This document summarizes the work Colin Wallace, along with his advisors (Doug Duncan and Ed Prather), have done on the project "Understanding Students' Difficulties with Cosmology" funded under a 2009-2010 iSTEM Award.

To explore which topics in cosmology are difficult for students, we developed and simultaneously administered four different surveys in introductory astronomy (hereafter ASTR 101) classes at both CU and Arizona. These surveys focus on four key constructs, where "construct" is defined as "the concept or characteristic that a test is designed to measure" (AERA/APA/NCME 1999; see also Wilson 2005). The four constructs are:

1) *Models of the Expanding Universe*: This construct focuses on students' conceptualizations of the expanding universe and the Big Bang.
2) *Hubble Plots*: This construct focuses on whether or not students can use Hubble plots to reason about the age and expansion rate of the universe.
3) *Evolving Universe*: This construct looks at whether or not students can explain how the universe has changed over time as a result of its expansion.
4) *Dark Matter*: This construct probes whether a student can reconstruct the causal chain of reasoning linking a galaxy's observed rotation curve with the distribution of matter within the galaxy.

Each of the four surveys focuses on one of these constructs. We developed these surveys in a systematic manner by following Wilson's (2005) four building blocks for survey development. First, we created construct maps for each construct, which are hypothetical tabular representations of how students vary along each construct. Second, we wrote and selected items for each survey that help us distinguish between students at different levels of the construct. Third, we thought about how each item should be scored and created scoring rubrics for each survey. Fourth, we applied standard psychometric measurement models, such as classical test theory (CTT) and item response theory (IRT), to our scored responses (see Hambleton and Jones 1993 for a comparison of CTT and IRT). This sequence helps ensure that each survey's construct definition, items, scoring procedure, and score interpretation are tightly-linked. This sequence also helps us construct validity arguments for each survey, since the validity of a survey is intimately related to how one interprets it (AERA/APA/NCME 1999). Wallace learned about these methods of survey development, CTT, and IRT as part of a two-semester graduate-level course sequence he took from the School of Education during the 2009-2010 academic year. Wallace was able to devote the time necessary to take these classes in part due to the support he received from his 2009-2010 iSTEM fellowship.

These written surveys are augmented by classroom observations and cognitive interviews Wallace conducted with individual students. While observing three different ASTR 101 classes at CU during the fall 2009 semester, Wallace developed a formal procedure for the classroom observations he's currently conducting of three ASTR 101 courses at CU. In this procedure, Wallace writes down each topic the instructor covers, how long the instructor spends on that topic, how much detail the instructor provides, and how the instructor teaches the topic. Wallace pays special attention to the number and nature of interactive teaching strategies employed by each instructor, since these techniques have been linked to high student learning gains (Hake 1998; Prather *et al*. 2009). He developed a rubric for codifying his observations of interactive teaching, based on his prior observations and the research of others (e.g. Turpen and Finkelstein 2009). These classroom observations help Wallace understand what students have ostensibly learned in class and it helps him make sense of any patterns he might observe in students' survey responses.

To date, Wallace has conducted cognitive interviews with ten ASTR 101 students. During these hour-long think-aloud interviews, Wallace has the interviewee answer every question to three of the four surveys. As part of his survey methodology coursework, Wallace constructed an interview protocol based on the best-practices recommendations of Patton (1980) and Wills (2005). These cognitive interviews serve two purposes: 1) They provide additional evidence about what students find difficult about cosmology; and 2) They help Wallace detect survey questions that are unclear or repeatedly misinterpreted by ASTR 101 students. Wallace in currently in the process of transcribing these interviews for further analysis.

Wallace has also coauthored a new suite of lecture-tutorials devoted to cosmology. Each lecture-tutorial was designed based on the same constructivist framework underlying the development of the highly successful *Lecture-Tutorials for Introductory Astronomy*. (Prather *et al* 2005; Prather *et al* 2008). A lecture-tutorial is a two to six page worksheet of Socratic dialogue-driven questions that helps students construct their own understandings of a topic that research has shown students struggle with (Prather *et al* 2005). Below is a brief description of the content and goals of each of the five new cosmology lecture-tutorials.

1) *Making Sense of the Universe and Expansion*: This tutorial helps students make sense of the balloon analogy for the expanding universe.
2) *Hubble's Law*: This tutorial helps students understand Hubble's law and Hubble plots for the expansion of the universe.
3) *Expansion, Lookback Times, and Distances*: This tutorial helps students understand the effect the expansion of the universe has on lookback times.
4) *The Big Bang*: This tutorial helps students realize that the Big Bang refers to the expansion of space from an initially hot and dense state and not to the explosion of pre-existing matter into empty space.
5) *Dark Matter*: This tutorial helps students understand one piece of evidence for the existence of dark matter in spiral galaxies.

These five lecture-tutorials were developed following a procedure similar to Prather *et al*'s (2005) description of the development of the original *Lecture-Tutorials*. We selected the topics for the five cosmology lecture-tutorials based on students' responses to a preliminary survey administered in the fall of 2008, our knowledge of common ASTR 101 cosmology topics, and our experiences teaching ASTR 101 students. Wallace used the first drafts of these lecture-tutorials in a summer ASTR 101 course he taught from July-August 2009.

These cosmology lecture-tutorials have since been significantly modified. Some modifications were made in response to feedback we received from our colleagues in order to refine the clarity, content, and scope of the tutorials. Other modifications were made in response to our observations of students working on the tutorials. Like the original *Lecture-Tutorials*, we want the cosmology lecture-tutorials to function in large enrollment (approximately 100+ students) classes. We assume if they can work in this environment, they can work in smaller classes. In order to function in large classes, the cognitive steps between each question must be small and students must be able to progress through the tutorials with minimal instructor assistance (Prather *et al* 2005). Based on these criteria, we flagged a question for revision if it confused many students, led many students to the wrong answer (either on the tutorial itself or on one of our assessment questions), or required the instructor to help many groups of students. We have also revised the tutorials in conjunction with revisions to ensure that they are linked to the same constructs. Finally, we have videotaped actual ASTR 101 students working on these tutorials as a part of class; we are in the process of transcribing and analyzing these videotapes and we expect they will help us make further revisions to help the tutorials better align with the needs of ASTR 101 students. We expect this iterative procedure will yield cosmology lecture-tutorials that are suitable for publication in future editions of the *Lecture-Tutorials for Introductory Astronomy* (Prather *et al* 2008).

To date, each of the five cosmology lecture-tutorials is in its third draft version or beyond. The four assessment surveys are currently in their eighth version and we have evidence supporting the validity of all but three questions.

In addition to Wallace's summer course, these tutorials have been used in large lecture hall ASTR 101 classes offered at both CU and Arizona, including an 800 student class currently being taught by Ed Prather at Arizona. By the end of the spring 2010 semester, approximately 1450 students will have used these tutorials.

Do these lecture-tutorials have any effect on students? A preliminary analysis indicates the answer is "yes." For example, 65% of students who used the lecture-tutorials were able, by the end of the fall 2009 semester, to correctly identify the rotation curve of a spiral galaxy. Rotation curves were emphasized in all classes (both using and not using lecture-tutorials). Yet, of all the students not using the lecture-tutorials, only 19% could select the correct rotation curve by the end of the semester. While results such as these are suggestive about the efficacy of the tutorials, additional work must be done before we can give a definitive answer to this question.

These lecture-tutorials and our research into students' difficulties with cosmology appear to be meeting a need in the broader community. In addition to the poster he presented at the inaugural STEM Education Symposium at CU-Boulder, Wallace also presented an updated poster and gave a contributed talk at the January 2010 meeting of the American Astronomical Society. Several attendees requested copies of the tutorials and/or offered to help the project by using the surveys and tutorials in their classes. In April 2010, Wallace will took part in a roundtable discussion on the use of IRT in his research at the International Objective Measurement Workshop. In August 2010, Wallace and Prather will run a workshop for ASTR 101 instructors at the triannual Cosmos in the Classroom Symposium; the workshop will focus on how to implement the cosmology lecture-tutorials. Finally, Wallace will presented his research as one of the invited speakers to the January 2011 meeting of the American Association of Physics Teachers.

This work would not have been possible without the iSTEM award. Receiving the award freed up time Wallace would have otherwise spent as a TA and allowed him to take the classes and conduct the research and design work discussed above.

References

American Education Research Association, American Psychological Association, and the National Council on Measurement in Education. 1999, *Standards for educational and psychological testing*, Washington, DC: American Educational Research Association.

Hake, R. R. 1998, "Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses," *Am. J. Phys.*, **66**(1), 64-74.

Hambleton, R. K., & Jones, R. J. 1993, "Comparison of Classical Test Theory and Item Response Theory and Their Applications to Test Development," *Educational Measurement: Issues and Practice*, **12**(3), 253-62.

Patton, M. Q. 1980, *Qualitative Evaluation Methods*, Thousand Oaks, CA: SAGE Publications.

Prather, E. E., Rudolph, A. L., Brissenden, G., & Schlingman, W. M. 2009, "A national study assessing the teaching and learning of introductory astronomy. Part I. The effect of interactive instruction," *Am. J. Phys.*, **77**(4), 320-330.

Prather, E. E., Slater, T. F., Adams, J. P., Bailey, J. M., Jones, L. V., and Dostal, J. A. 2005, "Research on a Lecture-Tutorial Approach to Teaching Introductory Astronomy for Non-Science Majors," *Astron. Educ. Rev.*, **3**(2), 122-36.

Prather, E. E., Slater, T. F., Adams, J. P., Brissenden, G., Dostal, J. A., and the CAPER team. 2008, *Lecture-Tutorials for Introductory Astronomy*, San Fransisco, CA: Pearson Addison-Wesley.

Turpen, C., and Finkelstein, N. 2009, "Not all interactive engagement is the same: Variations in physics professors' implementation of *Peer Instruction*," *Phys. Rev. ST: Phys. Educ. Res.*, **5**(2), 020101.

Wills, G. D. 2005, *Cognitive Interviewing: A Tool for Improving Questionnaire Design*, Thousand Oaks, CA: SAGE Publications.

Wilson, M. 2005, *Constructing Measures: An Item Response Modeling Approach*, Mahwah, NJ: Lawrence Erlbaum Associates.