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

omy Education Keview

Volume 3, Oct 2004 - Apr 2005 Issue 2

What Are Essential Concepts in "Astronomy 101"? A New Approach to Find Consensus from Two Different Samples of Instructors

by **Michael Zeilik** University of New Mexico **Vicky J. Morris-Dueer** University of New Mexico Received: 11/10/04, Revised: 02/15/05, Posted: 03/22/05

The Astronomy Education Review, Issue 2, Volume 3:61-108, 2005

© 2004, Michael Zeilik. Copyright assigned to the Association of Universities for Research in Astronomy, Inc.

Abstract

In the summers of 1997, 1998, and 1999, we gave attendees $(N = 44)$ at a workshop called Teaching Astronomy Conceptually a cognitive task: to rank 200 concepts often taught in "Astronomy 101." Prior to these workshops, we asked an expert panel $(N = 18)$ of Astronomy 101 teachers to also rank these concepts. Among the workshop participants, the electromagnetic spectrum ranked the highest; among the expert panel, mass held the top spot. We then requested the expert panel to perform a cognitive task of judging the relatedness of pairs of terms, and ranked the results based on concepts that were most frequently chosen. We conclude that there is reasonable consensus about essential topics in Astro101 that can be reached using ranking and relatedness tasks.

1. INTRODUCTION

What do "Astronomy 101" (hereafter, Astro 101) instructors teach in their one-semester courses that "cover the Universe"? Clearly, such a cosmic goal is impossible. The implicit sense in the community is that a consensus cannot be reached because each instructor has his or her "pet" topics. This attitude has resulted in little guidance for faculty, especially because most Astro 101 instructors do not have a degree in astronomy, and many are teaching the class for the first time (Fraknoi 2001; Zeilik 1997).

Astronomy textbook publishers would be overjoyed to have such a list; their market research reinforces the notion that the books include many topics that any one instructor may not choose to teach. The marketing goal is to offer a range so that instructors may make selections for their courses, and no crucial

topic is left out that might result in the loss of a large adoption. Nothing gets the attention of an editor more than a sales representative calling about a large adoption and asking if Zeilik's book (*Astronomy: The Evolving Universe*) includes orbital resonances in the asteroid belt (it does not). Then the author is under extreme pressure to add that topic just for that one adoption! My experience since 1975 is that peer reviewers are quick to add topics but extremely reluctant to cut them out. This dynamic again results in books with pretty exhaustive (and similar) topical coverage, but does not converge on a core set of concepts.

This article presents an innovative process, based on cognitive ranking and relatedness ratings tasks, to reach a consensus from two different sample populations. This method can be applied to larger populations to reach a much-needed community consensus.

2. PRIOR EFFORTS

We searched three areas for research into the issue of what is taught or what should be taught in Astro 101:

- 1. Grades 9-12 curriculum content standards in the United States
- 2. Groups of individual instructors of Astro 101
- 3. Efforts by professional organizations

The 1996 National Science Education Standards (NSES) proposed by the National Research Council aggressively outline the rules for effective classroom instruction, age-appropriate guidelines for curriculum materials development, authentic assessment procedures, and professional development programs for teachers (National Research Council 1996). These goals were articulated in *Science for All Americans* (1989, 1990), the *Benchmarks for Science Literacy* (1993), and *Atlas of Science Literacy* (2001) for Project 2061 of the American Association for the Advancement of Science.

In addition to the emphasis on describing science through unifying concepts and processes, the NSES provide specific learning content objectives. For K-12 astronomy, the NSES suggests 11 major astronomy objectives. These objectives include (1) describing the objects and motions of the sky (grades K-4), (2) the characteristics of gravity and the Solar System (grades 5-8), and (3) the origin and evolution of stars, galaxies, and the Universe (grades 9-12). These objectives are found in both the Earth/space science content strand and in the themes of unifying concepts and processes (Adams & Slater 2000; Slater 2000). The connections among these concepts were "back mapped" in the *Atlas of Science Literacy* (2001). Table 1 in Slater (2000) summarizes these for each grade grouping; grades 9-12 are most relevant to Astro 101. However, to the best of our knowledge, Astro 101 instructors have not inspected these topics. Our sense is that the college instructors are largely ignorant of these recommendations.

Slater et al. (2001) took another approach in two phases. First, they asked faculty preregistered for a workshop on teaching introductory astronomy to provide three main goals. Second, they examined 37 course syllabi and compared their content with 67 topics extracted from popular textbooks. In their Figure 2, Slater et al. gave the topics most commonly taught; we include these in Appendix A. We note that these are not very fine grained, but give a global indication of conceptual content.

The American Astronomical Society (AAS) took a different method (Partridge & Greenstein 2003). It convened two workshops for chairs and other department leaders from selected major research universities. Their rationale for this choice was that if any departments play a leadership role in the astronomical community, it is these research institutions. If their teaching practices are systemically reviewed and improved, it is more likely that those in two- and four-year colleges will follow. In addition, writers and publishers of textbooks are more likely to pay attention to systemic reforms at such large and influential institutions. (This assumption may not be true. For example, Arrny, Zeilik, Chaisson, Seeds, Pasachoff, Fix, Moche, and Fraknoi do not teach at prestigious Research I institutions.) The workshops involved three dozen participants from 30 institutions. The content goals are even less fine grained than those found by Slater et al. (2001); see Appendix B.

Lippert & Partridge (2004) adopted 13 of the AAS goals verbatim (see their Table 1) for a small ($N = 38$) Astro 101 class at Haverford College. At the end of the course, they asked the students to rate the 13 goals on a Likert scale from 1 (*not important or effective*) to 5 (*very important or effective*). The class mean was 3.3 ± 0.2 , just above midpoint. The students ranked goals that focused on the broad understanding of astronomy and on building a positive self-image as scientists. They also were requested to rank how well the course met these goals, and generally rated the effectiveness in meeting these goals higher than their interest in the goals. The #1 importance of a "cosmic perspective" matched its success, an indication that overall, the course matched student expectations. However, we find it hard to generalize from such a small, select sample.

3. PROCEDURE AND PROCESS

As a key part of the Conceptual Astronomy Project at the University of New Mexico (Zeilik et al. 1997; Zeilik, Schau, and Mattern 1998, 1999), we wanted to design the course around community-based core concepts. The following list outlines what we did to uncover these concepts.

- 1. We extracted concepts from Zeilik's textbook, *Astronomy: The Evolving Universe* (6e), using the glossary (roughly 400 concepts; about 20 per chapter).
- 2. We used three astronomy colleagues at the University of New Mexico, and Hirsch, Kett, & Trefil (1993) to reduce these concepts to about 200. Specifically excluded was astronomical technology (e.g., telescopes). Most of the reduction came from subsuming concepts under broader ones.
- 3. These 200 concepts were sent to an expert panel of 18 Astro 101 instructors to rank from 1 (*highly essential*) to 5 (*not at all essential*).
- 4. From these were extracted 120 concepts that had the lowest means and standard deviations.
- 5. The expert panel then appraised the relatedness of pairs from the 120 concepts, and we then ranked the concepts by how often each was related to others on the list.

The expert panel work was started in summer 1994 and completed in fall 1995. We then took the 200 concepts (Appendix C) and requested that participants in an American Association of Physics Teachers (AAPT) workshop on teaching astronomy conceptually also complete the ranking task during AAPT summer meetings in 1997, 1998, and 1999. Of these attendees, 44 completed the task on a volunteer basis. Appendix D gives the complete results. We summarize them in Tables 1 and 2, which give the workshop participants' ratings of the top 10 and bottom 10 concepts.

For comparison, we present in Tables 3 and 4 the top 10 and bottom 10 concepts ranked by the Astro 101 expert panel. We note that these results are much more fine grained than in any previous survey. Keep in mind that these two samples were very different; the expert panel consisted exclusively of people who identified themselves as astronomers and had extensive experience teaching Astro 101. In contrast, the AAPT workshop sample had a 50-50 mix of novice and expert instructors, with only one quarter identifying themselves as astronomers.

The overall mean for the AAPT workshop participants was 2.42 ± 0.59 (*SD*), which is about a half step higher than the *moderately essential* midscale judgment. For the expert panel, the mean was 2.19 ± 0.62 (*SD*), an indication that the experts judged more of the concepts to fall on the more positive end of the scale. Note that the spreads in the means are almost the same for the two very different groups. Also note from the standard deviations that more consensus occurred for the *highly essential* end of the scale.

Table 1. Top 10 Concepts Ranked by AAPT Workshop Participants ($N = 44$).

Table 2. Bottom 10 Concepts Ranked by AAPT Workshop Participants ($N = 44$).

Table 3. Top 10 Essential Concepts Rated by the Astro 101 Expert Panel (*N* = 18) for the UNM Conceptual Astronomy Project.

Table 4. Bottom 10 Essential Concepts Rated by the Astro 101 Expert Panel (*N* = 18) for the UNM Conceptual Astronomy Project.

Note that in both groups, no concept had a rating below 4.0. Survey participants were unwilling to give a rating of 5 (*not at all essential*). We expect that 5 would be a reluctant choice given the origins of the list, and this result indirectly supports its validity. Also note that the standard deviation of each group was equivalent, indicating that the overall spread in the responses was pretty much the same for each sample. Comparing the top 10 concepts from each group, we find a 50% overlap, with the following concepts in common: electromagnetic spectrum, H-R diagram, distances, mass, and atoms.

A commonly computed standard error is the *standard error of the mean*, (*SE*), defined as (population standard deviation)/SQRT(N), where *N* is the sample size. We note that the *SE* for the workshop results is 0.09, and for the expert panel, 0.15. We use this *SE* value to compare rankings from different populations, here the AAPT participants and the experts. For two concept rankings to be statistically different at a 95% confidence level, they must differ by two standard errors, or 0.18 for AAPT and 0.30 for experts.

From the expert panel, we sorted the responses by the lowest means and the smallest standard deviations. This sort resulted in 120 essential concepts, which are given in Appendix E and listed by mean (highest to lowest) from the expert panel data. The expert panel included recipients of teaching awards, members of AAS education committees, and people known to the first author to be dedicated instructors of Astro 101.

Are these *concepts*? One panel member complained that these terms are just vocabulary words. In isolation, that is correct. What counts, in a cognitive sciences sense, are the *relationships* among the terms—their connections. This statement is the underlying structural principle: that to be knowledgeable in a field entails understanding relationships among important concepts (see Hirschfeld & Gelman 1994

for a general description, and Chi, Feltovich, & Glaser 1981, for the pioneering study in physics).

As stated in Zeilik et al. (1997), the key question is, How do we represent and assess knowledge in a particular domain? The cognitive sciences have attacked this issue directly in the past two decades. From semantic memory, artificial intelligence, and expert-novice research, we know that knowledge has structure. The crucial issue, then, is how to elicit and represent *structural knowledge*, as it is called in cognitive psychology (see Jonassen, Beissner, & Yacci 1993, esp. chapter 7). Structure is the organizational property of knowledge; similarity among concepts defines that structure. These conceptual relationships become more structured with expertise, and a community consensus can be discovered by an idealized representation. These ideas imply that a novice's knowledge structure will be simple and contain misconceptions (Goldsmith, Johnson, & Acton 1991), and that as novices become more knowledgeable during a course, their representations will become more structured overall and more similar to the expert consensus.

We did the following assessment of the concepts' relatedness, following well-known procedures in cognitive psychology (Acton, Johnson, & Goldsmith 1994; Johnson, Goldsmith, & Teague 1995). We divided the concepts in Appendix E into two equal groups ($N = 60$) from an alphabetical list. As the second step in the process, we asked the expert panel to assess the relatedness of pairs of terms within each of the two groups. A computer program presented pairs of terms on the screen and asked the participants to judge their relatedness on a scale from 1 (*unrelated*) to 7 (*highly related*). We define the *scope* of any term as the number of times it was associated with any other term, summing the expert panel results for all terms. The higher the scope score, the greater the number of expert associations in this task. This quantifies the *generality* of a concept.

As far as we know, this result is the first study of its kind that sorts and prunes the data through three sieves: importance, scope, and generality. The end result of these filters are the key concepts, then, that can form of the core of an Astro 101 course, based on the responses of a panel of expert teachers. See Table 5 for the top 10 concepts based on scope scores alone. Appendix F gives the complete results.

Table 5. Top 10 Concepts Based on Expert Panel Relatedness Ratings Sorted by Scope Score (number of times paired with any other concept).

4. DISCUSSION

The results in Appendixes D, E, and F should provide reliable and valid guides to core concepts for Astro 101. In fact, the expert ratings and relatedness pretty much defined the content in the Conceptual Astronomy Project at UNM, especially the concepts with scope scores greater than 200. We intentionally included these 32 terms in instructor-constructed concept maps (see Zeilik 2002 for examples of these maps in a textbook).

We then took the work one additional step. To quantify the relatedness of pairs of concepts, each pair was given an association score that was the number of times across all experts that those two concepts were associated as a pair. We define the *concept relatedness* as the pair scores from the association task. The concept relatedness results are proximity data in that they show the closeness of pair connections. This kind of association task is a well-developed tool in cognitive psychology (Acton et al. 1994).

We then transformed the experts' similarity ratings into a conceptual map, using a tool developed in cognitive psychology. The algorithm, called *Pathfinder*, constructs a network map (Figure 1) in which related concepts are in near vicinity, and concepts with high scope scores have the greatest number of links (see [http://interlinkinc.net/index.html\)](http://interlinkinc.net/index.html). From this analysis, we found that 60 concepts with highest scope scores cluster around four nodes: electromagnetic spectrum/photons, stars, mass, and cosmology. We interpreted this result as a community consensus on the overall concepts, and used these results as global guides in course development at UNM.

Figure 1. Pathfinder network map constructed from responses by the expert panel ($N = 18$) to the association task that resulted in pair relatedness ratings. See Appendix D for the abbreviations of the concepts.

Note that the result is *not* a concept map! A concept map ideally has a hierarchy of concepts, usually from top to bottom, and their links have a direction and are named by verbs (see Novak & Gowin 1984).

From my own (MZ) experience, I sense that about 100 concepts sufficiently define a core for Astro 101. I have accomplished the pruning by dropping the "planets as places"—yes, the Solar System!—from my course and placing more emphasis on stars, stellar evolution, galaxies, and cosmology. The planets only serve as "test masses" for Newton's and Kepler's laws. I also do not cover astronomical technology as such, but do examine results (many of the cooperative learning team activities are based on real data). I do not include tools and telescopes, the Solar System, and stellar magnitudes (compare with Appendix A); I use fluxes instead. I am moving away from stellar spectral classification as such, and focusing on physical properties such as color and luminosity. Of course, students still complain on their UNM course evaluations that the course covers "too much"!

In this work, we used the expert panel to define the core concepts and obtain a concept relatedness score among pairs of these concepts. These data are proximity judgments, and a number of valid analyses can be applied to these results. We chose to use the Pathfinder algorithm (Schvaneveldt 1990), which transforms the proximity matrix into a network in which each concept is represented by a node, and the proximities are represented by how closely the concepts are linked. We used the Pathfinder analysis to discover that the structural network has four main nodes.

Acknowledgments

We thank Timothy E. Goldsmith for fruitful discussions. Kathleen (Kim) Teague and Nancy Mattern did much of the original effort with the expert panel and the fall 1994 Astronomy 101 class. This work was supported in part by National Science Foundation Grant DUE-9253983. MZ again thanks his colleagues on the expert panel.

Appendix B—From the AAS "GOALS FOR 'ASTRO 101'" Content

A cosmic perspective—A broad understanding of the nature, scope, and evolution of the Universe, and where the Earth and Solar System fit in

An understanding of a limited number of crucial astronomical quantities, together with some knowledge of appropriate physical laws

The notion that physical laws and processes are universal

The notion that the world is knowable, and that we are coming to know it through observations, experiments, and theory (the nature of progress in science)

Exposure to the types, roles, and degrees of uncertainty in science

An understanding of the evolution of physical systems

Some knowledge of related subjects (e.g., gravity and spectra from physics) and a set of useful "tools" from related subjects such as mathematics

An acquaintance with the history of astronomy and the evolution of scientific ideas (science as a cultural process)

Familiarity with the night sky and how its appearance changes with time and position on Earth

Appendix C—Essential Concepts for Introductory Astronomy

Please mark on this paper your judgments of the degree to which a student's understanding of each of the following concepts is essential by the end of a one-semester introductory course in astronomy for non-science majors.

APPENDIX D—AAPT Workshop Results

We performed the data analysis with SPSS version 6.1 for the Macintosh and SPSS version 10.2 for Windows. This table gives the variable name as coded in SPSS, the full concept name, the mean of the ratings, and the standard deviation of the mean. The number of ratings per item ranged from 37 to 44 (not every participant rated each item).

APPENDIX E—200 Top Concepts Rated by Expert Panel (*N* **= 18)**

This list is ranked by mean, with 1 = *highly essential*, 5 = *not at all essential*.

APPENDIX F—Top 120 Concepts Rated by Expert Panel (*N* **= 18)**

These are sorted by scope score (highest to lowest).

References

Acton, W. H., Johnson, P. J., & Goldsmith, T. E. 1994, Structural Knowledge Assessment: Comparison of Referent Structures, *Journal of Educational Psychology*, 86, 303.

Adams, J. P., & Slater. T. F. 2000, Astronomy in the National Science Education Standards, *Journal of Geoscience Education*, 48, 39.

American Association of the Advancement of Science and the National Science Teachers Association. 2001, *Atlas of Science Literacy*, Washington, D.C.: American Association for the Advancement of Science.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. 1981, Catagorization and Representation of Physics Problems by Experts and Novices, *Cognitive Science*, 5, 121.

Fraknoi, A. 2001, Enrollments in Astronomy 101 Courses: An Update, *Astronomy Education Review* , 1(1), 121. http://aer.noao.edu/AERArticle.php?issue=1§ion=4&article=2.

Goldsmith, T. E., Johnson, P. J., & Acton, W. H. 1991, Assessing Structural Knowledge, *Journal of Educational Psychology*, 83, 88.

Hirsch, E. D., Jr., Kett, J. F., & Trefil, J. 1993, *The Dictionary of Cultural Literacy*, Boston: Houghton Mifflin..

Hirschfeld, L. A., & Gelman, S. A. 1994, *Mapping the Mind: Domain Specificity in Cognition and Culture*, Cambridge: Cambridge University Press.

Johnson, P. J., Goldsmith, T. E., & Teague, K. W. 1995, Similarity, Structure, and Knowledge: A Representational Approach to Assessment, In *Cognitively Diagnostic Assessment*, P. D. Nichols, S. F. Chapman, & R. L. Brennan (Editors). Mahwah, NJ: Erlbaum, 221.

Jonassen, D. H., Beissner, K., & Yacci, M. 1993, *Structural Knowledge: Techniques for Representing, Conveying, and Acquiring Structural Knowledge*, Mahwah, NJ: Erlbaum.

Lippert, N., & Partridge, B. 2004, To Hear Ourselves as Others Hear Us, *Astronomy Education Review*, 1(3). http://aer.noao.edu/AERArticle.php?issue=5§ion=2&article=4.

National Research Council. 1996, *National Science Education Standards*, Washington, D.C.: National Academy Press.

Partridge, R. B., & Greenstein, G. S. 2003, Goals for "Astro 101": Report on Workshops for Department Leaders, *Astronomy Education Review*, 2(2), http://aer.noao.edu/AERArticle.php?issue=4§ion=2&article=3.

Rutherford, F. J., & Ahlgren, A. 1989, *Science for All Americans*, New York: Oxford University Press.

ÆR 61 - 108