Development of an observation protocol to measure implementation of Scientific Teaching
Brian Couch, MCD Biology

**Background:** Growing demands for skilled scientists and general science literacy have prompted scientists, educators, and policy leaders to issue several national calls for improving undergraduate science education [1]. In 2003, the National Research Council’s report, *BIO2010*, recommended several changes to undergraduate biology education in light of the increasingly interdisciplinary and quantitative nature of biomedical research [2]. This report inspired the founding of the National Academies Summer Institute for Undergraduate Education in Biology (SI) by Jo Handelsman (Yale) and Bill Wood (CU-Boulder) under the philosophy that science education ought to reflect the nature of scientific inquiry, incorporate our understanding of how people learn, and ensure equal learning opportunities for students of diverse backgrounds [3].

With on-going support from the Howard Hughes Medical Institute, the SI continues to bring together faculty members from diverse institutions to develop their teaching skills and to discuss recent innovations in undergraduate biology education. In 2007, the SI curriculum was formalized with the publication of the book *Scientific Teaching* [4]. Written in plain language, this book synthesizes a large body of education reform literature and presents the reader with a coherent framework for implementing research-based teaching practices. *Scientific Teaching* has gained widespread influence in the education community, serving as the basis for professional development workshops and as a resource for individuals interested in improving their teaching.

**Problem Description:** There is growing evidence that faculty adoption of Scientific Teaching can significantly improve student learning [5]. With this approach increasingly being used to guide reform efforts, a tool is needed to objectively measure the degree to which Scientific Teaching has been implemented in the classroom. Such a tool will be essential for capturing behaviors associated with this model and will ultimately lead to better understanding of how faculty incorporation of Scientific Teaching affects student learning.

The goal of this proposal is to develop a valid and reliable observation protocol to objectively measure the extent to which a faculty member implements Scientific Teaching.

Several classroom observation protocols have been developed to date. The Teaching Dimensions Observation Protocol (TDOP) provides a rich portrait of pedagogical strategies employed by an instructor [6]. The Reformed Teaching Observation Protocol (RTOP) measures the extent to which a class has been reformed with inquiry-based instructional methods and activities [7]. While each of these instruments has the ability to detect instructional changes, they do not completely align with the nuances of the Scientific Teaching model. For example, one component of this model is the use of backward design to develop course learning goals and objectives [8]. Since the TDOP and RTOP are limited to classroom observation, they are not well-positioned for evaluating this particular aspect of Scientific Teaching. Furthermore, many of the items on these current protocols are difficult to observe in large introductory classes, a central focus of most science teaching reform efforts.
Methodology and Evaluation: The development of an observation protocol to measure Scientific Teaching will proceed through the phases outlined below.

Phase 1: Defining the essential elements of Scientific Teaching. We have generated an exhaustive list of the observable and implicit components contained in Scientific Teaching. We will refine this list to consolidate redundant areas and more clearly define the core elements. We will consult the literature and Scientific Teaching content experts to provide greater substance for any underdeveloped areas.

Phase 2: Identifying observable behaviors associated with each element. The principles of Scientific Teaching can be applied in the classroom in diverse ways. Once we have defined the essential elements, we will use classroom observations, literature review, and faculty/student interviews to further elaborate a spectrum of observable behaviors associated with each element. This approach will allow us to refine the protocol to highlight elements that can be identified through observable behaviors in the context of a course.

Phase 3: Developing and refining the observation protocol. Our protocol will consist of 25-30 items with five-point Likert scales to rate degrees of implementation. These items will apply to different course features, such as syllabus content, in-class instruction, out-of-class activity, and assessment. For example: To what extent did the faculty member express what students should know and be able to do by the end of the course? 1 – Syllabus lists a number of topical areas to 5 – Syllabus contains explicit learning objectives that define both the intended knowledge scope and the expected performance level. Content validity of the items will be determined by asking faculty with experience in Scientific Teaching to rate the appropriateness of each item on the instrument. We will test our initial items on sample course materials and diverse classroom environments, including large/small classes, traditional/reformed classes, and introductory/upper division courses. We will use multiple raters for the same materials and classes, and the protocol will be refined until we achieve high inter-rater reliability. We will also empirically determine the number of observations required to achieve stable measurements for each item.

Phase 4: Measuring instrument reliability and validity. Once we have confidence in the reliability of our items, we will measure the reliability and validity of the instrument as a whole by using it to measure elements of Scientific Teaching in several courses. Each course will be rated by two independent observers, and the reliability of the instrument will be calculated by linear regression of the scores from the different observers. Construct validity of the instrument will be ascertained using regression analysis to calculate how well each category predicts the overall score.

The Development Team: This instrument will be developed by a team of researchers in MCDB, led by myself (Brian Couch) and supervised by Jennifer Knight and Bill Wood. We will be supported by our collaborators—and my past mentors—at the Yale Center for Scientific Teaching, including Jo Handelsman (PI) and Mark Graham (Evaluation Director). The MCDB research team includes two undergraduate students, Aubrey Pierce and Tyler Schelpat, who will be essential to instrument development and testing.
**Project Impact:** Scientific Teaching has become a centerpiece of biology education reform, with the SI expanding from one institute into seven regional institutes held at campuses nationwide, including CU-Boulder. The proposed instrument will help guide education reform efforts by providing an objective characterization of classroom behaviors associated with Scientific Teaching. We envision that this tool will be utilized by investigators across science disciplines for broad research applications. This tool can also be employed on a pre-post basis to gauge how instructors’ teaching practices change as a result of participation in professional development workshops, such as the SI [9]. Thus, this project will further CU-Boulder’s reputation as a national leader in education with a particular strength in developing tools of broad utility to the entire education community.

On the departmental level, we expect that the faculty interviews and surrounding discourse will serve to promote a culture that values teaching excellence. In the longer term, since this instrument will explicitly define research-based instructional practices, we foresee using it as the basis for faculty development workshops as well as distributing it as a tool to help individual faculty improve their own teaching practices.

**Timeframe:** We anticipate that a preliminary draft from Phase 1 will be complete by early summer. Phases 2 and 3 will commence over the summer, with summer courses serving as a testing ground for initial protocol items. Phase 3 will continue into the fall semester, and Phase 4 will begin the following spring semester, with parts of this phase, such as statistical analyses, extending beyond the proposed budget period.
References: