived them easily on either the 3D printed *texture* or *elevation* maps. The gas plus dust combination, represented by triangles, was more problematic as 75% of the testers could identify the texture on Swell Form but only 30% on either 3D printout. The *elevation* maps received mixed reviews – but participants did understand that the elevation represented brightness and could distinguish what features were dominant. Participants appreciated using the Swell Form as a starting point. However much of the problem with the textures was ultimately traced to the fact that our initial 3D prints were half the size of the Swell Form piece of paper, thus severely limiting usability.

At the STEM-X event, there were 20 high school students who participated in the evaluation. STEM-X is a program offering opportunities for inquiry-based learning, so this audience was already interested in science. At this initial stage, interacting with a somewhat science literate audience was helpful so that we could concentrate on the texture and printout evaluation rather than "teaching astronomy" *per se*. The materials used by the high school students were slightly different than for the Convention group as a result of that first group's feedback. Rather than having evaluators compare the 3D prints to a single composite Swell Form page, the image was deconstructed into a series of individual Swell Form pages, to better familiarize the testers to each unique texture. In this evaluation, successive Swell Form images represented an individual texture, culminating with the composite image. This strategy greatly helped the participants to distinguish individual features on the combined Swell Form and 3D print materials. Nevertheless, it appeared the small format of the half-sized 3D prints caused most of the difficulty with the *texture* and *elevation* maps.

The major results were that, while enthusiastic about the concept of the 3D prints, both groups made it clear that a size similar to an 8x11-inch piece of paper would be more appropriate to use, and would nicely match the Swell Form illustrations. Also, using slanted and horizontal lines in the same print in our first textures (seen in Figure 2a) was considered too confusing. As a result of this feedback and before we generated a larger set of prints (which take 12-15 hours to produce), we made test pieces with various roughness' and texture spacings as well as design elements such as pattern heights, thicknesses, and textures relative to the substrate. These objects were examined informally by various individuals mostly associated with local NFB chapter members who were keen to have an opportunity to provide feedback. We also found that larger prints had to be produced in pieces because the printer stage was too small to fabricate the desired (approximately 8x11-inch) size. This is relevant for any individual who wishes to print these on their own, or in large quantities, as some assembly is required before use.

Second Iteration -Focus Groups

We tested the 3D *texture* and *elevation* maps at the 2013 NFB Connecticut and Maryland State Conventions using nearly the same procedure as the initial testing, and results are also presented in Table 1. The new textures (Figure 2b and Figure 3) symbolized stars as circles with an indentation resembling a tiny bowl, to distinguish them from other raised patterns. Gas was represented as a dotted pattern and dust as horizontal lines, while the dust + gas combination was a roughened dotted texture. We included a 3D-printed tactile legend separately from the maps, and this seemed to help greatly by eliminating the necessity (according to the participants) of using the Swell Form illustrations as a texture introduction. Participants examined the legend and *texture* map to gain familiarity with the shape and overall composition of NGC 602's features, and reported whether they could locate and distinguish stars, gas and dust. The 3D *elevation* map was examined by each person subsequently, and then they answered the survey questions. The participants in Connecticut decided immediately that they did not need to use the Swell Form print at all so we quickly abandoned its use. Since the new 3D printouts were larger format and higher resolution, the users felt that the *texture* map was redundant to, and with more distinctive textures, than the Swell Form.

The Connecticut test group consisted of 11 people who were born blind, or somewhat visually impaired, and a number of people who lost their sight later in life. Nearly 100% of the participants easily identified textures on the *texture* map 3D printout. Most people could also identify the areas on the *elevation* model. For this test group, we found that participants who lost their sight later in life preferred the flat *texture* map, or preferred to use it before the *elevation* map, and thought the texture for gas was harder to find in the elevated model. Conversely, participants who were tactile readers throughout their life often preferred the *elevation* model, and remarked that in the future they would not need a *texture* map. All agreed that having both versions of the tactile material offered non-visual access to the broadest audience.

The Maryland test group included a few staff from the National Federation of the Blind headquarters. This small subset of participants had more experience using tactile graphics and often chose the elevated model as their favorite. A total of 40 participants were surveyed at the Maryland Convention, and were a range of ages (from 10 to 81 years) with a wide range of backgrounds, some with science, technical, and teaching backgrounds, and others from other professions. In addition to asking survey questions, we videotaped the interviews with participants to get their impressions.

ANALYSIS

The results from the focus groups indicated that once we produced larger format 3D prints and improved textures, nearly all participants could find textures on the *texture* map and nearly everyone could identify the textures on the *elevation* map as well. As mentioned above, individuals who had been blind from birth, or for a very long time, tended to be more familiar with Braille and had more experience with tactile materials, and therefore preferred to use the *elevation* model at the outset. Other individuals felt that the legend, the *texture* map, and the *elevation* map were most useful when used in sequence. Also, interestingly, those very familiar with tactile materials thought that "others with less experience" would benefit from the legend and *texture* maps. These skilled individuals had spent time in the company of people who needed some mentoring with tactile materials, so they were able to see the value of the progressive use of more complex materials.

Most Maryland respondents provided numerous additional comments and suggestions. Regarding the textures, a common comment was that the texture for the dust should be more intuitive and evocative (horizontal lines were hard to understand and remember, and did not evoke the mental image for "dust"). The testers felt the tight rougher triangular pattern was appropriate for the dust. The gas pattern was softer, and the respondents reacted favorably to its texture. Participants also felt that the texture used for the gas+dust regions should be similar to the dust but somehow also remind them of the gas. The current textures we are using, demonstrating, and informally testing reflect this input (Figure 3**)** see also https://sites.google.com/site/3dastronomy/). Creating textures that are evocative of the physical features apparently assists in stimulating, building, and reinforcing a person's mental model of the NGC 602 cluster region. Participants reported that the *elevation* map did help them understand the star cluster structure, and some individuals remarked they could "see" the cluster, a remarkable outcome .

Some participants in the Maryland group reinforced our idea that tactile materials could assist all kinds of learners. In particular, several parents and teachers remarked that the materials would be especially useful for students on the autistic spectrum, and also the home schooled. We definitely would like to broaden the applicability of the 3D materials to many groups as our project matures, and when we are able to provide a library of printable files to the public.

SUMMARY AND FUTURE WORK

Summary

We have demonstrated and tested the results of the first step in a new process for converting visual imagery and associated scientific data from Hubble Space Telescope observations to 3D tactile print products. From scientific analysis of HST data, we created tactile 3D printouts mapping the features of the celestial star-forming region, NGC 602, into textures. We tested suites of textures printed with a low cost 3D printer (a Makerbot Replicator 2) to determine which textures were most distinguishable, most useable, and feasible to print either as a flat textured image representation, that is, a *texture* map, or as undulating surfaces in an *elevation* map, representing NGC 602 features and light intensity from those textures or elevations**.** We found that a sequence of print materials--a tactile legend, a *texture* map and an *elevation* map used in succession--provides information to visually impaired individuals and can well convey the key elements of the HST imagery. The printouts helped the participants build a mental model of the star formation region being depicted.

These 3D prints will be of use to the visually impaired, and generally to anyone who finds tactile materials useful for understanding astronomy. We envision that the printouts can be used as primary materials for presenting astrophysical concepts to learners who can make use of tactile materials, supplemental resources for news releases through internet distribution of the 3D print files for printing by the user, as materials that can augment informal science displays of astronomical information, exhibits and planetarium shows, and as supplemental materials for science classes covering astronomical and science topics. Home schooling parents we interviewed suggested they would use these materials, especially if they can print objects at home. Others with special needs or autistic students that respond to tactile materials may use these to augment their instructional tools.

Limitations

During the course of this project, we found that producing 3D tactile representations of HST data for astronomical objects necessitated the use of existing software augmented with hand-tuning and custom algorithms. The CAD-type programs such as MeshLab, Google Sketchup, Tinkercad and others (http://www.3ders.org/3d-software/3d-software-list.html) suitable for production of 3D print materials can be used to represent physical objects in digital space and then print them, but they are not set up to ingest our kind of numerical data (indeed, as far as we know there is *no* CAD program that will ingest a spreadsheet). Also, although most 3D printers use the same file format, even in industry, there are specific methods and procedures that pertain to each of the types of printer being used. Print order and object orientation on the printer are important, but this is a complex issue highly dependent upon the design, printer, structure and plastic and sometimes 3D printouts need to be hand polished, painted and otherwise treated after production.

In our case, because we use a low-end printer that could be a typical device used in schools, libraries, museums, and for home use, we placed importance on experimenting with printing test pieces to assess surface quality, and accuracy. The low-end printer is not capable of the highest resolution one might desire, and other nuances arose, such as the roughness and uniformity of textures being a function of the orientation of the print piece on the printer bed, among others.

Next Steps

Our initial process for translating HST images and science measurements into 3D prints has been successful, producing materials that are usable and understandable by the visually-impaired and other tactile learners. We are currently tuning the process so that it will be possible to offer 3D print files to the public as part of the regular production of public imagery from HST. Our goal is to produce a library of such files for printing by anyone.

To utilize additional scientific data on other star clusters and celestial objects, we intend to improve the data ingest method so that scientists producing imagery and associated measurements can manipulate their data into the appropriate file formats for 3D printing. This will involve a custom interface for scientists, and a translation mechanism from that interface and science data to a program that can produce 3D print files. We will still target the lower end 3D printers as they will be more widely available to the public and educators. Eventually when we are satisfied with resolution and accuracy, we can produce test prints for higher end devices to create 3D exhibit quality products. We foresee that 3D science data from HST researchers will be useful not only for 3D printers, but also for internet 3D renditions as well as new emerging technologies such as gesture technologies (with/without gloves), heads-up technology, and others.

The next phase of 3D tactile representation development is to create distance layers, mimicking a visual fly-through from near to far, or to create a physical object that an individual can hold an accurate representation of a celestial object in their hands. The real challenge is not only to transform the scientific measurements and graphical information into a format for the printer, but also impose uniform, clear textures on curving, undulating surfaces. When we achieve production success, we will be in a good position to broaden 3D Astronomy printing and integrate these products into learning modules for both formal and informal education. As well, we will be able to create a library of STL files of HST data for use by anyone.

ACKNOWLEDGEMENTS

Support for Program number HST-EO-12915.12 was provided by NASA through a grant from the Space Telescope Science Institute (STScI), which is operated by the Association of Universities for Research in Astronomy, Incorporated, under NASA contract NAS5-26555.

REFERENCES

Beck-Winchatz, B., and Ostro, S. (2003). Using asteroid scale models in space science education for blind and visually impaired students, *Astronomy Education Review*, *2*, 118-126.

Beck-Winchatz, B., and Riccobono, M.A. (2008). Advancing participation of blind students in science, technology, engineering, and math, *Advances in Space Research*, *42*(11), 1855–1858.

Braille Authority of North America. (2012). Guidelines and standards for tactile graphics, from http://www.brailleauthority.org/tg/web-manual/index.html .

Carlson, L. (2007). Progressive star formation in the young SMC cluster NGC 602, *Ap.J. 665*, L109.

Christian, C.,A., Nota, A., Grice, N.A., Sabbi, E., Shaheen, N., Greenfield, P., Hurst, A., Kane, S., Rao, R., Dutterer, J., and deMink, S., (2014). *You can touch this! Bringing HST images to life as 3-D models*, AAS Meeting #223, #244.16.

Diaz-Merced, W. (2014). Making astronomy accessible for the visually impaired, *Scientific American*, from http://blogs.scientificamerican.com/voices/2014/09/22/making-astronomy-accessible-for-the-visually-impaired/ .

Dion, M., Hoffman, K., and Matter, A. (2000). *Teacher's manual for adapting science experiments for blind and visually impaired students*, from https://www.tsbvi.edu/attachments/1877_Manual2.doc .

Cignoni, M., Sabbi, E., Nota, A., Tosi, M., Degl'Innocenti, S., Moroni, P. G. Prada, Angeretti, L., Carlson, L. R., Gallagher, J., Meixner, M., Sirianni, M., and Smith, L. J. (2009). Star formation history in the Small Magellanic Cloud: The case of NGC 602, *The Astronomical Journal*, *137*, 3668–3684.

Frankel, F. (2004). The power of the pretty picture, *Nature Materials*, *3*, 417-419.

Grice, N., Beck-Winchatz, B., Wentworth, B., Winchatz, M. R. (2004). Touch the universe: A NASA braille book of astronomy, *ASP Conference Series*, *319*, 301.

Grice, N., and Mutchler, M. (2010). *A tactile Carina Nebula*, AAS Meeting #215, #445.08.

Grice, N. (2012a). *Everyone's universe* (second edition). New Britain, CT:You Can Do Astronomy LLC.

Grice, N. (2012b). Accessible astronomy: Astronomy for everyone. In Andre Heck (Ed.), *Organizations, People and Strategies in Astronomy (Vol.1)*, Duttlenheim, France:Venngeist, 199-205.

International Astronomical Union (2012). *Astronomy for development strategic plan.* from http://www.iau.org/education/strategic_plan/ .

Kessler, E. (2007). Resolving the nebulae: the science and art of representing M51, *Studies in History and Philosophy of Science, Part A*, *38*(2); 477-491. Also available at http://dx.doi.org/10.1016/j.shpsa.2007.03.001 .

National Research Council (2010). *New worlds, new horizons in astronomy and astrophysics*, Washington, DC:National Academies Press. http://www.nap.edu/openbook.php?record_id=12951 .

National Science Foundation (2010). *Preparing the next generation of STEM innovators: Identifying our nation's human capital*. Washington, DC:National Science Board http://www.nsf.gov/nsb/stem/innovators.jsp .

National Aeronatics and Space Administration (2014). *Senior review of operating missions: Hubble panel report,* http://science.nasa.gov/media/medialibrary/2014/05/15/Final_Report_HST2014_SeniorReview_Panel_.pdf .

Smith, L. F., Arcand, K .K,. Smith, J. K., Smith, R. K., Bookbinder, J., and Watzke, M. (2014). Examining perceptions of astronomy images across mobile platforms. *Journal of Science Communication*, in press: http://arxiv.org/abs/1403.5802v1 .

Snyder, E. (2011). The eye of Hubble: Framing astronomical images, *A Journal Of Visual And Material Culture*, *1*, 3-21, available at http://www.framejournal.org/view-article/9#_ftnref .

APPENDIX

SURVEY QUESTIONS FOR FOCUS GROUPS

Relevant Survey Questions (Tests 1 and 2 - NFB Orlando National Convention and MD NFB STEM-X):

The visual image was made available to the participants and a verbal description of NGC 602 was provided to the participants before each individual examined the materials and responded to the survey.

1. Can you locate the circular star symbols in the tactile graphic (Swell Form) and in the 3D maps?
2. Can you locate the dust texture (thick horizontal lines) in the tactile graphic (Swell Form) and in the 3D maps?
3. Can you locate the combined dust and gas (triangle) texture in the tactile graphic (Swell Form) and in the 3D maps?
4. Can you locate the gas texture (slanted dotted lines) in the tactile graphic (Swell Form) and in the 3D maps?
5. Does examining the tactile graphic (Swell Form) help you better understand the 3D maps?
6. How would you improve the tactile graphic – Swell Form?
7. How would you improve the 3D maps?
8. If you are low vision or sighted, does examining the color image of NGC 602 help you to better understand the tactile graphic and 3D model?
9. Does examining the tactile graphic (Swell Form) help you understand the 3D maps?
10. What is your experience using tactile graphics and/or 3D models?
11. Any questions or comments for the tactile designers?
12. What is your age and highest degree of education/occupation?

Relevant Survey Questions (Tests 3 and 4 – CT and MD NFB State Conventions)

The visual image was made available to the participants and a verbal description of NGC 602 was provided to the participants before each individual examined the materials and responded to the survey.

Please use the tactile texture key first so you are familiar with the textures.
Please answer the following questions. Comments are welcome at any time.

1. Do you have any experience with Braille or tactile materials?
2. Can you find the circular star symbols in the *texture* map? In the *elevation* map?
3. Can you find the dust texture (horizontal lines) in the *texture* map? In the *elevation* map?
4. Can you locate the soft gas texture in the *texture* map? In the *elevation* map?
5. Can you find the combined dust and gas (rough) texture in the *texture* map? In the *elevation* map?
6. Are the different textures harder or easier to identify on the *elevation* map as compared with the *texture* map? Why?
7. Can you detect by touch, any difference in brightness (energy output) of the different features on the *elevation* map?
8. Which particular features do you think are the brightest/most intense on the *elevation* map?
9. How would you describe the strengths and weaknesses of each tactile map of NGC 602?

SOLAR WIND MODELING: A COMPUTATIONAL TOOL FOR THE CLASSROOM

Lauren N. Woolsey, *Harvard University*

Received November 13, 2014; Revised January 30, 2015;
Accepted April 10, 2015

Abstract: This article presents a Python model and library that can be used for student investigation of the application of fundamental physics on a specific problem: the role of magnetic field in solar wind acceleration. The physics included in the model, The Efficient Modified Parker Equation Solving Tool (TEMPEST), is laid out for the reader. Results using TEMPEST on a magnetic field structure representative of the minimum phase of the Sun's activity cycle are presented and discussed. The paper suggests several ways to use TEMPEST in an educational environment and provides access to the current version of the code.

Keywords: Specialized – Graduate - Solar Physics - Technology in Education - Magnetic Fields - Python

INTRODUCTION

The Sun is a laboratory for a wide range of physics, including thermodynamics, plasma physics, fluid dynamics, and electricity & magnetism. Solar physics is a way to connect topics taught in the classroom with the real world, on a scale far greater than other example applications. One of the aspects of the Sun that is of direct interest to any courses covering magnetic fields, waves, and plasmas is the constant outflow of matter called the solar wind. The use of the solar wind as a teaching example can also highlight the scientific process and the ongoing generation of knowledge; this outflow has been studied for over half a century (Parker, 1958), yet there are still lingering mysteries. The most prominent of these open questions is the identification of the processes that generate and accelerate the solar wind. For this paper, I concentrate on the use of turbulent heating as an acceleration mechanism, which is one of the primary suggested processes to accelerate the wind (see reviews by e.g. Klimchuk, 2006, and Cranmer, 2009).

The general idea is that the upper layers of the Sun act like the surface of a pot of boiling water; energy is constantly being brought up from the lower layers through the process of convection. This convection creates a granulation pattern, and it can jostle magnetic field lines to create longitudinal sound waves and a special type of transverse wave called Alfvén waves (Alfvén, 1942). The field lines that emerge from the surface of the Sun often bundle into structures called flux tubes. When a flux tube extends outward beyond several solar radii, it is considered “open” to the heliosphere.

The model presented in the remainder of this paper, The Efficient Modified Parker Equation Solving Tool (TEMPEST), applies the physics of wave-driven turbulence to open flux tubes. TEMPEST solves a modified version of the Parker Equation (Parker, 1958) to provide the properties of the solar wind generated by a single open flux tube, and the only required input is the magnetic field profile from the flux tube. See the Appendix for information on TEMPEST.

APPLICATIONS FOR STUDENT INVESTIGATION

The majority of physics deals with topics that students often have trouble visualizing, either due to the vast scales of distance and time involved in astrophysics or the infinitesimal scales of quantum mechanics. This is especially true when dealing with magnetic fields, as scientists cannot directly observe the fields. The “invisible” nature of magnetic and electric fields can be hard to visualize for students (Herrmann, Hauptmann, and Suleder, 2000). Magnetic fields near the Sun, however, can be more easily visualized since plasma that emits in the ultraviolet traces the magnetic field lines in high-resolution images of the Sun.[1] The Sun is also an astronomical object to which students in all types of physics courses can relate, since it is easily observable and has important connections to Earth that can provide motivation for its study. For example, high-speed solar wind streams produce a greatly increased electron flux in the Earth's magnetosphere and can disrupt satellite communications and power grids on the ground (Verbanac, Vrnak, Veronig, and Temmer, 2011).

Models like The Efficient Modified Parker Equation Solving Tool (TEMPEST) make progress towards the ability to predict the nature of plasma streams that will be rotating towards Earth long before they can damage space-based equipment or create geomagnetic storms. Whether as part of a course in physics, independent study, or directed research, students can use TEMPEST to further their understanding of the role of magnetic fields in this bigger picture.

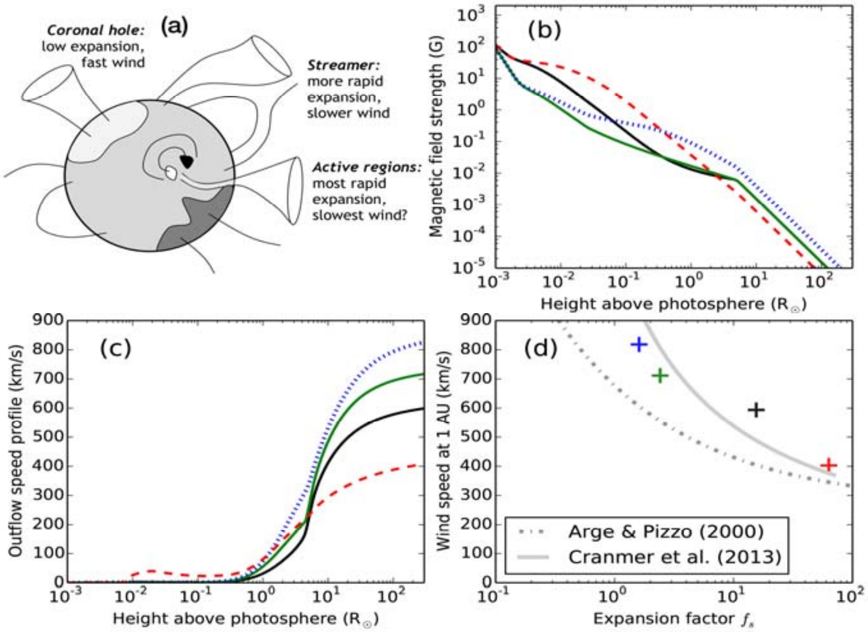


Figure 1. (a) A cartoon of possible field structures in the corona. The figure then shows TEMPEST results from four models: (b) magnetic field inputs, (c) wind speed outputs, (d) a comparison of TEMPEST with analytical relations with expansion factor.

---

[1] Free and easy access to images of the Sun can be found at http://www.helioviewer.org/

Magnetic Field Structures Throughout the Solar Cycle

The Sun goes through an 11-year cycle of high activity and low activity. During low activity periods, called solar minimum, the Sun's magnetic field looks considerably like a dipole field. The poles of the Sun during solar minimum are covered by large “coronal holes,” which are regions of open flux and lower plasma density in the corona. The equator is home to the streamer belt, where the opposite polarities of the two hemispheres join together. However, during solar maximum the magnetic field of the Sun is incredibly chaotic, with small or virtually non-existent polar coronal holes and many active regions where strong bundles of magnetic field have pushed up out of the solar interior. Figure 1a presents a sketch of some of the many types of structures.

The solar wind is present at all times, but as the Sun goes through different points of the cycle, the properties of the solar wind that reaches Earth change dramatically. Figure 1 shows four models from the results of Woolsey and Cranmer (2014). Predictions of solar wind speeds used in space weather forecasting often rely on the Wang-Sheeley-Arge (WSA) model (Wang and Sheeley, 1990; Arge and Pizzo 2000), based on a defined quantity called the expansion factor. This factor is a measure of the amount of cross-sectional expansion from the photospheric base of a flux tube to a “source surface” at a height of 1.5 solar radii above the Sun's surface, where field lines are set to be purely radial and defined as open to the heliosphere. An expansion factor of 1 refers to perfect radial expansion (i.e. an inverse-square relation of the magnetic field strength), while larger expansion factors mean more rapid “super-radial” expansion and vice versa. Several simple analytical relations between expansion factor and wind speed have been put forward. Two have been plotted in Fig. 1d; one maps the outflow speed at the source surface. The points all lie above it because the wind continues to accelerate above this height (Arge and Pizzo, 2000). The other relation was defined to match results from the model on which TEMPEST is based (Cranmer, van Ballegooijen, and Woolsey, 2013).

Students can use TEMPEST to investigate the many different magnetic field profiles that might appear throughout the solar cycle. Similar to the analysis by Woolsey and Cranmer (2014) on a large grid of synthetic models, students could investigate how different magnetic field profiles of their own creation can lead to varied solar wind solutions, and how well the WSA model holds for a variety of structures.

Dependence on Temperature Profile

I mentioned earlier that TEMPEST does not solve the energy conservation equation, because it has set up temperature profiles based on the results from ZEPHYR. Students could investigate how a different temperature profile, due to possible other sources of heating, would affect the solar wind. Using the function in TEMPEST called Miranda, which does not include the wave pressure term, and a different temperature profile, students could accurately use the momentum conservation equation for different coronal heating sources, including mechanisms that do not use wave-driven turbulence. Figure 2 shows a quick study of changes to the default TEMPEST temperature profile and profiles that reach higher or lower temperatures at large heights. As Parker (1958) originally demonstrated, a hotter corona generally produces a faster wind.

Students can investigate many aspects of stellar winds. How hot does the corona need to be in order to have a steady-state supersonic solar wind? This also allows TEMPEST to be used for other stars, if the gravitational term in the momentum equation was also modified based on the mass and radius of the star. While this goes beyond the scope of most physics courses, it could be a topic of independent study or undergraduate research.

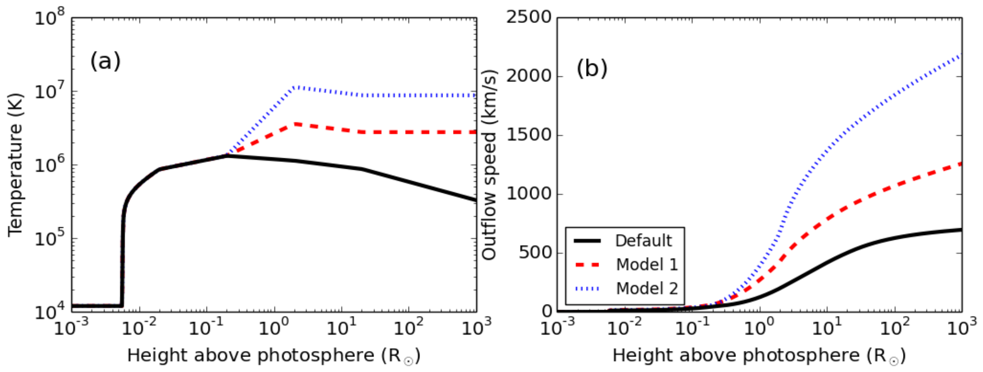


Figure 2. The relation between (a) temperature profile and (b) outflow speed follows Parker's (1958) original theory**.**

Comparing Observations to Models

A larger undertaking would be to use TEMPEST to compare real observations to the model output, exposing students to the process of science validation. The magnetic field profile would need to be determined from extrapolations of magnetograms, a map of magnetic field strength and polarity on the Sun's surface. This is most easily achieved using the solar models that can be obtained from the Community Coordinated Modeling Center.[2] Once the magnetic field profiles of Earth-directed flux tubes are run through TEMPEST, predictions from the model of wind speed, density, and temperature at 1 AU can be compared to in situ measurements by the array of solar spacecraft available. The Space Weather Prediction Center (SWPC) provides a wealth of resources to access available observations and measurements.[3]

CONCLUSIONS

TEMPEST models the steady-state solar wind from open flux tubes, using only the magnetic field profile of the flux tubes as input. While the topic of the model itself is specialized, the physics involved is covered in many basic undergraduate courses. I have discussed the fundamental physics used in TEMPEST (for further detail, see Woolsey and Cranmer, 2014; for a list of general solar physics resources, see the resource letter by Pasachoff, 2010), presented the use in an example coronal hole, and provided a few possible student applications of full code and of the associated Python library.

While I have given several possible uses of TEMPEST for student projects, the beauty of science is that success can often be found at the end of a long and widely-branching path. There are likely countless other ways in which TEMPEST can be used for teaching and exploring interesting science; I hope by making it publicly available on GitHub it can find such new uses. I have introduced TEMPEST in this paper, and I intend to work with students on practical applications of it in the near future. Those results will be published here in a follow-up paper.

ACKNOWLEDGEMENTS

The author gratefully acknowledges and sincerely thanks Steven R. Cranmer generally for his incredible mentorship and stellar advising over the past four years and specifically for Fig. 2a and helpful comments and suggestions for this paper. This material

---

[2] Community Coordinated Modeling Center: http://ccmc.gsfc.nasa.gov

[3] Space Weather Prediction Center: http://www.swpc.noaa.gov

is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1144152 and by the NSF SHINE program under Grant No. AGS-1259519.

Lauren Woolsey, an NSF Graduate Fellow, can be reached by email at lwoolsey@cfa.harvard.edu .

REFERENCES

Alfvén, H. (1942). Existence of Electromagnetic-Hydrodynamic Waves. *Nature, 150*(3805), 405–406.

Arge, C. N. and Pizzo, V. J. (2000). Improvement in the prediction of solar wind conditions using near-real time solar magnetic field updates. *Journal of Geophysical Research. 105*, 10465-10480.

Cranmer, S. R. (2009). Coronal Holes. *Living Reviews in Solar Physics, 6*(3).

Cranmer, S. R., van Ballegooijen, A. A., and Edgar, R. J. (2007). Self-consistent Coronal Heating and Solar Wind Acceleration from Anisotropic Magnetohydrodynamic Turbulence. *Astrophysical Journal Supplement Series, 171*(2), 520-551.

Cranmer, S. R., van Ballegooijen, A. A., and Woolsey, L. N. (2013). Connecting the Sun's High-resolution Magnetic Carpet to the Turbulent Heliosphere. *Astrophysical Journal, 767*(2), 125.

F. Herrmann, H. Hauptmann, and M. Suleder (2000). Representations of electric and magnetic fields. *American Journal of Physics, 68*, 171.

Hundhausen, A. J. (1972). *Coronal Expansion and Solar Wind.* Springer-Verlag:Berlin.

Jacques, S. A. (1977) Momentum and energy transport by waves in the solar atmosphere and solar wind. *Astrophysical Journal, 215*, 942-951.

Klimchuk, J. A. (2006). On Solving the Coronal Heating Problem. *Solar Physics, 234*(1), 41-77.

Parker, E. N. (1959). Dynamics of the Interplanetary Gas and Magnetic Fields, *Astrophysical Journal, 128*, 664-676.

Pasachoff, J. (2010). Resource Letter Sph-1: Solar Physics, *American Journal of Physics, 78*, 890.

Verbanac, G., Vršnak, B., Veronig, A., and Temmer, M. (2011). Equatorial coronal holes, solar wind high-speed streams, and their geoeffectiveness. *Astronomy and Astrophysics, 526*, A20.

Wang, Y.-M. and Sheeley, Jr., N. R. (1990). Solar wind speed and coronal flux-tube expansion. *Astrophysical Journal, 335*, 726-732.

Woolsey, L. N. and Cranmer, S. R. (2014). Turbulence-driven Coronal Heating and Improvements to Empirical Forecasting of the Solar Wind. *Astrophysical Journal, 787*(2), 160.

APPENDIX

OVERVIEW OF TEMPEST

Background Physics Included in the Model

TEMPEST was designed to be a simpler version of a turbulence-driven coronal heating model called ZEPHYR (Cranmer, van Ballegooijen, and Edgar, 2007; Woolsey and Cranmer, 2014). While ZEPHYR solves the mass, momentum, and energy conservation equations of magnetohydrodynamics, TEMPEST uses extensive calibration from a large grid of models to boil down the problem to only the momentum conservation equation (Woolsey and Cranmer, 2014). In fluid dynamics, this equation in spherical coordinates is $\partial u/\partial t + u(\partial u/\partial r) + (1/\rho)(\partial P/\partial r) = -GM_{\odot}/r^2 + D$ (Equation 1) where u is the outflow velocity, ρ is the mass density of the fluid, P is the pressure, and the right-hand side represents the effects of gravity and waves. The time-averaged effects of waves, D, has separate terms for Alfvén waves and acoustic waves (Jacques, 1977). Combining the time-steady momentum conservation equation, Equation (1), and mass conservation yields the equation of motion, which represents a modified version of the Parker equation: $[u - (u_c^2/u)](du/dr) = -GM_{\odot}/r^2 - u_c^2(d\ln(B)/dr) - a^2(d\ln(T)/dr) + Q_A/(2\rho(u + V_A))$ (Equation 2). This form drops terms that included only the acoustic waves, as they do not contribute a significant amount of coronal heating and outflow acceleration (Cranmer et al., 2007). In Equation (2) above, B is the magnetic field profile, T is the temperature profile, a is the sound speed, Q_A is the Alfvénic heating rate, V_A is the Alfvén speed, and u_c is the critical speed, whose radial dependence is defined by $u_c^2 = a^2 + U_A/(4\rho)[(3u + V_A)/(u + V_A)]$ (Equation 3), where U_A is the Alfvénic energy density. The Parker "critical point" is defined by the height where the wind transitions from speeds below this critical speed u_c to speeds above it. This is similar to the sonic point, where the wind is subsonic ($u < a$) below and supersonic ($u > a$) above the sonic point.

Typically, one must solve the equation of motion self-consistently with an internal energy conservation equation to obtain the temperature profile T. TEMPEST avoids this because it builds profiles of the temperature and wave heating rate Q_A from strong correlations between the magnetic field profile provided by the user and these variables (Woolsey and Cranmer, 2014). However, it is possible to determine more accurate profiles for the temperature and reflection coefficient. Such a study is discussed in the Applications section. I now turn to direct application of the physics I have just discussed.

Results from an Example Coronal Hole Flux Tube

The following example of an open flux tube is based on the modified solar-minimum magnetic model used by Cranmer et al. (2007). This magnetic field profile is shown in Figure S1a. The strength at the photosphere reaches as high as 1.4 kG (0.14 T), because open magnetic field arises from the extremely narrow lanes between granulation cells. For a radially expanding magnetic flux tube, conservation of magnetic flux requires that the field strength decrease as the tube's cross-sectional area increases. The field strength decreases with height, eventually following an inverse square law.

TEMPEST is uniquely capable of separating the two main effects of waves on the solar wind: 1) heating of the corona through a turbulent cascade of energy from large scales to small scales, where the energy can dissipate as heat; and 2) the additional momentum term provided by the outgoing waves themselves. This is shown in Figure S1b. The reason TEMPEST can separate these effects is the calibration of the temperature profile from ZEPHYR, which by design includes the heating done by sound waves and Alfvén waves already. The Python library is set up, however, to allow the user to submit a different temperature profile if desired (see Applications section).

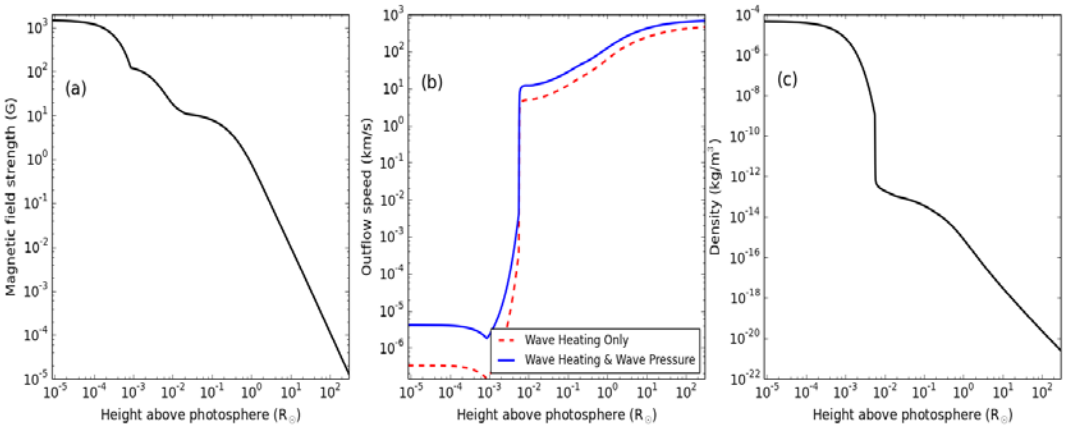


Figure S1: Results of TEMPEST using (a) the magnetic field profile of an open flux tube. TEMPEST provides (b) the outflow speed and (c) density profiles of the solar wind from this flux tube.

The current version of TEMPEST saves four files in the numpy ".npz" format throughout the calculation. They are named automatically based on functions within the TEMPEST library, with a prefix that can be specified by the user (default is T):

- T_inputs.npz saves the height profiles, magnetic field profile, and temperature profile.
- T_miranda.npz saves the outflow profile without the wave pressure term, the sound speed profile, the right-hand side of Eq. (2) without the Alfvén wave term, and the sonic point.
- T_fullRHS.npz saves the density profile, Alfvén energy density profile, Alfvén wave amplitude, critical speed profile (see Eq. (3)), and the full right-hand side of Eq. (2)
- T_prospero.npz saves the final outflow solution (the blue curve in Figure S1b) and the Parker critical height.

I wrote the code in Python because it is by nature completely modular. It is also conveniently open-source, so that students do not need to pay for costly software licenses. Within the TEMPEST library, there is a main() function that automatically reads in the magnetic field profiles, runs all steps of the code, and saves the above files along the way. However, the library also contains functions that can be used on their own, given the proper inputs. Students can determine the goals of the individual functions included in the code, because each defined function is documented with a) a short description of the function, b) a listing of its inputs and outputs, and c) any required outside packages (e.g. numpy). The code is available on GitHub[4] which allows clear version control and the ability for others to contribute to the code.

---

[4] TEMPEST GitHub Repository: http://github.com/lnwoolsey/tempest

Instructions for Authors

The *Journal and Review of Astronomy Education and Outreach (JRAEO)* seeks to publish the highest quality articles that advance the understanding of astronomy teaching and learning and are likely to have a significant impact on the field of astronomy education and outreach. *JRAEO* welcomes manuscripts from authors around the world and from all theoretical frameworks and research methodologies.

Types of Manuscripts

You may submit a variety of materials for publication in *JRAEO*, provided they are not currently under consideration by other journals, nor previously published. Manuscripts can be in but not limited to any of the following type:

Research: All articles in *JRAEO's* Section A are peer-reviewable scholarly research. We accept quantitative, qualitative and mixed methods research articles. All undergo rigorous double-blind reviewing and are held to the same high standards for sample sizes, statistics, inference making and so on.

Articles should discuss the importance of the study to the field. A relationship between the study and previous work is generally needed. Methodology should be reported concisely.

All other articles are placed into *JRAEO*'s Section B. These include:

Commentary [Hypothesizing]: Invited or unsolicited 'white papers' on a topic of interest to the astronomy education, formal or informal, communities.

Ideas [In the Astronomy Education Universe]: These are items that may be of value to teachers or to those seeking research ideas, but are not necessarily peer-reviewable materials. Among these are ideas that have been tried out in classrooms as in-house experiments that could be more rigorously investigated later.

General features: Similar to those found in the editor's column [Field Observations] that take a more informal look at some item of interest to the community.

Other types of articles may be accepted.

While many articles have practical considerations for a teacher, *JRAEO* does not currently accept most practitioner articles, such as how-to-do or resource list stories. These may be submitted to a Hermograph sister publication, *The Classroom Astronomer* (www.classroomastronomer.com) .

Language

The language of the journal is English. We accept articles using both American and British English conventions. Articles from non-native English writers are also accepted, in English, though we request that all possible attempts to perfect the writing be done before it is submitted. We will do our best to make linguistic corrections unless doing so effectively and substantially rewrites the text. An additional abstract in the other language is welcome.

Formatting and Submission

Articles should be pre-formatted in APA style, in any version as far back as version 5. We have some "local" rules as well. The list of these is maintained on the *JRAEO* website, but include such things as having two spaces after a full stop (a period, question mark, exclamation point, or colon), certain subheadings which can be bold, and a space follows a URL and then the period, etc.

Articles can be submitted at any time by emailing them in a zipped, archive file to JRAEO@toteachthestars.net. An Article Submission form, also found on the website, which contains contact information, preferred keywords, and other information, must be submitted at the same time.

Being an all-digital publication, there are no size limitations in terms of word count or page numbers. However, you have a better chance of a quicker acceptance if the author 'writes tight.' Superfluously worded articles may be sent back for rewriting even before the peer review process begins. Implications, backed by supporting evidence, are paramount.

Data tables should be easy to read and with all, and only, appropriate information. Color images (in graphs, graphics, or photographs) are welcome. All figures and tables should be sent in separate files though they may be also included in the text for help in final placement.

In addition to the instructions listed on the Information for Authors page, changes and current information can be found on the website at www.jraeo.com. If the first author is a subscriber, there is a reduced page charge/open access fee. Being a subscriber is *not* to be construed as a right to be published. To reduce the page fee, *the first author* must be a subscriber *before* the article is submitted.

Review

All Section A articles are reviewed in an anonymous double-blind peer process by at least two reviewers. Some Section B pieces are reviewed double-blind by one reviewer. The editors often raise questions, and partially edit the submission both before and after the peer review process in order to improve the submission's quality and its chances for acceptance.

As a journal with a publication schedule, and thus, deadlines, we try to put accepted articles into the next issue, but this is not guaranteed. We also try to have articles reviewed and sent back to the author(s) within about 10 days from submission. Authors need to make corrections and return the revised manuscript in no more than a week to ten days after receiving the reviewers or editor's corrections. Late returns can mean a change in publication schedule of the article, or rejection of the article.

DR. LAWRENCE KRUMENAKER is the Editor of *JRAEO*. Dr. Krumenaker has deep roots in the astronomy education community, having been an astronomy and physics instructor since the 1980s at various universities and colleges. His Ph.D. was in science education from the University of Georgia. The 2008 dissertation was entitled "The Status and Makeup of the Modern U.S. High School Astronomy Course in the Era of No Child Left Behind," which generated three articles in *Astronomy Education Review*, and others in *The Science Educator* and *The Planetarian*. His current research involves astronomy education in other countries.

Krumenaker has worked in journalism since the late 1980s, having started at ABC's *20/20* show as a researcher of science stories. He has been published in numerous astronomy, science, and computer magazines, and is currently the publisher and editor of *The Classroom Astronomer* magazine. Krumenaker is a member of the National Association of Science Writers and the International Science Writers Association. As an astronomer, he does stellar spectroscopy, from time to time, of cool and emission-line stars. One of his key research results was the discovery of the Milky Way's only microquasar, SS433.

DR. KRISTINE LARSEN is the Assistant Editor of *JRAEO*. Astronomy education and outreach have been the central focus of her decades-long career, beginning in 1981 when, as a college freshman at Central Connecticut State University, she began working as a planetarium lecturer and observatory assistant. While completing a Ph.D. in Theoretical Physics at the University of Connecticut, she returned to her alma mater as a faculty member in 1989, where she currently holds the rank of Professor in the Physics and Earth Science Department and serves as Faculty Coordinator of the Copernican Observatory and Planetarium. She teaches courses in introductory astronomy, observational astronomy, and astrophysics, as well as courses for pre-service teachers and interdisciplinary courses for both the University Honors Program and First Year Experience Program.

Larsen's scholarly work focuses on the intersections of science and society, especially science pedagogy and outreach, the history of women in science, misconceptions of science, and depictions of science in literature and film.

Journal and Review of Astronomy Education and Outreach

April 2015 Issue 3
Volume 1 Number 3

Section A – Research

Section B – General Articles

Hermograph Press LLC, Marietta, GA 30066 USA
www.jraeo.com