**FRIDAY**

**Long talk- Smith 100**

**Weeding out diversity in STEM? Lessons learned from 15 years of retrospective data and a meta-analysis of student achievement in active learning classrooms [abstract # 109]**

Elli Theobald\*, University of Washington; Scott Freeman, University of Washington; Rebecca Harris, Arizona State University; Michael Mack, University of Washington; Jasmine Bryant, University of Washington; Mariah Hill, University of Washington; Elisa Tran, University of Washington

Despite widespread effort to increase access to and diversity in STEM, women and minority students remain underrepresented in both STEM majors and STEM professions. Achievement gaps in college (i.e., differential performance between represented and under-represented students) likely contribute to this problem because lower performing students are more likely to drop out and less likely to major in STEM fields. Using four sources of evidence, we tested the hypothesis that active learning can close achievement gaps for historically under-represented students: First, we used retrospective data to quantify achievement gaps in a predominantly lecture-based STEM curriculum. Second, we updated the Freeman et. al 2014 meta-analysis, and identified types and contexts of active learning that are most effective. Third, we attempted to meta-analyze the instances of active learning closing achievement gaps for underrepresented students. Finally, we proposed an additional mechanism through which achievement gaps have been successfully narrowed or eliminated. Using 15 years of retrospective registrar data (2000–2016) we quantified achievement gaps across the predominantly lecture-based chemistry curriculum at a large, public, R1 university. With more than 130,000 student observations, from nearly 47,000 unique students, we demonstrated that these large and persistent gaps are not simply a continuation of the achievement gap on the SAT: rather, gaps get worse in college chemistry for all underrepresented students (URM, first-generation, women, and low-SES students). Notably, achievement gaps decrease from the first quarter of Intro Chemistry to the third quarter of Organic Chemistry, however these reductions are coincident with a reduction in the percentage of URM students in these courses. Taken in conjunction with the disproportionate rates of failing, it is apparent that diversity is being weeded out of this traditionally-taught chemistry sequence. Achievement gaps have consequences: they are correlated with attrition and with lower rates of majoring in STEM. Fortunately, research shows that evidence-based teaching practices improve student achievement compared to traditional lecturing. In a meta-analysis of more than 130 publications, we updated the Freeman et al. 2014 meta-analysis, using studies that have been published since their search concluded. We focused particularly on identifying characteristics of superior active learning. We found that active learning is effective across most contexts (e.g., class size, course level), but when the percentage of class time devoted to active learning is low (<30%), learning gains are no higher than in traditional lecturing. We further investigated the published and unpublished literature to ask if active learning can close achievement gaps for female and URM students. Two trends are apparent: 1) Relatively few studies investigate this critically important question: after contacting every corresponding author from the Freeman et al. 2014 meta-analysis and the current meta-analysis (totaling 297 studies), fewer than 20 authors can contribute data. 2) Within these datasets, the sample size of under-represented students is variable, ranging from 79 URM students to only three. Nevertheless, the variation in instances of closing the achievement gap cannot be explained by active learning alone. Finally, using the “Wise Schooling” framework developed by Claude Steele in 1997, we hypothesize that a combination of deliberate practice (vis-à-vis active learning) and instructor soft skills are necessary to narrow achievement gaps in university STEM classes. In all, our work suggests that historical achievement gaps in STEM are not destiny: they can be reduced or eliminated with evidence-based instruction.

**FRIDAY- afternoon**

**Involving URMs in Biology- Friday**

**Explaining the Dearth of African American Students in Evolutionary Biology as a Function of Religiosity**  Elizabeth Barnes\*, Arizona State University; Gale Sinatra, University of Southern California; Taija Hendrix, Arizona State University; Hayley Dunlop, Arizona State University; Sara Brownell, Arizona State University [abstract # 68]

In response to a nationwide call to increase the diversity of graduates with degrees in the fields of science, technology, engineering, and mathematics (STEM), there has been a major research thrust in science education aimed at increasing diversity in STEM and participation of underrepresented groups within STEM. However, rates of diversity in STEM can be field specific. Among biology disciplines, evolutionary biology is particularly lacking in racial/ethnic diversity. In 2011, a survey of Doctorate Recipients conducted by the National Science Foundation revealed that degrees awarded to African Americans in evolutionary biology are notably absent. While 5% of doctorates in 2011 were awarded to African Americans in neuroscience, molecular biology, and biochemistry, the field of evolutionary biology failed to award any doctoral degrees to African American students. Past research illustrates that religious beliefs tend to be an important part of African American culture and that commitment to religious beliefs is higher among African American populations compared to other racial/ethnic groups. Research also shows that higher religiosity is predictive of rejecting evolutionary biology. We hypothesized that African American college students would be more likely to reject evolution than their non-underrepresented peers and that this relationship would be explained by higher levels of religiosity among African American students. To test our hypotheses, we collected two separate sets of data in the spring of 2016 and fall of 2017 from over 2,500 students taking biology classes at a research-intensive university in the southwest. Using closed-ended previously published surveys administered to a total of 18 biology classes, we measured students’ acceptance of evolution and we also collected their religiosity and race/ethnicity. We used independent samples t-test and hierarchical linear regression to test the relationships between student race/ethnicity, student religiosity, and their acceptance of evolution. Our hypothesis was confirmed. We found that controlling for prior academic ability, African American students scored significantly lower on acceptance of evolution compared to their non-underrepresented peers ((Spring 2016) t=3.89, p<.001; (Fall 2017) t=3.69, p<.001). Additionally, African American students scored higher on religiosity ((Spring 2016) t=-4.50, p<.001; (Fall 2017) t=-4.20, p<.001) than their non-underrepresented peers. Further, our hierarchical regression indicated that low acceptance of evolution among African American students was explained entirely by their higher rates of religiosity; once religiosity is controlled for in the regression, differences between African American students and non-underrepresented students disappear ((Spring 2016) B=-.020, p=.56; (Fall 2017) B=-.07, p=.06). Given that past research illustrates that religious individuals may not feel comfortable in the evolution education community, this result implies a need for religious cultural competence in evolution education to foster ethnic and religious diversity in evolutionary biology.

**Disparities in URM evolution acceptance: implications for diversifying the biological sciences** Gena Sbeglia\*, Ross Nehm, Stony Brook University [abstract # 124]

STEM fields are characterized by low diversity even though underrepresented minorities (URMs) enroll in STEM majors at the same rate as White students. In particular, the field of Ecology and Evolution (E&E) has one of the lowest proportions of URMs among all STEM fields. As gateway courses to the life sciences, introductory biology classes have the potential to engage a diverse audience in terms of race, gender, first-generation status, and academic preparedness. Therefore, introductory courses are valuable settings in which to study the factors contributing to the disproportionate attrition of URM students. One hypothesis that has received considerable attention is that acceptance of evolution plays an important role in learning about evolution. Therefore, it is possible that a low acceptance of the central tenets of the discipline may explain the relatively poor persistence of URM students and ultimately, their disproportionately high rates of attrition. In this study, we aim to assess the relationship between URM status and evolutionary acceptance. Evolutionary acceptance was measured using the published Inventory of Student Evolution Acceptance (I-SEA), which assesses acceptance of evolution on three subscales: microevolution, macroevolution, and human evolution. Pre- and post-course acceptance measures, along with student-specific data (race, gender, age, previous biology courses taken, reading skills, english as a first language) were collected pre-post-course for six semesters of an introductory biology course (N = 2132, average consent rate=74%). Prior work using Rasch found the I-SEA to be multi-dimensional across subscales (as theorized) and to have acceptable fit statistics. Rasch-transformed subscale scores (i.e. person scores for each subscale) were used as outcome variables in regressions. We compared three regression models for each subscale, each of which controlled for student variables and had varying interaction effects among race, gender, and time period (i.e. pre/post). We addressed repeated measures by anchoring the pre-test Rasch item scores to the post-test, and treated student ID as a random variable. Pre-course acceptance scores were generally high but increased significantly post-course for all subscales (Micro: β = 0.546, df = 4178, p << 0.001, Cohen’s D = -0.31; Macro: β = 0.570, df= 4178, p << 0.001, Cohen’s D = -0.37; Human: β = 0.599, df = 4178, << 0.001, Cohen’s D = -0.29); these gains did not differ by race or gender (interaction effects n.s.). However, females had lower acceptance scores than males for all three subscales (Micro: β = -0.372 df=4178, p < 0.001; Macro: β = -0.286, df=4178, p << 0.001; Human: β = -0.570, df=4178, p << 0.001). Furthermore, Asian and URM students had lower acceptance scores than White students for all subscales (Asian vs. White: Micro β = 0.236, df = 4178, p < 0.05; Macro β = 0.319, df= 4178, p << 0.001; Human β = 0.319, df = 4178, p < 0.05; URM vs. White: Micro β = 0.476, df = 4178, p < 0.001; Macro β = 0.411, df= 4178, p << 0.001; Human β = 0.580, df = 4178, p < 0.001). We report a measurable disparity in evolution acceptance by gender and race in our large population of undergraduate biology students. Because evolution is a central tenet of E&E, lower acceptance of evolution by females and URMs may contribute to URM attrition. Our work, the first to use validated instruments and Rasch-transformed scores to measure acceptance pre-post course in large samples, calls for two actions: First, application of interventions targeting acceptance in undergraduate biology courses; second, empirical studies examining the degree to which acceptance disparities impact URM attrition rates. Funding was supported by the National Science Foundation Division of Undergraduate Education.

**Identifying the Community Cultural Wealth of Academically Successful Black Science Students** Darris Means, Omowunmi Oni, Birook Mekonnen,; Julie Dangremond Stanton\*, UGA [abstract # 140]

Although attrition of undergraduate students from science, technology, engineering, and mathematics (STEM) is an issue for all racial and ethnic groups, the rates of attrition are highest for Black\* students. Only about 35% of the Black undergraduate students who begin college as STEM majors earn a STEM degree. Researchers have suggested that poor academic preparation, lack of qualified teachers, and lack of rigorous middle and high school classes are major factors that negatively impact Black students' persistence in STEM. Additionally, Black students have to navigate feelings of isolation, racial stereotypes, and subtle and overt racism, all of which can hinder their persistence in undergraduate STEM degree programs. Preventing the loss of Black undergraduate students from STEM is critical not only for building the size of the STEM talent pool, but also for ensuring its diversity. While the reasons Black students leave STEM have been investigated, the strengths and strategies that Black students use to succeed in science are less clear. In order to help Black students persist in earning undergraduate STEM degrees, we must better understand their mechanisms of success. One way to accomplish this is by using the anti-deficit achievement framework to explore the experiences of high-achieving underrepresented minority (URM) students in STEM (Harper, 2012). This approach invites researchers to reverse deficit-oriented questions such as "Why do so many Black students leave STEM majors?" and reframe them as achievement-oriented questions such as "Why do Black students persist in STEM majors despite all the known barriers?" Using an anti-deficit achievement approach, we studied the strengths of academically successful Black undergraduate science majors. In particular, we investigated the community cultural wealth that Black undergraduates bring to their science majors. Community cultural wealth consists of six forms of capital or "knowledge, skills, abilities, and contacts" that students of color possess and can use for educational success (Yosso, 2005, p.77). Our research question is: What forms of community cultural wealth and previously undescribed capital do academically successful Black undergraduate students use to persist in science majors? To answer this question, we studied Black science majors in the final year of their bachelor's degree program at a doctoral university with highest research activity (n=20). We collected data using a demographic and pre-college survey and two semi-structured interviews. Using content analysis, we found evidence of participants' use of all six forms of community cultural wealth. For example, participants in our study used aspirational capital (the ability to maintain career aspirations while encountering setbacks) and navigational capital (the ability to find and use resources to navigate through higher education) in their science majors. We identified novel forms of capital, such as spiritual capital (the ability to draw upon spiritual beliefs for support), which contributed to participants' success. We also gained insights into the ways subtle and overt racism impacted participants' motivation to succeed. For example, some participants strove to positively represent Black science students in order to disrupt negative racial stereotypes. In the future, we will translate our results into products that promote awareness of Black science students' community cultural wealth. \*In this abstract, we use Black to describe individuals from the African diaspora, which also includes individuals who identify as African American.

**Determining Aspects of Active Learning and Classroom Community that are Important to Underrepresented Groups in a Large-Lecture Biology Course: The Student Perspective** Michelle Nugent\*, NC State University; Miriam Ferzli, NC State University; Miles Engell, NC State University [abstract # 185]

Issues of persistence and representation of diverse groups in science, technology, engineering, and mathematics (STEM) majors are well known. Research suggests persistence and student learning are associated with motivation and sense of community, among other factors. Large introductory science courses are often difficult for students as they adjust to college rigor. Underrepresented (URM) students can have an added struggle in that they often lack the role-models and the guidance necessary to learn how to be successful. We define an URM student as one who self-identifies as a first-generation college student, racial minority, need-based financial aid recipient, member of the LGBTQ community, transfer student, or student with disabilities. Much education research has focused on improving large lecture courses for all students. Among many pedagogical strategies, active learning has been identified as an effective approach for improving learning outcomes. However, claims of active learning effectiveness are often based on final course grades or exam scores, neither of which explain how, why, or for whom active learning strategies are beneficial. We explore the components of active learning that students identify as most beneficial in a large introductory biology course, and compare URM and non-URM perspectives. We also attempt to determine the perceived level of classroom community. The study captures the student perspective through the lens of the Social Cognitive Theory and Expectancy-Value Theory, using a convergent parallel mixed methods approach. The specific research questions are: 1) Which aspects of active learning tasks do students identify as most beneficial for learning? 2) Are there differences in URM and non-URM perspectives of active learning task values and the level of classroom community? The Classroom Observation Protocol for Undergraduate STEM (COPUS) was used to examine active learning strategies and student engagement during class. Immediately following in-class activities, students used a personal response platform, Top Hat, to describe aspects they found useful for learning. Students completed a questionnaire to Assess Student Perspective of Engagement in an Active-Learning Classroom (ASPECT) and the Classroom Community Scale (CCS), both Likert-type scales. Finally, students reported demographic data including URM status. Open-coding of Top Hat responses found that URM and non-URM students most frequently reported working with peers as beneficial, followed by practical use of notes, reiteration of material, and development of new study skills, among numerous others. T-tests compared factors related to task value, connectedness, and learning environment across different student groups from ASPECT and CCS surveys, for students who completed all tasks (n=159). ASPECT data yielded statistically significant differences (p<.05, ɑ=.05) in student perception of task value for students with disabilities, who found the activity more valuable and transfer students, who found the activity less valuable, as compared to non-URM students. Analysis of CCS for URM and non-URM students yielded statistically significant positive perceptions of connectedness and learning for students with disabilities, first-generation college students and LGBTQ students, while transfer students had negative perceptions of connectedness and learning compared to non-URM peers. The findings suggest that many URM students hold active learning strategies in higher regard than their non-URM peers. This study contributes to the limited understanding of the experiences and needs of URM groups in large-lecture classrooms by exploring the student perspective. It is necessary to further examine the transfer student population perceptions, as well as the connection between specific active learning strategies and the sense of classroom community.

**Towards a Critical Pedagogical Approach: Challenging the Dominant Enterprise of STEM Instruction through Inverted Course Lectures in Supporting Underrepresented Racial Minority Students** Julio Fregoso\*, Kevin Eagan, Edgar Romo, UCLA [abstract # 261]

Introductory STEM courses across colleges and universities have a reputation of disproportionately weeding out seemingly lower-performing, less-engaged students. Efforts to incorporate critical pedagogical practices to retain STEM students comprise a much smaller proportion of institutions providing such efforts for their student populations. As part of a federally funded grant initiative, researchers at a highly selective public university have transformed how its life science (LS) majors experience the core life science curriculum by implementing an inverted (“flipped) classroom format with goals of enhancing students’ motivation of learning through student-centered pedagogical practices. Such inverted classroom approach is an attempt to confront the nationally recognized STEM retention issue that is in part due to the current (antiquated) pedagogical approaches that remain as the dominant form of teaching enterprise in the nation. A flipped learning model allows faculty to supplement their teaching through an online learning environment, creating a space for instructors to adopt a student-centered pedagogical approach during course lectures. Online support provides high quality, interactive videos that contribute to the knowledge construction of students; it is designed to challenge students to actively engage with course content then provides immediate assessment results on students’ overall comprehension gains. We merged administrative data provided by the University registrar with pre and post concept test results to evaluate how moving from a traditional lecture format and toward active learning strategies affects learning gains and engagement in the LS Core courses, comparing outcomes for students of different race/ethnicity backgrounds, gender, and first-generation status. The LS core curriculum is comprised of four LS courses. For abstract purposes they shall be titled LSA, LSB, LSC, and LSd. Analyses computed were focused on LSB, C, and d science courses. Preliminary analyses show learning gains based on pre- and post-concept test scores by the type of learning environment and whether the student identified as an underrepresented racial minority (URM), which we define as Black, Latino, or Native American. The data utilized for this analysis has a sample size of N=10,463, where students in the URM recoded variable comprised of 21% of the student population. Such number of URM students is not an anomaly for the highly selective public institution location under inquiry. Analyses show that URM-LSB students in flipped learning classrooms make significantly stronger learning gains than their peers in more traditional lecture environments. Additionally, the achievement gap between URM students and their White and Asian counterparts is reduced between the pre- and post-tests for students in flipped learning courses than traditional lecture classes. LSC and d enrolled students demonstrate similar results in cognitive and affective gains per concept test scores. These findings challenge and disrupt the widespread assumption and socialization processes that incur when faculty adopt pedagogical practices that cater to the needs of URM students in flipped course learning environments. Preliminary regression analyses highlight significant predictors for students enrolled in LSd courses. Predictors such as GPA, Pell-grant recipients, and math AP scores help explain variance in our dependent outcome of concept test posttest scores, whereby the Pell-grant recipient variable was a significantly negative predictor in the regression model. Although URM students demonstrate higher cognitive gains in flipped courses than their URM non-flipped enrolled counterparts, the achievement gap present amongst URM and non-URM students based off of concept tests scores calls attention to the much needed pedagogical reform that addresses the issue of diverse learners within STEM.

**Undergraduate Research Experience- Friday**

**Authentic Research Connection: Assessment of Faculty Led Research Projects in the Introductory Biology Laboratory** Julianne Vernon\*, University of Michigan; Deborah Goldberg, University of Michigan; John Wolfe, University of Michigan [abstract # ]

A common cause of loss of students to pursing a major in science is a negative experience in introductory “gateway courses”. The Authentic Research Connection program (ARC) funded by the "X" aims to reduce this loss by engaging students in introductory biology and chemistry laboratories through faculty led research projects. For this talk I will only focus on the introductory Biology laboratory course (Bio 173), which is required of all biology majors. It is taken primarily by second year students and has a prerequisite of either AP Biology credit or previous or concurrent enrollment in both of the sequence of introductory biology lecture courses (Bio171 or Bio172). We recruited faculty researchers whose research areas over lapped with the content/skills sets learned within Bio173 and whose research could benefit from more participants without extensive technical training., in addition to the faculty researcher and graduate student instructors funded as part of the regular course, implementation of each research project included a postdoc from the research lab to oversee data quality and ensure appropriate protocols. We also provided an undergraduate learning assistant for each section of 24 students. This talk focuses on one of the research projects, led by Professor "X" on the human biome. Students sampled their own microbiome and investigated the introduction of different starch/fibers in their diet. I describe the impact of the ARC human microbiome project on student interest over five semesters from Fall 2015 to Fall 2017. A modified compilation of four independent published instruments was used, which we called Research Based Course Assessment (RBCA). The survey focused on five different factors, which include student’s attitude toward, interest in, and self-efficacy in Biology. All survey factors were based on five-point scales of strongly agree to strongly disagree. Survey instrument questions were validated through student focus groups during the first three semesters of the project. Surveys were administered at the beginning (pre) and end (post) of the semester in the ARC sections as well as in a control population (traditional biology lab sections). Student response to each factor compared between the pre and post surveys using paired t-tests. Additionally, we compared the research and control sections using unpaired t-tests. Cronbach’s alphas were computed for all factors in the pre and post surveys to test the reliability of the factors. We found that the students in the ARC sections had a significantly higher maintained interest in Biology and confidence in conducting laboratory related tasks relative to students in the control section.

**“Curing” CUREs: Understanding and alleviating barriers to implementation of course-based undergraduate research experiences** Elizabeth Genne-Bacon\*, Tufts University School of Medicine; Jessica Wilks, Tufts University School of Medicine; Carol Bascom-Slack, Tufts University School of Medicine [abstract # 29]

Undergraduate research experiences promote many positive learning outcomes and increase persistence in STEM fields. However, these experiences traditionally exclude many students, particularly those at schools without significant research infrastructure. Course-based Undergraduate Research Experiences (CUREs) have emerged as an effective way to expose more students to authentic research, but despite the benefits of CUREs, most laboratory instructors still use traditional cookbook methods. Research investigating barriers to CURE implementation reported by veteran CURE instructors reveals instructor time, class time within the semester, and access to equipment and materials as common challenges. However, little is known about the decision process of novice CURE instructors when contemplating adoption of a CURE. Understanding specific challenges perceived or experienced by instructors can facilitate design of CUREs to mitigate these barriers, thus allowing CUREs to reach more students. The Prevalence of Antibiotic Resistance in the Environment (PARE) project is a CURE designed to overcome reported challenges by using a flexible modular approach that is short duration, low-cost, and uses only simple techniques. Drawing dozens of new instructors per year, PARE provides an opportunity to study not only the perceived barriers/challenges for implementing CUREs in novice instructors, but also gauge efficacy of PARE’s approach to mitigating these challenges. To gauge perceived barriers and elements of PARE that draw instructors to implement, we undertook a qualitative study of new PARE instructors. Our study is framed in diffusion of innovations (DOI) theory, which posits that the decision to adopt an innovation is primarily influenced by 5 perceived features of that innovation: relative advantage, compatibility, complexity, observability, and trialability. For our purposes, we define both PARE and CUREs (in general, or specific other CUREs, depending on the instructor’s level of awareness) as innovations. We hypothesized that for instructors who are already convinced of the relative advantages of CUREs but who have previously struggled with the decision to implement, a module-based CURE like PARE could reduce complexity and increase compatibility and trialability, tipping the balance in favor of implementation. We conducted semi-structured interviews with 19 new PARE instructors from diverse institution types, who had not yet implemented PARE in their classes. Thematic analysis (with 2 independent coders) was used to code interview transcripts for DOI-related themes, as well as additional themes emergent from the text. We found that all interviewees made statements indicating their acceptance of the relative advantage of CUREs, particularly their potential to enhance student learning and engagement with science. Many instructors also made statements indicating that CUREs were compatible with their beliefs and values about education. PARE’s perceived compatibility was notably higher than other CUREs, particularly with respect to instructors’ course structure and content, funding and resources, and past teaching experiences. The perceived trialability of PARE was also higher than for other CUREs, while perceived complexity was lower. The most commonly reported barriers to implementing non-PARE CUREs were instructor time/bandwidth, time in the semester or having to transform the whole course, and access to equipment and materials, none of which were common PARE challenges. Our data suggest that, for these instructors, PARE did effectively lower barriers for implementing CUREs relative to other CUREs they had considered. Designing CUREs that specifically address these common barriers, for example, by using only techniques that instructors are likely to have experience with (i.e. compatibility with past experiences), could increase adoption of CUREs, especially at institutions with limited resources.

**Characterization of Student and Instructor Behaviors in CURE and Non-CURE Contexts: Impacts on Student Motivation and Science Identity Formation** David Esparza, University of Texas at El Paso; Dimuthu Fernando, The University of Texas at El Paso; Amy Wagler, The University of Texas at El Paso; Jeffrey Olimpo\*, The University of Texas at El Paso [abstract # 122]

Within the last decade, the advent of course-based undergraduate research experiences (CUREs) has significantly increased student access to authentic scientific opportunities within the STEM disciplines. Current evidence within the bioeducation literature indicates that, in contrast to traditional laboratory curricula, CUREs positively impact scholars’ development of experimental design skills, affect, and ability to “think like a scientist.” While imperative in enhancing the community’s understanding of the benefits of CUREs, these and other studies have historically focused on student and/or faculty outcomes within such learning environments. In contrast, few, if any, studies have characterized the behaviors and interactions occurring within CURE and non-CURE contexts and the impact of those behaviors on non-cognitive student outcomes. In an effort to address these concerns, a quasi-experimental, mixed methods design was employed to evaluate student and instructor behaviors in four CURE and four non-CURE sections of an introductory cell and molecular biology course at our institution in the Spring 2016 semester. Videodata were collected in each section during six instructional episodes (96 hours of total video), distributed evenly throughout the term, so as to generate a representative account of those interactions occurring within each classroom. Data were analyzed using an adapted version of the Laboratory Observation Protocol for Undergraduate STEM (LOPUS; Velasco et al., 2016), a periodic-interval measure which allows one to determine the frequency of common behaviors (e.g., teacher or student questioning; students engaging in a laboratory exercise) within the learning environment. Episodes were evaluated by two individuals with expertise in bioeducation, with high inter-rater reliability observed (κ = 0.94; p < 0.001). In addition, pre-/post-semester survey data were obtained from CURE (n = 47) and non-CURE (n = 60) students in the above-referenced sections, including measures of intrinsic motivation (Biology Motivation Questionnaire; Glynn et al., 2011) and science identity formation (Estrada et al., 2011). Kruskal-Wallis H analyses indicated that the median time graduate teaching assistant (GTA) instructors within CURE contexts spent posing questions, interacting with students, and following-up on student work was significantly greater than their non-CURE peers, who were mostly engaged in administrative tasks or monitored student work without interaction (p ≤ 0.018 for all comparisons). Similarly, CURE students were engaged in active behaviors (e.g., questioning; making predictions; one-on-one interaction with GTA) more frequently than their non-CURE classmates (p ≤ 0.038 for all comparisons). Principal Components Analysis further suggested that instructor behaviors could be collapsed into two factors: (a) global prevalence of student- vs. teacher-centered behaviors; and (b) low vs. high levels of one-on-one interaction with students. Student behaviors could be aggregated into a single factor describing their level of participation in the laboratory. General linearized mixed modeling approaches indicated that, after controlling for high levels of student-centered instruction and variation in student demographic characteristics, CURE participants exhibited more positive shifts in motivation and science identity formation over the course of the semester than their non-CURE peers. Behaviors indicative of high student engagement in the course were likewise found to result in more positive shifts in motivation within CURE vs. non-CURE contexts. Collectively, these findings provide novel insights into the contextual features inherent of CURE and non-CURE learning environments and the influence of these factors on non-cognitive student outcomes in the domain.

**Students as Annotators: emulating traditional research experiences** Melissa McCartney\*, Florida International University; Diane Ebert-May, Michigan State University [abstract # 61]

Engaging students in the process of research has become a key emphasis in science education. Undergraduate research experiences (UREs), including course-based undergraduate research experiences (CUREs), are one way to actively engage students in the scientific process and have been shown to increase retention and academic achievement. However, not all institutions have the resources to provide these experiences, and, even when they do, the demand far exceeds supply. How can we expand opportunities in scientific research to include a wider and more diverse array of students? We describe a novel method for student engagement in research. “Undergraduates as Annotators” (UAA) aims to emulate the more traditional URE and CURE models by engaging students with recent scientific discoveries and methodologies used to collect and analyze data. First, students deconstruct a research paper connected to the biological concepts taught in their course. Next, students meet with the authors of the research paper to have conversations within the traditional scientific discourse, a practice that further develops students’ understanding of the scientific community. Taken together, the two phases of UAA place students in a research mindset: first by having students work through the data and results of a research project, and then by engaging students in an authentic conversation with a member of the scientific community. By engaging students in these two types of scientific discourse, UAA provides learning about research through participation in the context of course content. In Fall 2017, eight instructors from four institutions participated in the pilot launch of UAA. Students worked in groups to annotate an assigned research paper. When applicable, research papers came from faculty members at the student’s institution. Each student completed a pre/post-course survey, annotated a paper, and participated in a discussion with the paper author (in person, when applicable). The pre/post-course surveys included two 5-point Likert-scale questions on communication and research skills, as well as short answer questions asking students to further describe their experiences with UAA. Our analyses have shown promising results. Gains were observed in research skills (i.e. ability to contribute to science, discuss science with others, understand journal articles, and engage in real-world science research). Self-reported gains in undergraduate student interaction with scientists at their own institutions were also observed, which is a key goal of the participatory learning included in UAA. UAA was designed for students to interview authors of the papers they are annotating, an understudied area of science education. Our Fall 2017 pilot resulted in 24 student-author interviews, generating hours of data to code. While we are just beginning to unpack these data, we will do so through the lens of learning gains seen in UREs and CUREs to see how UAA compares as a novel type of student research experience.

**The Effects of Prompted Group Discussion on Improving Student Science Process Skills** Alex Paine\*, University of Colorado [abstract # 85]

Students find mastery of science process skills such as analyzing data, evaluating data, and experimental design to be challenging. Course-based undergraduate research experience lab courses (CUREs) presumably give students experience learning science process skills, yet little is published on how they improve student competence. Based on a prior pilot study, we suspected that students would make little improvement without explicit practice of these skills, and that group discussion could be critical in this process. We studied 107 students in a first-year CURE course to determine how their experimental design skills changed over time, how students interacted with each other in groups, and whether these interactions affected future problem solving. Students answered modified published questions on experimental design as part of an pre and post assessment. The Expanded Experimental Design Ability Tool (E-EDAT) was used to score their answers. Between assessments, students worked on two assignments, each independently and then as a group, focusing on data analysis and experimental design. For the independent assignments, students were given written prompts to make a claim about data and provide evidence for that claim. In the following lab period, students engaged in group discussion, either with or without additional prompts. In the 15 prompted groups, students were given questions designed to encourage discussion and sharing of reasoning, while in the 13 unprompted groups, students were instructed only to discuss the problem. We analyzed three prompted and four unprompted groups via audio-recording (27 students). Prompted groups spent twice as long in discussion, had twice as many turns of talk, and each student contributed twice as much to the discussion in comparison to the unprompted groups. Despite differences seen between prompted and unprompted groups, students failed to improve their performance on the E-EDAT, scoring an average of 40% both pre and post. Similar to previous reports, the majority of students omit components of design such as sample size and repetition. However, when prompted with a follow up question that encouraged reflection of experimental design, students frequently added considerations for sample size, stated additional uncontrolled factors could affect their conclusions and described to control for multiple variables. Sixty percent of the 53 students who completed both pre and post assessments improved their final score by answering the reflection question. However, only students who engaged in prompted group discussions had significantly higher scores from pre to post + reflection (p<0.05). We will further investigate how, and how often students justify their ideas in these assignments and will use regression analyses to make predictions about the effects of demographic factors, group participation, and individual earlier performance on student’s final reasoning and experimental design abilities.

**Student Thinking and Behavior- Friday**

**Self-Regulated Learning (SRL) in an Introductory Biology Course: Impacts on early academic performance and the activities that promote its change** Holly Swanson\*, University of Rhode Island [abstract # 86]

Bandura’s Social Cognitive Theory describes the process of students shifting from using social sources of regulation, like observation and emulation, to self-initiated sources, like self-control and self-regulation. The process of self-regulation utilizes a reiterative cycle of forethought, performance, and reflection by the student to improve their learning. Through this cycle, the triadic relationship of personal factors (ie. motivation), behavioral actions (ie. the strategies a student uses to learn), and environmental factors (ie. course activities) playout in the social cognitive theory of self-regulated learning (SRL). Research supports that high school academic resources, such as student GPA and the quality of their school curriculum, are among the top predictors of college success and that approximately one third of all entering freshmen are ill-prepared, often requiring remedial work. While some programs, like summer bridge programs, are aimed at remediating select groups of students prior to their matriculation into higher education, efforts need to be made to help all students better regulate their own learning and to be able to do so earlier in their academic careers. Additionally, these efforts need to be made within the context already in existence within an institution. The purpose of this study was to investigate the changes in SRL over the course of the semester in an introductory biology course and its impact on early academic performance. Surveys were administered through an online platform the week prior to the first exam in multiple introductory biology sections and during the last week of the semester. Surveys included scales from the Survey of Academic Self-Regulation and the Strategies for Studying Science Survey. Data from both the survey and from the Office of Institutional Research was used to conduct two analyses: a Path analysis to test the possible mediational impact of student self-reports of SRL between high school GPA and early academic performance in an introductory biology course and an ANOVA to evaluate differences in group means for student changes in survey scores across their agreement or disagreement with the impact of various activities directed at SRL. For the Path analysis, three nested models were tested: a direct model with high school GPA predicting early performance as measured by z-scores for the first exam, a mediational model where survey scales measuring the learning strategies used by students and their level of intrinsic and extrinsic motivation mediating the relationship between high school GPA and exam z-scores, and a full model that included both the direct and mediational effects on early performance. Based on Chi-squared difference tests, the full model best described the data with a good to fair fit [ 2(426, N = 284) = 1017.810, p-value < 0.0001, CFI = 0.772, RMSEA = 0.070, R2 = 0.162 (95%CI = 0.085, 0.239)] and the standardized parameter estimates for intrinsic motivation and high school GPA suggest a small to medium effect size. ANOVA results showed a significant difference [F(1, 92) = 7.28, p = 0.0083] in the mean change in the strategy use scale over the semester based on student agreement (M = 2.44, SD = 5.59) and disagreement (M = -1.61, SD = 6.84) with the helpfulness of activities incorporated into their freshmen orientation course but there were no group differences in changes in motivation for any of the six categories of activities. These findings suggest that in addition to students’ high school academic preparation, their level of intrinsic motivation impacts their early success in an introductory biology course. However, our data also illustrated that none of the activities associated with the introductory biology course or the first-year experience changed students’ level of motivation over the semester. Finally, we will review the literature, listing common activities used in introductory biology courses while discussing intended constructs of change.

**Exploring connections between students’ perspectives on their learning experiences and their views on knowledge and learning in a studio-style science course** Sarah Andrews\*, University of New Hampshire; Eleanor Abrams, University of Massachusetts - Lowell; Serita Frey, University of New Hampshire [abstract # 258]

Student-centered learning approaches have consistently been shown to improve learning, but lecture-based courses still dominate undergraduate STEM education. Therefore, when instructors work to make their courses more student-centered, students may be resistant to these modes of teaching and learning that run contrary to their predominate experiences, particularly if they have internalized these experiences into a belief system about how learning happens. Our research addresses the following questions: 1) What are students’ perspectives on their learning experiences in a student-centered science course?, and 2) Do these experiences contribute to shifts in how students’ view knowledge and learning, particularly as it relates to the roles of instructors, peers, and themselves as learners? This research is grounded in social constructivism (a sociological theory in which knowledge is constructed through interaction with others) and utilizes an intellectual development model (the Epistemological Reflection Model) as a framework. The study context was a studio-style introductory soil science course where lecture and lab were combined and integrated with student-centered active learning strategies. Data were collected during the second year the course was taught in the studio style. Consenting students participated in two semi-structured interviews at the beginning and end of the semester and the Measure of Epistemological Reflection (MER), an open-ended questionnaire, was administered to all 70 students during class at mid-semester. Seventeen students participated in both interviews and completed the MER. Interviews and the MER were analyzed qualitatively. Open data-driven coding of interviews was utilized to identify students’ perspectives on the learning environment and their MER responses were interpreted to identify students’ ways of viewing knowledge. Initial interviews were examined for evidence to support this interpretation, and then compared to exit interviews to look for any evidence that these views had shifted by the end of the semester. Students described active learning (doing, thinking, sharing), the integrated nature of the course (being able to immediately apply what they learned in a lecture to a hands-on activity), a sense of community, and the variety of learning and assessment methods as positively influencing their learning experience. Aspects of the studio-course that they found challenging were acclimating to the new experience and having to rely more on themselves or their peers. Three students expressed changes in their views on the role of peers or instructors and two students described “sneaky learning,” which occurred when they realized they had learned something only when confronted with an assessment; they described feelings of anxiety prior to assessment (because they did not believe they learned anything) followed by relief when they discovered they actually had learned. These findings suggest that the studio structure may promote shifts in how students view knowledge and learning.

**How do students study in STEM courses? Findings from a light-touch intervention and its relevance for underrepresented students.**  Brian Sato\*, UC Irvine [abstract # 101]

With the nationwide emphasis on improving outcomes for STEM undergraduates, it is important that we not only focus on modifying classroom instruction, but also provide students with the tools to maximize their independent learning time. This is particularly relevant for underrepresented minorities (URMs) who may be less familiar with means to self-regulate their learning. Two beneficial study practices include (1) spacing, where an individual’s study time is separated into multiple successive sessions as opposed to a much smaller number of lengthier study sessions, and (2) self-testing, which involves assessing one’s understanding by solving practice problems. Laboratory studies have examined the benefits of spacing and self-testing for enhancing learning, although their value in the classroom has been demonstrated to a lesser degree. The potential impact of these practices in a college-setting is intriguing, especially in light of the need to improve the quality of education with low-cost interventions. In this work, we examine the effectiveness of a light-touch study skills intervention. Rather than alter the course structure to force a certain behavior, we aim to see whether having an instructor repeatedly recommend spacing and self-testing as effective study strategies can alter how students study. We collected a variety of data, including pre- and post-course survey responses, course grades, and student demographic data, from three sections of a second-year molecular biology course taught in spring 2016 and 2017. During both years of the study, students in one of the course sections received the study skills intervention, which consisted of a brief lecture on spacing and self-testing and consistent reminders to implement these strategies throughout the class. We aim to address the following research questions: 1. How do students in a large enrollment STEM course study, and to what extent do students report utilizing spacing and self-testing strategies? 2. Does using spacing and self-testing strategies vary for URMs and non-URMs? 3. Does a study skills intervention focused on spacing and self-testing impact student use of these strategies? 4. Does self-reported utilization of spacing and self-testing correlate with course performance? From our analysis, we learned that students report utilizing both beneficial and ineffective study practices, and that surprisingly, URMs report utilizing self-testing to a significantly lower degree (59.0% vs. 66.4%). We found that our intervention positively impacted student self-reported use of spacing and self-testing with students in the intervention section either maintaining or increasing their use of both strategies at the end of the course, compared to students in control sections who exhibited sharp decreases in their use of spacing and self-testing. The intervention also helped to partially ameliorate the gap in self-testing between URMs and non-URMs. The impact of the intervention was specific for these study strategies, as the use of flashcards and condensing notes, two other commonly reported strategies not discussed in the intervention, were similarly used at the end of the class by students in intervention and control sections. Similar to past work on spacing and self-testing, we also confirmed that self-reported usage of these strategies correlates with a higher course grade (an increase of roughly 5 percentage points). And while we observed a decreased use of self-testing for URMs, those who did report self-testing earned similar course grades as non-URMs who also self-tested, while there was a significant drop in performance for URMs who did not self-test relative to non-URMs who also did not self-test. Overall, we would encourage instructors to dedicate class time towards discussing the merits of beneficial study practices, especially for students that have historically underperformed in STEM disciplines.

**Using social network analysis to measure the dynamics of student interactions in group work** ALBERT CHAI\*, UNIVERSITY OF CALIFORNIA SAN DIEGO; JOSHUA LE, UNIVERSITY OF CALIFORNIA SAN DIEGO; ANDREW LEE, UNIVERSITY OF CALIFORNIA LOS ANGELES; Stanley Lo, University of California San Diego [abstract # 74]

Group work forms the foundation of many evidence-based instructional practices, such as peer instruction, peer-led team learning, and process oriented guided inquiry learning. Extensive evidence indicates that group work promotes learning measured by quantitative student outcomes, and existing learning theories, such as constructivism, can provide broad explanations for the theoretical basis of group work. However, the exact nature of how students interact and what about these interactions that promote student learning remains unclear. Some researchers have studied the qualitative nature of student discussions in group work through discourse analysis or examination of arguments formed by students. In this study, we developed and used a quantitative methodology to study the dynamics of student interactions in group work. We established our methodology through observations of 15 small groups of 4-8 people, including peer facilitators, in a peer-led team learning environment at a private research university. Data collection went through three iterations of refinement, and data were ultimately recorded in a question-and-response format that tracks the sequential order in which students talked. For data analysis, we treated the discussion as a network, with the sequential talking order as transitions between individuals. We adapted tools and parameters traditionally used for analyzing large social networks to quantify our discussion networks. Examples of these parameters include density, average degree, centrality, and subgroups. Density describes the extent to which students interact with one another in the discussion network on a standardized scale from 0.0 to 1.0, with 1.0 indicating that all students are talking with one another. Average degree refers to how many other persons each individual interacts with on average. Centrality measures how dominant the most central person is in the network based upon that person’s interactions with others. Subgroups can be mathematically identified when clusters of students interact more among themselves than with other students. We tested our final methodology on seven out of the 15 observed groups. Results were processed using a custom R script that we developed, which calls on functions from the igraph library package. From our results, even with the small sample size, we identified three archetype cases that have clear implications on learning and teaching. First, in a group of five people, the peer facilitator was a dominant figure (with high centrality) in the discussion network; the density of the resulting network is 0.4, and the average degree is 1.6, indicating that most students are talking with the facilitator but not with each other. Underrepresented minority students had quantifiably fewer transitions, meaning that they did not speak as frequently as others. After more analysis, we experience contrasting results. In contrast, in a second group of six people, the peer facilitator has a low centrality in the discussion network; the density of the resulting network is 1.0, and the average degree is 5.0, indicating that all students, including underrepresented minorities, are interacting with one another in a relatively equitable fashion. Finally, in a third group of eight people, the discussion network has intermediate features between the first two groups. More interestingly, this network is divided into two subgroups based on gender. The subgroups were detected by network mathematics and were not immediately obvious from qualitative observations. In this study, we developed a quantitative methodology to examine the dynamics of student interactions in group work, which is increasing in prevalence in undergraduate biology education. In our presentation, we will also discuss the implications on learning and teaching, such as training for peer facilitators or learning assistants who lead group work and creating inclusive learning environments that engage students equitably.

**Students reading real science: primary literature in the introductory biology classroom** Matthew Kararo, Florida International University; Melissa McCartney\*, Florida International University [abstract # 60]

Reading primary scientific literature is an authentic scientific practice requiring skills that are rarely developed in introductory undergraduate students. One way to introduce these skills is by incorporating annotated primary scientific literature as a pedagogical tool in first-year undergraduate science classrooms. Science in the Classroom (SitC), a collection of annotated research papers, makes primary literature more accessible to students and educators. SitC is an online tool that uses the original text of research articles along with a “learning lens,” designed to selectively highlight different parts of the text including: glossary, previous work, author's experiments, conclusions, news and policy links, and references and notes. Annotations provide an educational scaffold that helps students deconstruct scientific papers, giving them an understanding of experimental design and the presentation of results and conclusions. With the scientific language barrier removed, students are able to see how the authors identified a question, how data was collected and analyzed, and how the next question(s) was proposed, essentially introducing them to the non-linear and iterative nature of science. With the cooperation of introductory biology teachers and students at our institution, we have developed an implementation protocol for the use of SitC as a pedagogical tool using an iterative process over three semesters. We collected observational data (i.e. did students work in groups, how did the instructor introduce the paper), quantitative data (i.e. how long did students remain on the research article page, what annotations did they click on the most), open-ended questionnaire data (i.e. which learning categories were useful, how did students utilize the learning lens tool), and conducted multiple focus groups with students who participated in the SitC activity and were familiar with the annotation tool. Data collection was focused on both classroom implementation and additional SitC tool development. Research design consisted of introducing the SitC tool, allowing student’s time to read the same annotated research paper, and asking students to complete a post-activity questionnaire. Initial implementation efforts involved researchers entering the classroom and providing a lesson on how to use SitC as a way to read primary scientific literature. It was then determined that the next iteration should focus instead on the classroom educator with which the students had already developed a relationship. This iteration involved the classroom educator introducing and demonstrating the SitC annotation tool for their students prior to the students reading the annotated primary literature article. Again, the protocol was iteratively improved when an educator was observed that modified this implementation and incorporated digital clicker content questions as a follow-up activity, rather than the embedding content questions within the post-activity questionnaire. This final method of further integrating SitC annotated papers into the introductory biology curriculum will be our model of implementation moving forward. Future efforts will propagate this implementation protocol to additional general biology sections at our institution as we further refine the SitC tool and activity evaluations.

**SATURDAY**

**Long talk- Bruininks 220**

**A Multi-Year Study of Student Anxiety in Introductory Biology Classrooms**

Ben England\*, The University of Tennessee; Beth Schussler, "University of Tennessee, Knoxville"; Jennifer Brigati, Maryville College [abstract #53]

Recent shifts toward student-centered teaching in biology courses have generated benefits in student exam performance and lower course failure rates. However, this pedagogy can differentially impact students’ classroom experiences, resulting in varying emotional experiences for each student. An emerging area of research is student anxiety in lower-level biology courses that use student-centered learning. Though anxiety is not uniformly bad, negative effects such as reduced class performance and motivation and increased attrition have been documented. Investigating anxiety in introductory biology classrooms is important for understanding the sources of this anxiety, its impacts on students, and factors that may lessen its negative effects. Over the last 2.5 years, our lab has been studying student anxiety in introductory biology classrooms at a large research university. This work has been guided by the Control-Value Theory of Achievement Emotions, which posits that student emotions arise from one’s appraisal of the learning environment and can differentially impact student achievement. Our goal over this time period has been to identify causes of student anxiety and the impacts of this anxiety on course performance and persistence in the major. We started with two research questions: 1) To what extent do students experience anxiety towards active learning practices? & 2) Does anxiety impact performance or persistence? In spring 2016, we surveyed students in three introductory biology classes (N=327) and conducted student interviews (N=12) to collect initial data. We found certain types of active learning (cold calling and being asked to volunteer an answer to a question) caused more student anxiety than others (clickers, group discussion, or worksheets) (ANOVA, p<0.001, Cohen’s f=0.23), and overall “general” student anxiety and some active learning anxieties varied by instructor. Student interviews revealed that peer communication and graded formative assessments caused them anxiety. Student self-reported course performance and persistence were negatively impacted by high general anxiety, but not active learning anxiety. In 2016-2017, we then investigated whether other types of classroom anxiety (test, social, and communication) impacted student performance or persistence. We measured these anxieties at the beginning and end of the semester, and investigated actual final grades and persistence in the major (N=337, 7 classes). General, test, and communication anxiety were negatively correlated to student final course grade (MANOVA, p<0.01, Cohen’s f=0.38), as was anxiety about volunteering to answer a question (Kruskall-Wallis, p<0.0125). General anxiety at the beginning and end of the semester was negatively related to persistence (MANOVA, p<0.01, Cohen’s f ≥0.45) and was higher for freshmen, females, and those who had taken fewer AP courses. In 2017-2018, we explored what aspects of the class aligned with student anxiety. Students self-reported that 30-48% (varied by class) of their general anxiety was caused by active learning (N=693, 6 classes), with tests/quizzes, how to study, and bad grades being the most common other causes of anxiety. We also investigated instructor impacts on anxiety. High levels of autonomy-supportive practices (a motivation construct) and non-verbal immediacy (a communication rapport measure) were both associated with lower general and test anxiety, though high non-verbal immediacy was associated with an increase in communication anxiety. We interpret these results in the framework of the Control-Value Theory of Achievement Emotions: students appraise instructor support and communication, classroom pedagogy and assessment, and anticipation of their performance, generating an anxiety level that then impacts their course performance and persistence in the major. Our next steps are to create interventions related to classroom practices and student coping to modify anxiety and increase student success.

**Long talk- Bruininks 230-**

**It Takes a Village: Beyond Training Mentors to Empowering Research Trainees**

Janet Branchaw\*, University of Wisconsin - Madison; Amanda Butz, University of Wisconsin - Madison; Amber Smith, University of Wisconsin – Madison [abstract # 132]

Mentored research experiences are important to the success and persistence of novice trainees in STEMM, especially those from diverse backgrounds. Recent interventions to improve the quality of research mentoring relationships have focused on training mentors; while interventions focused on building the capacity of trainees to effectively navigate their mentoring relationships and succeed in their research experiences have been less common. To address this gap, we developed and tested a theoretically informed, active learning curriculum and developed and validated a new research trainee learning assessment tool. Four components of the project will be presented: 1) a framework of trainee development and associated learning objectives, 2) a curriculum that has been pilot tested nationally, 3) a validated mentor-trainee paired assessment tool; and 4) a facilitator training workshop. 1) The framework of research trainee development organizes and aligns the curricular materials, which include over 100 activities, and a learning assessment. Development of the framework was guided by the literature on undergraduate and graduate research trainee development and it was optimized for use in STEM research training programs by scientist practitioners. It includes seven areas of research trainee development: Research Comprehension and Communication; Practical Research Skills; Research Ethics; Researcher Identity; Researcher Confidence and Independence; Cultural Awareness and Skills; and Professional and Career Development Skills. 2) A team of 22 scientists and social scientists from diverse institutions across the country developed the curriculum. Pilot testing of the curriculum took place with undergraduate and graduate trainees in STEMM fields during summer and fall 2017, and spring 2018. Data from 281 trainees across 12 sites in 20 unique implementations from summer and fall 2017 were used for the preliminary analysis included in this abstract. Of the trainees who provided demographic information (n = 189), 58% identified as White and 60% reported their gender as female. Facilitators (n = 37) primarily identified as White (76%) and 68% reported their gender as female. All curricular activities tested were rated as effective by both facilitators and trainees. More than 81% of trainees were likely or very likely to recommend the activities and found them to be a valuable use of their time. Trainees who completed the research trainee learning assessment (N = 85) reported gains in all areas of development. 3) Originally, a 44-item assessment of trainee learning was developed and pilot tested with 193 trainees during spring and summer 2017. The first iteration of the scale showed evidence of structural validity and internal consistency (factor loadings .28 - 92; α .74 - .95). Assessment of item factor loadings, participant feedback on items, and ongoing development of the curriculum framework informed revision of original items and the creation of 14 additional items, yielding a revised assessment instrument with 56 items. Final validation of the revised instrument using confirmatory factor analysis will be performed in late April 2018 on data collected from approximately 450 trainees and their mentors. 4) Nearly all (96%) of the participants in workshops to train research program directors to build and implement custom curricula using the materials rated the quality of the workshops as very good or excellent (n=45). In a retrospective assessment, participants' confidence in their ability to utilize the curriculum and supporting resources went from an average of 2.82 to 6.27 on a scale ranging from 1 (not at all confident) to 7 (extremely confident), t(43) = 15.94, p< .001. Future research will investigate the effectiveness of strategies to disseminate and implement the curriculum and the differential impact that specific curricular activities have on different populations of students across career stages and institution types.

**Teaching Assistancts- Saturday morning**

**Stuck in the middle again: Investigating teaching and research anxieties in biology graduate teaching assistants**  Miranda Chen\*, Univ.Tennessee, Knoxville [abstract # 9]

Graduate students in the United States are experiencing increased levels of anxiety, affecting their overall mental health and attrition in graduate programs. One in three graduate students report being depressed, a rate six times higher than the general public. Yet we are only beginning to understand what contributes to graduate student anxiety. Biology Graduate Teaching Assistants (GTAs) occupy an “ambiguous niche,” with simultaneous roles as teachers, researchers, students, and employees. This can make it difficult to form a coherent sense of identity, contributing to anxieties regarding teaching and research responsibilites. GTAs teach over 91% of freshman Biology labs and discussions nationally. If GTA anxiety impacts their teaching, it may also broadly influence the quality of undergraduate education. As part of a longitudinal study, we investigated Biology GTA teaching and research anxiety at a large research-intensive southeastern university. We interviewed 23 Biology GTAs to probe their teaching and research anxiety and the causes of these anxieties. As the first qualitative study examining Biology GTA anxiety, our findings are framed by social cognitive career theory (SCCT). SCCT was developed to identify the cognitive variables that influence a person’s career interests and trajectory. We predicted that GTA teaching anxiety, for example, would mediate how learning experiences (e.g. teaching in a classroom) would build self-efficacy (e.g. “Can I teach well?”) and outcome expectations (e.g. “What will happen if I teach poorly?”). These two variables would then have downstream effects that may impact a GTA’s career interest, choices, and performance with regard to teaching- or research-focused careers. To identify themes and categories regarding teaching and research anxiety, thematic analysis of interview transcripts was conducted using Fall 2016 interview data from the 23 graduate students. Each interview was ~60-90 minutes long with mostly experienced GTAs, 70% female, and 74% Caucasian participants from across three Biology departments. Four GTA anxiety themes (applying to both teaching and research) were identified from the data set: perception by others, impact on others, lack of control or self-efficacy, and role tension. Anxiety related to how others perceive you and your work was the most prominent theme, with 82% and 91% of the 23 participants having this theme emerge for teaching and research, respectively. This theme reflected a feeling of being unable to successfully fulfill the expectations of oneself or others in that role of teacher or researcher. An example of this theme was when a GTA feared “looking dumb” in front of students or faculty. The second theme was anxiety related to impact on others, or the consequences on others if you were an incompetent teacher (87%) or researcher (48%). Impeding student learning or disappointing your advisor were examples of this theme. Third, anxiety rooted in lack of control or self-efficacy for both teaching and research (100%) was related to GTA ability to carry out a task such as handling classroom emergencies, grading fairly, or running data analysis. Last, GTAs reported anxiety about role tension, which was associated with not having enough time to attend to priorities such as teaching or research (52%). Results from this study demonstrated that a GTA’s identity in teaching and research may be in conflict, reducing self-efficacy and outcome expectations related to each, and enhancing feelings of not meeting the expectations of students or advisors. Further qualitative analyses will examine the coping strategies employed to try to reduce these anxieties. In understanding the foundations of GTA anxiety, this work can inform future professional development or intervention activities for GTAs and encourage greater awareness and dialogue about the impacts of GTA mental health issues in academia.

**Responsive Teaching Training for Teaching Assistants: Examining Shifts in Framing** Matthew Simon\*, Tufts University; Julia Gouvea, Tufts University [abstract # 78]

Responsive teaching in science has been shown to enhance students’ understanding of scientific content (e.g, Coffey et al, 2011), promote student agency and voice in the classroom (e.g., Hammer, Goldberg & Fargason, 2012) and elicit more equitable student participation (e.g., Warren et al, 2001). The ability to notice and interpret students’ beginning scientific ideas is at the core of responsive teaching, and prior work has shown that this skill can take practice to develop. Instructors may too quickly dismiss potentially productive ideas if they are focused on superficial indicators of correctness. Experts in responsive teaching can notice and draw out these beginning ideas, allowing students to develop and refine their thinking. Our primary research aim was to investigate how novice instructors – graduate and undergraduate teaching assistants (TAs) – engage in noticing and interpreting student thinking in biology. We developed a professional development (PD) course for biology TAs that foregrounded responsive teaching practices: noticing, attending and responding to the substance of student thinking and ideas. The PD ran concurrently with the TAs’ teaching assignments as introductory biology lab instructors and was structured around TAs sharing video clips from their lab sections and discussing the substance of student ideas and activities. In this talk, we address two questions: 1) What do TAs notice when asked to view and interpret video of student thinking? 2) How do the things TAs notice change with different forms of student activity? We conducted an analysis of the discourse of the nine participating TAs during a 75-minute class session early in the PD (week 2) as TAs watched a clip that featured different types of student activity (discussing ideas and designing experiments). We used evidence from TA talk and tone to describe two different “frames” that they applied to this activity (Scherr & Hammer, 2009). An evaluative frame emphasized judgments about the correctness of students’ ideas or activity (e.g. “I think the problem [students] have is like, they're trying to interpret results that they don't have instead of trying to do a design to get the results”). A descriptive frame emphasized identifying and attempting to find the scientific merit in students’ thinking (e.g. “You have one guy developing the model, and you have inputs from the rest of the group, and therefore you have this group learning”). We found that all nine TAs applied both frames during the discussion of the video. Each TA made between one and thirteen utterances that indicated a descriptive frame and between one and nine utterances that indicated an evaluative frame. We also found that the context of the student ideas being discussed changed the frequency of statements coded as descriptive or evaluative frames. When students were articulating conceptual ideas or proposing hypothetical results, TAs were more likely to notice and attend to the substance of that thinking (descriptive). However, when students were engaged in experimental design, TAs were more likely to judge the ideas for viability and correctness (evaluative). Our results make two contributions to the BER community. First, we provide preliminary evidence that novice TAs with minimal training can prioritize noticing student thinking, suggesting that PD focused on responsive teaching may be a promising and accessible approach for biology TAs. Second, we have identified specific types of student activity that may impact the degree to which TAs are able to attend to student thinking, suggesting that it may be important to design PD that begins with more discussion-based clips of student thinking before considering how to interpret students’ ideas in the context of experimental design. Experimental design may be a more challenging activity to interpret and may require more facilitator support for TAs to stay focused on interpreting and responding to the intellectual work of students.

**Impact of a graduate teaching assistant (GTA) training program on GTA approaches, self-efficacy, and knowledge of student-centered learning** Heather Vance-Chalcraft\*, East Carolina University [abstract # 133]

Graduate Teaching Assistants (GTAs) deliver a substantial amount of instruction to biology undergraduate students during their first year, yet relatively little attention has been paid to the training of graduate students for this responsibility. GTA training, when provided, can be a time-intensive effort yet the effectiveness of this training is rarely assessed. I examined the impact of a GTA training program on GTA self-efficacy and GTA knowledge of inquiry-based labs and student-centered learning. For this study, GTA training included a week-long intensive summer workshop and weekly laboratory preparatory meetings with a specific 20-minute professional development piece added. Multiple published, validated surveys were administered before and after GTA training to assess changes in teaching approaches, knowledge, and self-efficacy. The Approaches to Teaching Inventory (ATI) was used to measure GTA approaches to teaching from a conceptual change/student-focused approach to an information transfer/teacher-focused approach. The STEM GTA Teachers’ Sense of Efficacy Scale (STEM-GTA TSES) was used to assess GTA self-efficacy in establishing an active and engaging learning environment and GTA self-efficacy about instructional strategies. Finally, the subset of questions designated as inquiry-focused from the Knowledge Survey was used to assess GTA perceptions of their understanding of concepts and topics related to inquiry instruction. A comparison of pre- to post- survey results showed that GTAs, on average, firmly identified with an information transfer/teacher-focused approach before training but moved closer to a conceptual change/student-focused approach after training. Self-efficacy of experienced GTAs was less impacted by training than GTAs with less experience. New GTAs, however, were variable in how they responded to training and some even became less confident about their ability to create an engaging learning experience or use appropriate instructional strategies. GTA knowledge gains were largest in the areas of providing effective explanations and pre-class preparations. These results provide insight on which GTAs are most impacted by this training program and in what areas the training is most effective. For example, GTA knowledge and self-efficacy about student-centered learning did not improve as much as other areas. Thus, changes to the GTA training program may be required to assist GTAs in internalizing the value of student-centered approaches. Follow up work is examining the impact of GTA training on teaching practices, as determined by classroom observations. Additionally, a planned expansion of this project to another institution will allow me to determine how generalizable these results are to other GTA training programs.

**Using Activity Theory to Examine Active Learning in Learning Assistant Supported Undergraduate STEM courses** Laurel Hartley\*, Cu Denver; Robert Talbot, University of Colorado Denver; Jeffrey Boyer, North Dakota State University; Laird Kramer, Florida International University; Hagit Korneich-Leshem, Florida International University; Leanne Doughty, University of Colorado Denver; Mary Naeyma, Florida International University; Paul Le, University of Colorado Denver; Amreen Nasim-Thompson, University of Colorado Denver; Andrew McDevitt, University of Colorado Denver; Chelsey Grassie, University of Colorado Denver; Hannah Huvard, University of Colorado Denver [abstract # 266]

The use of active learning methods has been shown to be effective in undergraduate science teaching, as has the use of Learning Assistants (LAs). As with many innovations in teaching, it is difficult to parse the effect of any one intervention. In our research program, we observe, characterize, and interpret the active learning tasks and methods employed in a large sample of LA supported science courses at three different research universities. We study both the active learning tasks and the interactions that occur in the LA supported classroom. Because we are interested in both student level outcomes and the nature of social construction of knowledge, we have chose the metatheoretical framework of Activity Theory to guide data collection and analysis. Applying Activity Theory, we view a course as a system. The core activity within the system is “learning science”. The “subject” is the student, the “object” is student success, and the “outcome” represents the way student success is operationalized (e.g., learning gains, course grades). The “mediating artifacts” are the active learning tasks meant to promote learning and engagement. In this session, we will focus on elements or interactions between elements in our activity system, which we hypothesize are related to student outcomes. Those elements are: 1) the mediating artifacts (active learning tasks), 2) the actions that LAs engage in when interacting with students, 3) faculty beliefs and practices related to engagement with LAs and students, and 4) the engagement of students in classroom interactions. We collected data in over 50 undergraduate STEM courses at three institutions. Data sources included 1) classroom observations using the Classroom Observation Protocol for Undergraduate STEM, 2) active learning tasks (e.g., worksheets, clicker questions) 3) video data of LA-student interactions using Point of View (POV) cameras worn by LAs, 4) data on who students interacted with in class and the perceived value of those interactions for their learning, 5) data about the classroom climate, 6) instructor perceptions of their LA supported teaching based on the Teaching Practices Inventory and follow up interviews, 7) student outcomes including grade in course and gain on conceptual assessments, 8) student attributes (e.g., race, ethnicity, gender identity, major). Not all data sources were collected from all sample classrooms. Data from classroom observations revealed high variability among courses with respect to time spent in interactions (student-student, student-LA, and student-faculty). Analysis of active learning tasks using Bloom’s Taxonomy and the Three Dimensional Learning Assessment Protocol suggests that courses with lower concept inventory gain scores and less time spent in active learning utilized no or very few activities that were 3 Dimensional (i.e., core idea, cross-cutting concept, and a scientific practice) or rated higher than “understand” on Bloom’s Taxonomy. Student social network survey analysis shows that the odds of passing a course increase by 10% or 8% when students report directly communicating with an LA or the instructor respectively. LA point of view video analysis revealed that facilitation of activities (LA directed or LA guided) appears to be the main category of LA action in the classroom. However, LAs also interact with students in other ways such as offering hints and tips to succeed in the course and discussing non course related events. This suggests that LAs play additional roles in the classroom beyond just facilitation of task related discussion. Analysis of faculty interviews reveals that instructors view LAs as valuable primarily for providing student support and feedback about instruction. However, LAs also contribute to changes in instructor ideas about their students’ lives and student learning, and LAs co-create active learning tasks in some contexts.

**Preconceptions/Misconceptions- Saturday morning**

**The effects of an introduction to biological research course on novice students’ views on the nature of science** Kelly Schmid\*, Syracuse University; Ryan Dunk, Syracuse University; Jason Wiles, Syracuse University [abstract # 67]

Previous research has shown that understanding the aims, processes, history, and philosophy of science (a branch of knowledge collectively referred to as the nature of science [NOS]) is important for understanding and acceptance of a variety of scientific phenomena, perhaps especially those that are socially – but not scientifically – controversial topics such as climate change and evolution. Our prior research has shown that NOS understanding is one of the most important factors in explaining variation in evolution acceptance among university students, especially among those who have completed a year of biology instruction. In addition, change in acceptance of evolution over a year of introductory biology was most significantly correlated with change in student understanding of NOS, particularly with regard to understanding science as amoral, unified, and testable. Here, we seek to extend our work through a qualitative inquiry into how an introduction to biological research course early in students’ undergraduate education may impact their understanding of the nature of science. The goals of the course were to give a broad introduction to biological research; to showcase the types of research being done in the university’s biology department; and to promote skills in reading, writing, and discussing science. This was a small, seminar-style course in which students read and summarized scientific articles, engaged in small and large group discussions, explored faculty research labs, and learned to write about science for both general public and scientific audiences. To assess the impact of this course on students’ NOS conceptions, four questions from the Views on the Nature of Science- C (VNOS-C) questionnaire were administered at the beginning and end of the semester. Specifically, the students were asked to define science and distinguish science from non-science, describe how scientific theories may change, explain their views on the sociocultural aspect of scientific inquiry, and discuss the possibility for creative input in scientific investigations. For this project, we used a grounded theory framework to explore students’ conceptions of the nature of science, with the additional goal of identifying student misconceptions of NOS. We undertook our qualitative analysis by reading all responses, then used a random subsample for initial coding. Initial coding was done independently, and all coders met to confer and review codes. From there, all responses were coded using a constant comparative method until saturation was reached. We then took the codes and developed emergent themes, and compared the themes and codes from the beginning of the semester and the end of the semester. Here, we present the results of this effort, with special attention given to changes in early science students’ conceptions of the nature of science from before to after the introduction to research course. This study will allow us to better understand how engaging in this type of course early in their undergraduate career might help develop students’ understanding of the nature of science in preparation of future science courses.

**Mixed Students Ideas about Tracing Matter across Biological Scales in the Context of Human Weight Loss** Kamali Sripathi\*, MSU; Rosa Moscarella, University of Massachusetts Amherst ; Rachel Yoho, MSU; Hye Sun You, MSU; Ross Nehm, Stony Brook University; Mark Urban-Lurain, MSU; Kevin Haudek, MSU; John Merrill, MSU [abstract # 199]

Undergraduate biology education is undergoing a transformation, as outlined in “Vision and Change in Undergraduate Biology Education.” (AAAS, 2011). This report has identified “pathways and transformations of energy and matter” as a Core Concept in biology, underscoring the complexity of molecular processes. Wilson et al (2006) studied the tracing of matter as a strategy that students can learn and apply to understand such processes. To build upon this and other work on student thinking about matter and energy, we investigated the following research question: What molecular mechanisms and types of matter do students invoke to explain the process of weight loss? Our group created a constructed response (CR) question designed to encourage connection of cellular respiration to weight loss. CR assessment items elicit more complete pictures of student thinking than forced choice items (e.g., Birenbaum and Tatsuoka, 1987). We collected 1192 student responses from introductory undergraduate biology courses from 3 public research universities. Using emergent coding on this data set, we developed an analytic scoring rubric with 7 categories that included both students’ normative and non normative ideas about weight loss. We trained six biology expert scorers on the rubric, and divided the scorers into pairs. To refine rubric definitions, each pair was assigned sets of 100 student responses to score, and disagreements were resolved by an external tiebreaker. Pairwise scoring of response sets continued until the average Cohen’s Kappa for each rubric category was 0.6 or above (average = 0.7; range 0.69 - 0.82). Our rubric categories revealed how students did or did not trace matter and mechanisms in weight loss (category names italicized and in parenthesis). The nature of analytic rubrics allows for tracking of multiple categories within each response. We found that 47.0% of responses correctly traced matter by stating carbon dioxide as a byproduct of fat breakdown (Correct Products), and 12.7% of responses correctly described cellular respiration as the means of mass transformation (Molecular Mechanism). In 54.9% of responses students stated that the weight was exhaled (Exhalation). Our rubric also captured incorrect or vague ways of tracing matter. For example, 22.8% of responses stated that mass left the body through excretion (Excretion and Waste); 33.2% of responses unclearly stated that the mass was somehow converted to energy (Matter to Energy) without specifying a cellular process. Additionally, 25.9% of responses stated that mass was used up by exercise (How to Lose Weight), again without cellular explanations. Because our rubric spanned multiple scales from organismal (e.g., Exhalation) to cellular (e.g., Molecular Mechanism and Correct Products), we were curious to see how students combined these scales in their responses. We found that 36.2% of responses linked ideas about organismal-level Exhalation and cellular Molecular Mechanism or Correct Products. We also found less correct co-occurrences of ideas across scales. For example, 11.5% of responses had ideas about Correct Products and Excretion and Waste, and 12.6% of responses discussed ideas about vague Matter to Energy conversions and How to Lose Weight. Such heterogeneity was dominant in our data set: overall, we found that 69.2% of responses contained 2 or more ideas from our rubric categories. Our analyses thus reveal the heterogeneity in student thinking about weight loss: students typically have many ideas of varying correctness about biological processes across scales. The item we developed has already been used at a variety of institutions, demonstrating its future widespread usability. Our results underline the need for targeted instructional interventions aligned with the heterogeneity in student thinking. Our work also serves as an important foundation for future learning progressions monitoring the development of student proficiency of this core concept.

**Context as a source of student difficulty in human physiology**  Tara Slominski\*, North Dakota State University; Jennifer Momsen, North Dakota State University [abstract # 236]

Students struggle to succeed in Human Anatomy and Physiology (HA&P) for several reasons. Many students enter science classrooms with naïve, informal knowledge constructs that conflict with the concepts they are learning in class. These ideas may originate from observing natural phenomena in the real world and often conflict with formal instruction. As a result, a student must undergo considerable conceptual conflict and reformation of that knowledge if they are to be a successful learner. Findings from earlier research on student difficulties suggest students may rely on intuitive reasoning and experiential knowledge to frame their thinking about physiology. Superficially, building links to existing knowledge structures may help students reason about biological phenomenon; however, this type of framing may lead students to make incorrect assumptions and reason inappropriately. Although previous research provides some evidence that students use experiential or intuitive knowledge to shape their reasoning about neurophysiology, the results from this single study are limited. My current research directly investigates the role of students’ intuitive ideas in HA&P classrooms. Specifically, I am researching if, when, and how students use intuitive reasoning strategies when presented with a physiology problem or scenario and how these reasoning strategies affect student understanding of HA&P content. I used surveys and interviews to investigate how the context of human physiology impacted student reasoning. Through a collaboration with physicists, I developed a set of isomorphic tasks that asked identical questions, presented in different situational contexts. By comparing student responses in each context, I can identify the impact of context on student reasoning. In addition, this experimental design enabled me to control for student ability and isolate the effect of physiology context. I collected data in multiple iterations of HA&P and algebra-based introductory physics courses at large, public institution, resulting in over 600 data points. In addition, I conducted semi-structured, think-aloud interviews with 12 HA&P students to substantiate the inferences made from our survey data. Each participant was given one set of the isomorphic task and asked to reason through the question set. During the interviews, students’ explanations were probed to gather deeper insight into their reasoning patterns. Quantitative analysis suggests assessment context does affect the way a student answers a problem. Students ranked the fluid flow rates of three systems differently depending on whether they were reasoning about physics or biology (X-squared = 23.456, df = 1, p-value = 1.278e-06). Based on the differences in rankings, we believe students may be vulnerable to context – or the surface features – of an assessment. Through analysis of the written component of our isomorphic prompt, we observed distinct differences in the language students used to explain their reasoning. Our early analysis suggests student use of teleological reasoning is affected by assessment context, with more instances observed in the biological setting opposed to the isomorphic physics setting. We support this evidence of student reasoning by triangulating the written responses with the ranking task and the semi-structured interviews with HA&P students. These data provide new insight into how students reason about human physiology and the role of item context in assessment. More importantly, the results from this study bring to light the unique challenges students experience when learning physiology.

**Process versus Object: How do students think about mutation?** Fangfang Zhao\*, University of Minnesota; Anita Schuchardt, University of Minnesota [abstract # 37]

Mutations as the origin of variation underlie students’ understanding of two core areas of biology: evolution and inheritance. Textbooks and traditional structure of teaching mutations emphasize the status of a mutation as a static object instead of the process of how a mutation is generated. Educational researchers have theorized that, in general, students tend to ascribe object properties to processes. If biology students tend to perceive mutations primarily as objects, subjugating the processes that produce them, they may have difficulty with fundamental evolutionary or inheritance concepts. While much research has been done on students’ ideas about mutation in the context of natural selection, students’ thinking about the object/process nature of mutation has been underexamined. This study explored students’ responses to an open-ended question, designed to elucidate their object/process thinking: Explain how mutations originate in populations of [organism]. Sixty-four biology majors enrolled in an introductory biology course answered this question after they had received instruction in both mutation and DNA replication. Because it was thought that context could matter, half of the students received a question that asked about lizards and half received a question that asked about bacteria. Codes were developed based on student responses (Cohen’s kappa = .85, indicating an excellent inter-rater reliability). Three common themes emerged: a) students mentioned a mechanism involved in mutation production, b) students defined or categorized mutation, and/or 3) students discussed the ways in which mutations are spread through a population. For responses that mentioned mechanism of mutation and the spread of mutation, which are both processes, a second round of analysis was performed to compare the ways in which students described these processes. Specifically, did students tend to discuss mechanism or spread as a defining characteristic, or as a process which included steps and interactions between relevant objects? Our analysis showed that 23 of 64 (36%) students did not mention a mechanism for the origination of mutations in a population. Instead they defined the types of mutations and/or described the ways mutations were spread. For those who mentioned a mechanism for producing mutations, only 13 of 41 (32%) students included steps or interactions in the DNA replication or induction process. For those who described the process of spreading the mutation in the population, 26 of 33 (79%) students were able to include steps and/or interactions. No substantial difference was seen for the object/process analysis between students who received the bacteria or the lizard context. Two conclusions were reached from this analysis. First, in response to an open-ended question about how mutations originate, only twenty percent of 64 students described the process of mutations production by including steps and interactions between relevant objects. Most students tended to treat the process as a defining characteristic of mutations or treat mutations as something that just appeared without an associated mechanism. Second, many students were capable of discussing the steps and interactions in the spread of mutations, but they didn't do this with the mechanism of mutations. The tendency to favor the characterization of mutation as an object as opposed to a process may have important implications for students’ understanding of other areas of biology especially those that require an understanding of how mutations are generated (e.g. the ontogeny of cancer, or sources of variation in evolution). Therefore, we suggest reexamining the traditional structure that educates students about mutation, perhaps placing more emphasis on the process as opposed to the categorization of mutations as a way to counterbalance students’ tendencies to objectify processes.

**Identity- Saturday morning**

**Investigating the Instructional Transition of Students From High-School to First-Year College STEM Courses** Emma Toth\*, University of Maine; MacKenzie Stetzer, University of Maine; Marilyne Stains, University of Nebraska-Lincoln; Brian Couch, University of Nebraska-Lincoln; Michelle Smith, University of Maine [abstract # 129]

Recent national reports have cited ongoing issues in undergraduate STEM education. Fewer than half of first-year undergraduate students who start in STEM fields graduate with a STEM degree six years later. Most of this attrition occurs between the first and second year of college, and students often cite instructional practices used in introductory college courses as a prominent reason for leaving. Furthermore, students from backgrounds that are underrepresented in STEM fields, including first-generation college students, leave STEM majors at higher rates than their classmates. Recent data show that the instructional practices used in introductory college STEM courses differ significantly from those used in high school science classes, suggesting that incoming college students may hold expectations that are not well aligned with actual instructional practices. To more fully understand this prediction, data were collected from online surveys given to students enrolled in large introductory STEM courses at two universities; including biology, chemistry, and physics courses. Pre-semester and mid-semester surveys asked students about their expectations and perceptions regarding the teaching practices used in undergraduate courses, how class time would be spent, any differences they expected to see between their high school and university STEM courses, as well as concerns they had about this instructional transition. This presentation will report on whether specific student backgrounds, such as being first-generation, alter first-year student expectations. Previous reports have shown that first-generation students have unique experiences and struggles throughout this instructional transition compared to their continuing-generation counterparts. However, our data suggest that first-generation and continuing-generation students hold similar expectations about the instructional practices in introductory STEM courses. This work will also address how well these student expectations and perceptions align with observations of teaching practices that were collected using the Classroom Observation Protocol for Undergraduate STEM (COPUS). The data from this project will be used to inform Faculty Learning Communities (FLCs) that are focused on the instructional transitions students make during their first year of STEM classes. The ultimate goals of these FLCs are to provide insight into what factors are involved when students, especially those who are first-generation, struggle during the first year of college and to develop classroom materials about this instructional transition that could be used in a variety of courses.

**Comparative Confirmatory Factor Analyses of Measurement Traits between Intended Biology Majors and other First Year College Students** Ryan Dunk\*, Syracuse University; Jason Wiles, Syracuse University [abstract # 159]

Current studies in evolution education have identified a diversity of factors associated with acceptance of evolution among students, including their knowledge of evolution, the strength and personal importance of their religious beliefs, and their understandings of the nature of science. While several of these studies agree on individual factors, there is not yet a consensus on which factors are the most salient to study with an aim toward a general improvement in evolution acceptance. Part of the reason for this may be changing scales: evolution acceptance and associated factors have been studied in biology students, general undergraduates, and the general population. However, it is unclear if the measurements used in one population are even applicable to another population. It may be the case that the same measurement has a different underlying factor structure for differing populations of students. Specifically, in this study we study the factor structure in three different measurements in first-year intended biology majors versus other first-year university students. Participants were drawn from all sections of first-year orientation courses at a large, private, research intensive (R1) university in the northeastern United States. All surveys were completed during the beginning of the students’ first university semester. The surveys analyzed include the Measure of Acceptance of the Theory of Evolution (MATE), the Nature of Scientific Knowledge Survey (NSKS), and three measures (Evolutionary Misconceptions, Evolutionary Knowledge, and Genetic Literacy) from the short form of the Evolutionary Attitudes and Literacy Survey (EALS-SF). Confirmatory Factor Analysis (CFA) models of each survey item were conducted using the lavaan package for R. Each of these CFA models were ran separately for first year students who reported an intended biology or related measure and for all other students. Here, we present the results of these analyses and compare the factor structure of the given measurements between intended biology majors and intended non-biology majors. These results add important nuance to the use of the described measures and inform any study that seeks to compare biology majors and nonmajors in their conceptions of evolution, biology knowledge, and understanding of the nature of science. Further, these results help to inform the generalizability of results of studies that use biology majors and the specificity to biology majors of results of studies that use general undergraduate populations.

**Figuring out our place in science: how are students’ science identities produced and shaped?** Paul Le\*, University of Colorado Denver; Laurel Hartley, Cu Denver; Leanne Doughty, University of Colorado Denver; Amreen Nasim Thompson, University of Colorado Denver [abstract # 249]

Understanding the process of identity production is important in science education because we want to foster spaces where our students can engage with perceptions of who they are and how that intersects with how they see themselves as science people. This is in light of evidence that students’ identities impact their course experiences and willingness to persist in science and NSF data suggesting gender and racial disparities despite our best efforts. Identities are complex because they impact the individual level (How do we construct our sense of self?), interactional level (How can we be “othered”?) and institutional level (How are identities impacted by organizational practices?). In other words, our differences matter, and differences in identities can affect how we view ourselves as science people and our formation of a science identity. We use Holland et al.’s theory of identity and focus on the roles of (1) conceptual identity production (CIP) and (2) procedural identity production (PIP) in understanding how students produce science learner identities. CIP refers to the students’ conception of who they are and who they want to be (e.g., What are their fixed identities? What are their lived experiences and aspirations?). PIP refers to “the performance” of their science learner identities (e.g., How do students engage with science intellectually? Culturally? Socially?) This talk provides an overview of how science learner identities are produced using CIP and PIP. Students in our first-year general biology courses completed Rovai’s Classroom Community Scale, which is a survey that asks students about their perceptions regarding classroom climate and interactions. We combined their responses with demographic data to purposefully recruit a variation of students of different demographics and perceptions to interview. Our final sample included 24 students at one institution. The interview focused on identities, influential experiences, perceptions of science, and classroom community. We used a grounded approach to develop hierarchical coding schemes which focused on describing students’ CIP and PIP. In our coding, we started with 2 umbrella statements (e.g., Students engage in CIP) and created 28 codes that are more specific but fall under those umbrellas. Students’ CIP showed that all interviewees were impacted by past familial and school experiences. These experiences were positive (e.g., siblings who share a love of science) and negative (e.g., parental pressure to pursue medicine) in facilitating students’ science identities. Additionally, students who had clear goals and a plan to achieve those goals presented stronger science identities than those who did not. Interestingly, some students acknowledged systems of oppression and privilege related to their identities, experiences, and overall representation in science. Students’ PIP indicated that those who positively engaged with science intellectually, culturally, and socially were more able to develop stronger science identities. For example, students who reported majority positive interactions with others and become leaders in groups reported stronger science learner identities. Many students struggled to engage culturally because of conflicts in their perceptions of science; though many students believe that science is for everyone, they reiterated traditional science discourses, such as science being for the smartest students. Overall, our findings revealed that students’ conceptions of who they are in science are dynamic. There is not a well-defined pattern in science identity production, and future work should document experiences that occur outside of the classroom. We believe that this contributes to an increasing knowledge base of student science identity and that this study articulates two modes of identity production that allows for thoughtful conversations on how to create spaces that allows for students to produce more positive identities and thrive.

**Gaining Insight into the UC System’s Tenure-Track Lecturer Position** Ashley Harlow\*, University of California, Irvine; Brian Sato, UC Irvine [abstract # 120]

A nationwide push to improve undergraduate STEM education has been motivated by a lack of diversity in STEM programs and the variance in quality in the undergraduate classroom. In response to these concerns, discipline-based education research (DBER) fields have identified interventions aimed at improving STEM education, including active learning lectures and social-psychological interventions. However, implementation of these practices has yet to occur on a larger scale. One potential mechanism to generate institutional change in regard to teaching practices is the use of change agents within a department who possess both discipline-specific and teaching expertise. An example of this is Science Faculty with Education Specialties (SFES), a heterogeneous group of individuals employed in the California State University (CSU) system. Previous research found SFES to be fulfilled by their job, but many were still considering leaving their positions possibly due to a misalignment in job expectations between SFES and CSU stakeholders. A related position is the Lecturer with Potential Security of Employment (L(P)SOE), a tenure-track faculty line within the University of California system that was intended primarily to meet the instructional needs of the university. While a long-standing position, it has recently undergone a transformation to require participation in scholarly activities, such as DBER, and has seen a spike in hires across the UC system over the past decade. As a precursor to an exploration of L(P)SOEs as change agents, the current study aims to answer the following questions: Who are L(P)SOEs and does their training match the new expectations for the position? Are L(P)SOEs satisfied with their roles and do they feel like they belong within their institution? Are there impediments to the successful integration of L(P)SOE faculty within their institution and to their professional success? To answer these questions, we released an online survey to nearly 200 STEM L(P)SOEs across the UC system (and obtained a 75% response rate) as well as a survey of a smaller group of UC stakeholders, including deans, department chairs, and faculty involved with L(P)SOE hiring. The former instrument measured background training, job expectations, job satisfaction, sense of belongingness, and perception of value, while the latter asked similar questions from the departmental perspective. From these surveys, we found that both L(P)SOEs and those involved in L(P)SOE hiring agreed that the majority of their time should be spent teaching, with roughly 20% dedicated towards scholarly activities and another 20% to university service. L(P)SOEs were overwhelmingly trained within their discipline (46.3% had discipline-specific PhDs, and 43.8% discipline-specific postdocs) relative to formal DBER training (7.6% DBER Phds, and 0.8% DBER postdocs). As opposed to the SFES studies, L(P)SOEs were overwhelming satisfied and were not planning to leave their position (3.7% noted they would leave their position in 5 years) and felt a sense of belonging within their department (on average respondents reported values of 4.45 or higher out of 6 on scales measuring belongingness related to collaboration, academic support, and personal support). However, L(P)SOE faculty did cite factors that may cause them to leave, including better opportunities, an unreasonable workload, and disrespect within their department. The above findings were similar across discipline, gender, tenure-status, and UC campus, but we did identify discrepancies that we will elaborate on in our talk. Our results highlight that L(P)SOE faculty were overwhelming positive about their position within their home institution, a key trait that would be expected of individuals who act as change agents. This work is the first step in determining whether the L(P)SOE faculty position is a model that can be used to address the need for institutional change in higher education across the country.

**SATURDAY afternoon**

**Student Reasoning- Saturday afternoon**

**Exploration of Students' Systems Thinking** Narmin Ghalichi\*, University of Minnesota; Anita Schuchardt, University of Minnesota; Gillian Roehrig, University of Minnesota [abstract # 41]

One of the central questions in ecology is how the sum of local ecological processes gives rise to ecological patterns at a larger spatial scale. Systems thinking is important because it allows the consideration of how objects interact within a system. It was theorized that one way of advancing students’ system thinking is to incorporate learning about the biological and chemical processes that connect objects within the system. A curriculum was developed that asked students to investigate the mechanisms behind molecular transformations within the nitrogen cycle. This curricular unit focused on the impact of agricultural practices on ecosystems. It was implemented during a one-semester long elective course on wild-life ecology available to high school seniors. This presentation intended to investigate how the ways in which students connect objects reflect their systems thinking. In order to evaluate the effect of instruction in this unit on systems thinking, students were provided pre- and post concept maps that included molecular and macro level objects from nitrogen and carbon cycles within the context of agroecosystem. Forty-two concept maps (21 pre- and 21 post-) were coded for accuracy of connections, and assessed for the patterns of connections. Based on patterns, a theoretical framework was developed to show how students connect objects within and across nitrogen and carbon cycles. Development of this framework addressed multiple goals. First, it provided a clear definition for broad categories of links that connect objects within the nitrogen and carbon cycles (intra-subsystem) and across the two cycles (inter-subsystem). Second, it identified whether students are linking objects at the macro and molecular levels or only on one level, reflecting students’ ability to move across these two levels. This framework served as a basis to explore patterns in the ways in which students connect matter objects within and across subsystems. Few concept maps demonstrated nitrogen intra-subsystem category in the pre-assessment (9 out of 21), while most students connected objects within the carbon cycle (20 out of 21). This difference probably reflects a relative lack of familiarity with the nitrogen cycle as opposed to the carbon cycle. On the post-assessment, more concept maps (20 out of 21) demonstrated nitrogen intra-subsystem category. The number of concept maps that showed connections within the carbon cycle remained about the same pre- to post-assessment (21 out of 21). The shift to more nitrogen cycle connections in post-assessment concept maps was accompanied by a shift in the number of concept maps making inter-subsystem links across nitrogen and carbon cycles (from 5 pre-instruction to 17 post-instruction). Moreover, there is a pattern between the level of complexity in intra-subsystem links within the nitrogen cycle and ability to link nitrogen and carbon cycles. Sixteen out of seventeen post-instruction concept maps that show links connecting macro and molecular scales within nitrogen cycles also create inter-subsystem links connecting nitrogen and carbon cycles. The research suggests that instruction in a unit that focuses on developing understanding of the mechanism connecting objects in one cycle (the nitrogen cycle) can improve students’ understanding of how objects are connected within that cycle. Concomitantly, students show improved understanding of how two cycles are connected within an ecosystem. Thus, pedagogically, developing mechanistic understanding within a subsystem has the potential to increase systems understanding at a larger scale, an educational principle that is potentially broadly applicable to the undergraduate biology curriculum. Moreover, the theoretical framework on object connections within and across systems is an instrument that could be used by instructors and researchers to evaluate students’ systems thinking at both the high school and college level.

**Developing a learning progression in physiology to characterize how students reason about ion movement** Jennifer Doherty\*, University of Washington; Emily Scott, Univ. Washington; Jack Cerchiara, Univ. Washington; Jenny McFarland, Edmonds Community College; Joyce Parker, MSU; Mary Pat Wenderoth, University of Washington [abstract # 102]

To gain expertise in a field is to recognize, understand, and effectively use underlying disciplinary principles. Too often students rely on rote memorization to solve problems rather than apply the appropriate principles of physics that governs biological phenomena, that is, use principle-based reasoning. Student who rely on memorization can students list the steps of generating an action potential or stomatal opening but cannot reason to a correct prediction when changes are introduced in the system, e.g. when a toxin is applied. A learning progression can help us understand how students’ use of key principles are developed, refined and strengthened over a curriculum. Our goal is to develop a learning progression to describe how students reason about the fundamental physiological principle of flux (movement of a substance is directly proportional to the size of the gradient and inversely proportional to the resistance) to explain ion movement. To develop our learning progression, we used interviews and short-answer questions to solicit explanations and predictions in six contexts: plant tropism, plant transport, neuromuscular, renal, cardiovascular, and respiratory physiology. We interviewed 90 students and collected short-answers from 2500 students ranging from freshman to seniors at both an R1 and Community College. We used a constant comparative approach to uncover emerging patterns of reasoning. We propose the following learning progression: Lowest level, L1: Students provide explanations that tell stories in a non-principled or non-mechanistic way that often contain teleological, “these are the steps” or “these things move like this naturally” ideas. L2: Students use limited mechanistic reasoning. They attempt to explain the “why” or “how” by using relevant course content, often accurately, in a limited way. For example, they use “high to low concentration” or “opposites attract” reasoning where appropriate. L3: Students use emergent principle-based reasoning. In their explanations, students independently address ideas about chemical and electrical gradients, resistance, and ion species, or they make errors when integrating the components. The errors they make are of application, not omission. L4: Students use principle-based reasoning with incomplete consideration of the relationships among interacting components. They construct mechanistic causal explanations that accurately integrate some ideas about gradients, resistance, and multiple ions (i.e., they make errors of omission, not of application). Upper level, L5: Students use principle-based reasoning with full consideration of the relationships among interacting components. They construct mechanistic causal explanations that cohesively integrate ideas about gradients, resistance, and multiple ions. Students reconcile the impact of opposing forces (e.g., electrochemical gradients) or opposing ion fluxes (e.g., Na+ and K+) by attending to key quantitative values (e.g., equilibrium potential). This learning progression will afford faculty insight into how students’ principle-based reasoning develops in the context of ion movement. This insight can inform their physiology course design to better support students’ learning.

**Comparing problem-solving cognitive processes across multiple content areas of genetics** Betsy McIntosh\*, Jennifer Avena, Oscar Whitney, Jenny Knight, University of Colorado, Boulder [abstract # 165]

Although solving complex problems is a desired learning outcome of undergraduate biology courses, students still struggle with this skill. When compared to experts, students tend to use ineffective strategies, such as focusing on surface features rather than broad concepts, and rarely evaluate their approach or check their work. When coupled with a lack of content knowledge, these faulty processes may prevent students from recognizing and correcting their mistakes. To help students navigate these pitfalls, we must first understand their strategies for solving complex problems. We are working to identify which cognitive processes characterize correct versus incorrect answers, and how these processes differ by content area. From undergraduate students enrolled in an introductory genetics course, we have collected 1900 written “think-aloud” responses on four different challenging topics: probability analysis using recombination, probability analysis using a pedigree, determining the mode of inheritance using a gel and pedigree, and evaluating non-disjunction in meiosis. We characterized student cognitive processes into the categories of orientation (e.g. noticing, identifying question type/concept, recall), task execution (e.g. using information, integrating multiple sources, calculating, drawing), reasoning (e.g. reasoning, justification), metacognition (checking work), and drawing conclusions (claim, elimination). In addition to many similarities across content areas, there are several notable differences. “Calculating” is observed in only recombination and probability responses, where it is significantly associated with correct answers. “Elimination”, in which a student narrows down possible answers, is only significantly associated with correct answers in gel/pedigree questions, where students evaluate the nature of inheritance of a particular phenotype, and non-disjunction questions, where students evaluate which phase of meiosis was affected by chromosomal mis-segregation. Statements about creating drawings were significantly associated with correct answers in non-disjunction problems, but with incorrect answers in recombination problems. In the latter case, use of “drawing” largely corresponds to students attempting to solve problems by using a Punnett square, thus failing to consider the chance of recombination between two linked genes. Perhaps most notable is that between 23% and 40% of all student responses from gel/pedigree and recombination problems, respectively, contained justifications with correct logic, but an incorrect final answer. This indicates that students are likely missing a pertinent fact or making an error which then results in an incorrect answer despite understanding the problem-solving process. We plan to further identify patterns or sequences of cognitive processes that correlate with incorrect and correct answers, with the intent of ultimately teaching students how to utilize successful problem-solving pathways.

**Exploring content and process in student problem solving in genetics** Jennifer Avena\*, University of Colorado Boulder; Oscar Whitney, University of Colorado Boulder; Betsy McIntosh, University of Colorado Boulder; Jenny Knight, University of Colorado, Boulder [abstract # 182]

Previous research has shown that students have difficulty solving higher-order problems in many disciplines and that experts commonly employ different techniques than novices. We are studying the cognitive processes students use to both correctly and incorrectly solve genetics problems and testing whether interventions focused on content can help students improve this skill. Students enrolled in an introductory genetics course completed online practice assignments that consisted of multiple sets of questions addressing a specific genetics topic: probability using pedigrees, probability in the case of recombination, modes of inheritance, and chromosome separation errors during meiosis (nondisjunction). Students were asked to provide written documentation of their problem-solving process as they solved three questions on each of four possible content areas. Previously, we determined that a content-focused hint can help some students (36%) who initially answer a question incorrectly to answer a subsequent question correctly. To better understand how students were getting to the correct solution, we further examined students’ problem-solving processes in 132 answers in which students ultimately answered a question correctly after a hint compared to 324 answers in which students continued to answer incorrectly. Compared to students who continued to answer incorrectly despite a hint, students who answered correctly were more likely to have used the content of the hint in their description (54% versus 86%, respectively, p<0.001), made a smaller number of errors and omissions per answer (1.3 versus 0.7, respectively, p<0.001) and were more likely to have used the processes of integrating data and justifying reasons in their answers (p<0.05). We will discuss in detail the mistakes students make that are common to all problems and those that are specific to certain content. We anticipate that this research will assist us in identifying the process of student problem-solving in genetics as well as potential interventions focused on potentially both genetics content and problem-solving processes.

**"How do I solve real biochemical problems?" What analyzing students' domain-specific problem solving reveals about their biochemical ideas**  Paula Lemons\*, University of Georgia; Cheryl Sensibaugh, University of Georgia; Stephanie Halmo, University of Georgia; Sophia Jeong, University of Georgia; Robert Idsardi, University of Georgia; Kush Bhatia, University of Georgia [abstract # 228]

Undergraduate students persistently struggle with biochemistry problem solving because they must integrate and apply their knowledge in both biology and chemistry, which is frequently superficial and flawed. Yet, biochemistry coursework is a required component of many undergraduate life science degree programs and may be a major choice-point for undergraduates as they decide whether to continue in a life science degree program. We are taking a multi-pronged approach to understand how students solve problems in the domain of biochemistry and ultimately to develop and test instructional interventions aimed at improving students' biochemistry problem solving. This type of domain-specific problem solving operates on the assertion that the best problem solvers use a well-organized knowledge base to discern the underlying principles of authentic problems and efficiently apply that knowledge. Using this perspective, we are investigating undergraduates' normative and non-normative ideas about biochemistry concepts in the context of a longitudinal study of 700 STEM students who initially enrolled in introductory biology at a research-intensive institution. At the outset of the study, students completed a biochemistry problem-solving assessment called BioSTEPS, which is comprised of true/false and constructed response items useful for revealing student ideas about macromolecular structure and metabolic pathway dynamics. A subset of students also participated in think-aloud interviews. We used a combination of deductive and inductive coding to analyze student writing and interviews, coding in teams to reduce bias and increase rigor. We identified a number of critical features of student thinking. Regarding macromolecular structure, we found that students: (1) can correctly categorize amino acid R groups without recognizing that this categorization is the key to solving problems; (2) can provide definitions of hydrogen bonds, but often misidentify hydrogen bonds in a macromolecular structure; (3) think that nonpolar groups repel each other, participate in hydrogen bonds, or do not interact at all; (4) express certainty in their predictions when tentativeness is more appropriate given the information available in the problem. Regarding metabolic pathway dynamics, we found that students: (1) think an arrow depicting feedback inhibition indicates recycling of a product; (2) use visual cues to decide if a step in the pathway is reversible; (3) think primarily about isolated steps in a pathway rather than about the impact of that step on the rest of the pathway. Our study is among the first to uncover students' thinking about macromolecular structure and metabolic pathway dynamics in the context of biochemistry problems. Based on our results to date, we recommend instructional interventions that focus student thinking on causal, mechanistic reasoning and underlying principles, which we are testing in ongoing research. We expect this talk to be of interest to SABER attendees who are interested in the integration of conceptual knowledge and scientific practice and those who focus on biochemistry and molecular biology.

**Using Models and Representation- Saturday afternoon**

**Developing tasks that assess student ability to use chemistry ideas to explain biological phenomena** Becky Matz\*, Michigan State University; Alex Kararo, Florida International University; Brittney Pardinas, Florida International University; Laura Santiago, Florida International University; Cait Herr, Michigan State University; Amelia Gotwals, Michigan State University; Kristin Parent, Michigan State University; Sonia Underwood, Florida International University [abstract # 128]

Explaining biological phenomena often requires the application of chemistry ideas, yet connections between introductory chemistry and biology courses are usually not explicit for students. In fact, students often view these courses as isolated experiences resulting in students segmenting the knowledge gained from each of these courses. Further, introductory science courses tend to be broad instead of deep in material covered, typically reflected in a lecture-based pedagogy and assessing rote memorization and factual recall via forced choice exams. Thus, while it is desirable for students to integrate their scientific knowledge across disciplines, students are typically given few purposeful opportunities to do so. In this project, our research question is: In what ways do students use their chemistry knowledge to explain biological phenomena? To answer this question, we are building, testing, and evaluating assessment tasks for introductory biology courses that examine students’ abilities to connect core chemistry ideas with biological phenomena. Our approach to assessment design uses the NRC Framework for K-12 Science Education (Framework) as a foundational guide. The Framework offers a way of thinking about science education aimed at positively affecting students' abilities to both use their knowledge and make connections across disciplines. This vision integrates three dimensions: scientific practices (how students apply their knowledge), disciplinary core ideas (what students really need to know), and crosscutting concepts (themes across science disciplines). The integration of these dimensions is referred to as "three-dimensional learning” (3DL), and we are developing the assessment tasks to incorporate 3DL. Having identified numerous potential areas of connection between chemistry and biology through a semester of course observations at two universities, we surveyed our local introductory cell and molecular biology instructors to determine which areas of connection they most valued. Based on these survey results, and drawing on principles of evidence-centered design, we drafted assessment tasks that ask students to apply core chemistry ideas to biological phenomena as well as use scientific practices. For example, one task centers on students applying their understanding of probability and entropy to explain osmotic balance across cell membranes, and another focuses on how intermolecular forces and DNA structure can be used to explain the relationship between temperature, hydrogen bonding, and DNA stability. These items were administered through an online platform to approximately 600 students across two universities, including students who had taken both relevant chemistry and biology courses, and, as a control group, students who had only taken chemistry. In this presentation, we will report on findings from the faculty survey and administration of the assessment tasks to students. Biology education reports have underscored a fundamental need for scientists and nonscientists alike to be able to connect seemingly disparate information with biology for the purpose of addressing large-scale challenges. Indeed, the Vision and Change Pathways and Transformations of Energy and Matter core concept specifically cites that biological systems function precisely because of the ways that chemical pathways are governed by thermodynamics, and even undergraduate students themselves recognize the need for more connections between chemistry and biology. Yet, assessments that cross this disciplinary boundary, testing student understanding of the chemical basis of biological systems, are few and far between at the college level. The assessments developed and tested in this project are designed to meet this need. Through additionally incorporating the principles of 3DL, the tasks will help the research community understand how students use their chemistry knowledge to explain biological phenomena.

**3-Dimensional models reduce biochemistry learning gaps** Michelle Howell\*, University of Nebraska - Lincoln; Brian Couch, University of Nebraska-Lincoln [abstract # 96]

Understanding the relationship between macromolecular structure and function represents a central goal of undergraduate biology education. AAAS, ASBMB, and the NRC all emphasize that students need to master this fundamental goal. However, how instructors teach visual literacy is not organized, leading to the existing educational problem that many students struggle to make strong mental models of complex 3-dimensional (3D) structures from 2-dimensional (2D) images. Thus, the Biomolecular Visualization theoretical framework developed by Dries et al outlined core content and competencies to guide instruction in visual literacy and in the ability for students to go back and forth between the 3D model and the 2D image. 3D models have been suggested as one of the most useful teaching tools in science education and long-term learning because they can stimulate students to ask more sophisticated questions, while more effectively capturing student interest. This contrast from 2D illustrations provided our rationale to specifically focus on 8 out of 12 of Biomolecular Visualization’s overarching themes to ask the following research question: Do 3D model-based instructional modules improve student performance on structure-function assessments? Our research involved three stages: 1) development, 2) implementation, and 3) assessment. First, we designed 3-D printed molecular models and developed three easily-incorporated instructional modules for any size third-year undergraduate biochemistry class. Each module was developed using backward design from a list of learning objectives and aligned assessments. The entire program was implemented and assessed using pre-post assessment instruments at two universities to demonstrate the impact on: 1) student understanding of structure-function relationships, 2) student learning of fundamental biochemistry concepts, and 3) student enthusiasm for the instructional method. The class was divided into groups of four where each group received one set of models. Students took roughly 45 minutes to complete the in-class, inquiry-based activity. Each student submitted individual answers via a Qualtrics survey with adaptive responses built in to promote learning. Results from the three 3D model-based instructional interventions indicate improved student performance on the material. Data is reported as normalized learning gain [(pre score – post score)/(1 – pre score)\*100]. The first module comparing DNA and RNA structure and function resulted in a 48% learning gain (N=114). Five specific learning goals for this module were improved with 50-81% gains (with an average of 66% gains). The second module, on concepts related to DNA supercoiling, produced a normalized learning gain of 45% (N=114). Nine learning goals for the DNA supercoiling module were improved by 29-76% (with an average of 49% gains). The third topic illustrated protein secondary structure-function relationships and resulted in student learning gains of 22% (N=15). Six specific learning goals for this module were improved with 13-58% gains (with an average of 31% gains). Additional survey responses followed by focus group interviews revealed that nearly 70% of the students found that the physical models made it easier to learn the material being taught. Thus, our evidence suggests that these 3D model-based instructional modules positively impact student understanding of macromolecular structure-function concepts. These modules can help STEM students develop mental frameworks that more accurately represent structure-function interactions. Moreover, this work supports the use of 3D models as effective learning tools that promote a deeper understanding of fundamental biochemistry concepts and better equip students to easily switch between 2D and 3D models.

**Coupling experimental design and computational modeling to support productive engagement with scientific uncertainty** Julia Gouvea\*, Tufts University; Aditi Wagh, Tufts University [abstract # 100]

Learning how to do science means learning how to make progress in the face of uncertainties. Scholars in science education have argued that students need practice navigating scientific uncertainties: to learn how to identify and characterize new problems, plan and design experiments, interpret complex datasets, and make and defend theoretical claims (Ford, 2005; Manz, 2015; Metz, 2004; Phillips et al., 2017; Reiser, 2004). A major challenge for biology education researchers is to understand how to design learning environments that balance allowing students to grapple with uncertainty with providing the support needed to make progress. In this talk we report on findings from a design-based research study (Cobb et al. 2003) of an inquiry-based introductory biology laboratory curriculum. A core design feature of this curriculum is that students conduct extended investigations into biological study systems using both experimental and computational methods. At multiple points throughout a lab unit they are supported in making connections between the two methods. The primary contribution of our work is that we propose theoretical relationships between designed features of the lab and observable student activity. Our central claim is that coupling experimental design and computer simulations can both generate scientific uncertainty and support students in making progress in their scientific investigations. We support this claim with a case study of how a group of students made progress during a three-week inquiry-based investigation on variable mutation rates in bacteria. In this unit, we asked students to investigate, using both experiments and computer simulations, the relative benefit to a population of bacteria mutating at a higher or lower rate. Our qualitative analysis of this focal group triangulates across video of students’ activity in lab, interviews with 3 of the 4 focal students and three writing assignments collected from all four students. We present our findings in two parts. In the first we describe the uncertainties the group encountered and the scientific progress they made. We describe, for example, how the group mapped their theoretical understanding of the study system, locating a problem in need of study in the process. We also describe how the group was able to make and justify choices about what to investigate both in designing an experiment and investigating the computer simulation. In the second part, we describe how the group’s progress was linked to the design feature of coupling experiments and simulations. We found that (1) students located a scientific research question by considering discrepancies between the simulation and experimental system; (2) students used the simulation to reason through the predicted results of their experiment; (3) students used their experimental results to question an initial prediction from simulation and began exploring the parameter space more thoroughly; (4) when the results of the experiment were difficult to interpret, the group used the simulation to continue thinking about their expectations and future experiments. Each method was both a source of uncertainty and a tool to support progress at different points in the investigation. While our analysis is limited in that we cannot generalize the activity of this group of students to all students, the strength of our analysis is that we propose a novel theoretical links that the community can consider and test. In this way we add to recent work that proposes to model the links between design and student outcomes (e.g. Corwin, 2015). Further, this work has practical implications for educators interested in designing open-ended inquiry-based labs that create opportunities for students to productively engage with scientific uncertainty.

**Understanding student reasoning between modeling tasks during a computer-based lesson** Gretchen King\*, University of Nebraska-Lincoln; Heather Bergan-Roller, Northern Illinois University; Nicholas Galt, Valley City State University; Tomas Helikar, University of Nebraska-Lincoln; Joe Dauer, University of Nebraska-Lincoln [abstract # 139]

As instructors utilize biological computer models to enhance student learning, it is important to understand how to best design modeling activities to engage students in challenging and diverse types of reasoning. The modeling process provides an opportunity for students to think deeply about the system through inductive reasoning, explanations, and evaluation of their thinking. Cognitive reasoning processes, such as inductive reasoning and explaining, identify how students are thinking about content, while the focus of their discussion shows what computational model features they are thinking about. This study employed two research questions: 1) How do students’ cognitive reasoning processes and discussion foci differ between students who build their own computational model (Build) and students who investigate a pre-fabricated model (Investigate)? 2) In what ways does each modeling activity phase (Construct (Build only), Revise (Build only), Simulate (Build and Investigate)) affect students’ reasoning processes and discussion foci? Ten groups (Build=7; Investigate=3, mostly quads) of undergraduate students were audio recorded in an introductory biology lab while completing a lesson on the lac operon system in the web-based modeling platform, Cell Collective Learn. Qualitative content analysis was used to categorize student cognitive reasoning and discussion foci. Reasoning processes (n=5,426) were coded (in increasing order of sophistication) as: read and paraphrase, analyze, inductive reasoning, explain, and evaluate. Discussion foci were categorized as task based (assignment or tool) or model based (component, relationship, mechanism, modeling actions, or model output). Two main findings emerged from this study: 1) Constructing and revising a model fosters deeper student reasoning and 2) Model revision prompts students to evaluate their own model. Students who constructed their own computational model (Build) used more in-depth reasoning than students investigating a pre-fabricated model (Investigate) by employing more sophisticated reasoning processes like explain (B=6%; I=4%) and evaluate (B=1%; I=0%). Investigate students read and paraphrased more (34%) than Build students (26%), suggesting that Build students became more familiar with the content during the Construct and Revise phases and did not rely as frequently on lesson materials. Some Build students employed metaphors to explain the system, a phenomenon not observed in Investigate groups. For example, a Build student compared the effect of the presence of glucose to an idling car, indicating they were fluent enough with system dynamics to generalize mechanisms of the lac operon system to other familiar systems. Build students also discussed system mechanisms more frequently (B=12%, I=9%), while Investigate students placed more emphasis on the components (B=16%, I=18%) and relationships (B=22%, I=25%) in the model. While the Construct and Revise phases required an average of 45 minutes longer, created a higher cognitive load, and fostered a higher percentage of off-topic discussions (B=30%, I=16%), there were substantive benefits for the Build students, including more sophisticated reasoning processes and focus on system mechanisms. Additionally, the Revise phase prompted Build students to evaluate their models and think more critically about the actions they performed while revising their model. During the Revise phase, groups focused on the model output (17%) and the ways in which they changed their model (19%) while evaluating their model quality. This attention to model output was not observed during any other lesson phases by either treatment group. Taken together, the findings emphasize that building and revising a computational model is beneficial to student reasoning, despite the higher cognitive load. Future research should examine how these phases of modeling prompt deeper reasoning to occur.

**Visualizing cognitive load: The effect of visual representation on student performance**  JESSIE ARNESON\*, Washington State University; Erika Offerdahl, Washington State University [abstract # 155]

A learner’s capacity for processing information is exceptionally limited because the working memory can accommodate only a few elements at a time. Through repeated practice over time, experts develop sophisticated mental schema that link multiple elements together, thereby functionally increasing the amount of information that can be processed by the working memory. As the transition from novice to expert is characterized by the development of such schema, instructional materials should be designed to support novices’ schema development without overwhelming working memory. To this end, cognitive theory of multimedia learning suggests using visual representations to reduce the strain on working memory and support schema development by allowing information processing to occur through both visual and verbal channels. The use of visual representations (e.g. graphs, diagrams, cartoons) is an integral part of science communication, and students frequently interact with visual representations in undergraduate science courses. Multimedia learning theory predicts increased student performance when information is presented visually and verbally. Conversely, cognitive load theory predicts that introducing visual representations to a task will increase cognitive demand, require additional processing to make sense of the representation, and result in lower performance. To test these hypotheses in the context of undergraduate science, we crafted paired questions in which one form conveys some of the information through a visual representation, while the other form strictly uses text. These questions were embedded in exams featuring biochemistry content in courses that spanned a molecular life sciences curriculum. Preliminary student performance data reveal mixed support for the hypotheses. For a subset of the questions, students were more successful at answering the form that contained a visual component. On other questions, students performed significantly lower on the visual form as compared to the strictly text form. However, many factors contribute to the cognitive demand of a task, including the Bloom’s level of the question, the type of visual representation used, the format of the question, and student familiarity with the provided representation. Further analysis of how these factors may influence students’ ability to answer these exam questions will better elucidate the ways and contexts in which students struggle to use visual representations. Better understanding of the impact of visual representations will allow us as educators to devise strategies for training students to interpret and use scientific images. Improving student visual literacy could lead to increased understanding of the content in STEM courses and greater retention of students in STEM disciplines.

**Classroom Practices- Saturday afternoon**

**Incorporating Mathematics into an Introductory Biology Course: Examining the Impact on Life Science Majors’ Attitudes Toward Mathematics** Melissa Aikens\*, University of New Hampshire; Sarah Andrews, University of New Hampshire [abstract # 82]

Quantitative skills are an integral component of biology, yet incorporating quantitative skills into biology courses can be challenging for a number of reasons. One particular challenge that instructors face is negative attitudes toward mathematics from some students. According to expectancy-value theory, these negative attitudes can impede student engagement in quantitative tasks, resulting in lower performance. Therefore, quantitative skills should be taught in a way that aims to engender more positive attitudes toward mathematics. However, the extent to which students’ attitudes can improve over the course of one semester is unclear. Moreover, studies in physics have demonstrated that students can actually have worse attitudes toward mathematics after completing a physics course. The goal of this study was to determine how life science majors’ attitudes toward using mathematics in a biology context change over one semester of an introductory biology course that integrated quantitative problem sets. Here, we focused on quantifying changes in students’ (1) interest in using mathematics to understand biology, (2) perceptions of the usefulness of mathematics for their life science career (utility value), and (3) perceptions of the cost (extra effort or worry) of incorporating mathematics into biology courses, which are task-values predicted by expectancy-value theory to be important for motivation and academic achievement. Five problem sets were given to the students over the course of the semester in lecture that involved testing a hypothesis by using a mathematical model and interpreting the output of the model. We measured interest, utility value, and cost in life science majors (n=103) at the beginning and end of the semester using the Math-Biology Values Instrument (MBVI), which contains items on a 1-7 scale. To better understand whether the quantitative problem sets were responsible for any observed changes, we asked students whether biology lecture, lab, and/or another experience during the semester influenced their responses to the MBVI on the post-survey. We also included non-equivalent dependent variables, which measured students’ attitudes toward physics in the context of biology, on the pre- and post-survey. Similar changes in both the MBVI scores and the non-equivalent dependent variables would suggest that general maturation, rather than any specific quantitative curriculum component, is responsible. We used linear mixed-models to account for repeated measures and included demographic and academic variables as controls. By the end of the semester, students were significantly more interested in using mathematics to understand biology ( =0.37, p=0.001, pre/post lsmeans: 4.00/4.38) and perceived significantly less cost associated with mathematics in biology courses ( =-0.31, p=0.02, pre/post lsmeans: 4.80/4.49). Utility value did not significantly differ between the pre- and post-survey, likely because it was high initially (pre/post lsmeans: 6.03/6.14). We did not see similar increases in interest or decreases in cost in the non-equivalent dependent variables, suggesting the observed changes in students’ math-biology attitudes are not related to maturation processes. Furthermore, 77% of students reported that lecture influenced their responses to the MBVI on the post-survey. However, 35% of students reported lab influenced their responses, and 15% reported that another course influenced their responses. Therefore, although the problem sets likely had some impact, they were not the only factor influencing the observed change. Our results suggest that incorporating quantitative skills into an introductory biology course can lead to more positive student attitudes toward mathematics, though in this case, the changes were small. However, it remains unknown what is driving this change, which opens up intriguing possibilities for future research that could lead to improved instructional design of quantitative curricula.

**Successful integration of statistics and programming in undergraduate biology courses using Cognitive Load Theory.**  Laura Guzman\*, University of British Columbia [abstract # 92]

Biostatistics courses are integral to many undergraduate biology programs. Biostatistics courses have often been taught using point-and-click software; now that biology is becoming more data and computationally intensive, these software programs are seldomly used by researchers or professionals. Therefore many biology programs are not equipping their students with the skills they will actually use as they move forward in their careers. A solution to this problem is to teach biostatistics students the programming language R which is the most commonly used software for statistics and data science. However, teaching statistics and programming together can overload the students and hinder their overall learning. Here we used cognitive load theory to develop assignments for students in two biostatistics classes (introductory and advanced). The assignments contained completed problems, partially completed problems, and code explained using the split-attention effect. To evaluate the effectiveness of the assignments we surveyed two cohorts, before and after the assignments were introduced. We also analyzed the students’ lab reports, midterms, and final exams to evaluate how well they were able to apply both the programming and statistical concepts they were taught. We found that students doing the assignments developed using cognitive load theory rated their programming ability higher and were more likely to put the usage of R as a skill in their CVs than students who just received a pre-completed R script. We also found that students were more motivated, less frustrated and less stressed when learning programming using the assignments. From the students’ answers to open ended survey questions, it was clear that they liked the assignments, and specifically, the explanations, step-by-step problems, and the examples. Overall, this work suggests that we can use cognitive load theory as a foundation to develop and teach high-interactivity-material such as statistics and programming to better prepare students for modern careers in biology.

**Learning Spaces Matter! The TEAL classroom increases the benefits of flipped classroom pedagogies.**  Marcos García-Ojeda\*, University of California, Merced; James Zimmerman, University of California, Merced [abstract # 47]

Multiple governmental and educational organizations have advocated for the adoption of student-centered active learning strategies (ALS) throughout the undergraduate STEM curriculum. These ALS increase student performance, enhance student mastery of class content and augment student affect, particularly in underrepresented minority students. Among these student-centered ALS, the flipped classroom pedagogy delivers class content to students outside the classroom through pre-recorded video lectures and utilizes in-class time for ALS, including collaborative learning. Course delivery in Technology-Enabled Active Learning (TEAL) classrooms also facilitates implementation of ALS by encouraging student engagement in team-based, collaborative learning. However, the effectiveness of flipped classroom pedagogies has not been evaluated in courses taught in TEAL classrooms. This study examined the impact of TEAL learning space in the effectiveness of flipped classroom pedagogies. We assessed student performance (exam, quiz, and participation) in a completely flipped, upper division microbiology course, delivered either in a traditional classroom or a TEAL classroom at a doctorate-granting university. The same instructor delivered both classes flipped, using the same video lectures and quizzes, while the exams had minor alterations (e.g. different temperature ranges or bacterial species). Also, the active learning activities and worksheets used were the same in both courses. The principal difference was the learning space where the courses were delivered: either a traditional classroom or a TEAL classroom. Our data indicates that final course grades were higher, with fewer students that failed, in the flipped class taught in the TEAL classroom as compared to the same class taught in the traditional classroom. There was a significant difference (F(1,281)=8.2, p=.004) on course grade, with those in the traditional classroom earning a lower course grade (MTrad=76.7, n=108) than those in the TEAL classrooms (MTEAL=80.2, n=175). The percentage of students earning A’s in the class was equal in both courses, but there was a trend towards more B's and fewer C's in the TEAL-taught class. Chi-square analysis showed that significantly fewer students failed the course (earned a D or F) in the TEAL classroom compared to the traditional classroom (𝒳2 = 5.439, p=.0197). To determine what component caused this shift in grades, lecture participation, discussion participation, exam scores and quiz scores were examined for significant differences between the two classroom settings. Lecture participation scores and exam scores were similar between the two cohorts. However, there was a significant difference on discussion participation (F(1,281)=4.4, p=.037), with those in traditional classrooms earning a slightly higher discussion participation grade (MTrad=73.0, MTEAL=71.8). The strongest contributor was revealed when quiz scores were examined. On average, the overall quiz scores were about 20 points higher in the TEAL-taught class than in the class taught in the traditional classroom, a statistically significant increase in performance (F(1, 281)=136, p<.001). Also, students scored high on Positive forms of motivation and Lower in Negative forms of motivation after completing the Self-Determination Motivational Assessment tool, which identified motivational modes active in the TEAL classroom. Taken together, our data indicates that when most important variables stay constant (instructor, pedagogy, quizzes, and exams), students are more motivated and perform better in a flipped course delivered in a TEAL classroom versus the same class delivered in a traditional room. This study indicates that although active learning pedagogies alone improve student performance, they are even more effective when delivered in TEAL classrooms, a finding which has important policy implications as universities look to redesign their biology classrooms.

**Implementation of open textbooks in two-year college biology courses: the Good, the Bad, and the Data** Kristyn VanderWaal MIlls\*, Saint Paul College; Kimberly VanderWaal, University of Minnesota Department of Veterinary Population Medicine [abstract # 46]

One challenge facing students today is high textbook costs. This problem is particularly difficult at two-year colleges where students typically have lower incomes and textbooks are a larger part of the overall cost of education. In order to ameliorate this expense, many advocate using open source textbooks, which are free online and low-cost for print. However, some faculty have raised concerns about the quality and efficacy of these open texts. In our study, we investigate these concerns by comparing student performance, utilization methods, and textbook perceptions in courses where traditional published or open textbooks are implemented. Specifically, we collected data from 33 general biology classes at four different two-year colleges on the effectiveness of an open textbook compared to traditional textbooks.  In addition, we surveyed biology students and faculty on their use and perceptions of open and published textbooks. In all univariable statistical models, book format (open vs. published) did not have a significant effect on measured student outcomes (p>0.2).  However, in both uni- and multi-variable models, semester and class size did significantly affect student outcomes.  Additionally, survey results suggest that student textbook use does not always align with faculty expectations.  For example, 28% of students report actually reading their text compared to 75-95% of faculty expecting that students read the text.  Finally, faculty who implement open textbooks have different opinions about how textbooks should be utilized as compared to faculty who use traditional textbooks. This research advances our understanding of resource utilization by biology students and faculty, the efficacy of different types of textbooks, and decision making processes on the adoption of open textbooks.

**The Intermediate Constraint Hypothesis: Using constraint as an axis for promoting active learning**  Eli Meir\*, SimBio; Denise Pope, Center for the Integration of Research, Teaching and Learning; Daniel Wendell, Massachusetts Institute of Technology; Kerry Kim, SimBio; Ling Hsiao, Massachusetts Institute of Technology [abstract # 252]

Providing active learning experiences for students is hard in large classes in part because keeping students challenged at the appropriate level, and providing timely feedback, are both difficult with high student-teacher ratios. While virtual learning environments can help, the definition of inquiry in such environments covers a wide range. Some virtual activities follow a pre-set path with highly constrained student decisions and multiple choice-style questions. Others offer very open-ended environments where students pose and answer their own questions under minimal constraints. In general, constraints in an activity allow feedback more specific to a student’s work and answers, and provide more guidance, at the cost of less room for student-directed exploration. What level of constraint optimizes student learning? Here we synthesize data from several studies involving over 7,000 students at over 45 high schools and colleges using five different simulation-based ecology and evolution activities. We examine how level of constraint, along with the associated ability to provide instant and specific feedback, impacts how students interact with inquiry-based educational software. Our data comes from several question types embedded in simulation-based labs including multiple-choice (highly constrained), LabLibs and WordBytes (essay-like question formats that have intermediate levels of constraint), and short essay (open-ended). We examined usage patterns for 17 questions in the two intermediate constraint formats compared to 4 short essay and 81 multiple choice questions. Separately, we also examined student work in a simulation exercise within a lab called “Understanding Experimental Design” where we varied levels of constraint. We use this data to investigate how student’s higher-order thinking and learning are affected by constraint and feedback. Across the question types, we find that intermediate levels of constraint often lead to more information on student thinking and more engagement, as measured by the range of answers students give, the number of attempts students make to reach a correct answer, the ability of these questions to predict quiz scores, and the perceived difficulty. With both intermediate constraint questions and intermediate constraint levels in the simulation activities, we see evidence that students learn the concepts being taught, that this learning is efficient, and that changes in understanding may be higher than with high or low constraint formats. From these data, we propose an “intermediate constraint hypothesis” wherein focusing on level of constraint in student exercises may be a good metric for helping target active learning experiences to maximize student learning. Perhaps unsurprisingly, intermediate levels of constraint (and associated feedback) may often promote the greatest learning, but each student may benefit most from a different level of constraint. Thinking in terms of constraint may be a practical path towards “tuning” student exercises to keep each student within their zone of proximal development.

**SUNDAY morning**

**Naure of Science- Sunday morning**

**The development and use of a universal rubric for assessing scientific arguments** Shannon Butler, University California, San Diego; Lisa McDonnell, University of California San Diego; Stanley Lo\*, University of California San Diego; Ella Tour, UCSD [abstract # 207]

Scientific arguments use data as evidence to make claims and provide justification that links the evidence to the claim. They are used to make sense of data, communicate, and construct hypotheses. Although scientific argumentation is aligned with Vision and Change, these skills are rarely taught explicitly to undergraduates. A recent (2010) survey of college educators found that a common barrier to teaching this skill is the lack of assessment tools. Scientific argumentation has been extensively studied, especially in the K-12 context. However, we lack a universal framework to assess scientific arguments. Many researchers use self-constructed frameworks or the Toulmin 1958 argumentation framework, which delineates the components of arguments but does not assess the depth or correctness of scientific arguments. The goal of our study is two-fold. First, we wanted to design a universal rubric for scientific arguments. Second, we wanted to characterize the argumentation skills of upper-division biology students so that it can inform the design of future interventions. We developed a universal rubric to assess scientific arguments by combining the Toulmin framework and Biggs’ Structure of the Observed Learning Outcome (SOLO). Biggs’ SOLO is broken down into five levels of increasing complexity, and these levels can be applied to connect the three main components of the Toulmin framework: claims, evidence, and justification. The first SOLO level is prestructural: These argument tends to miss the purpose of the question due to data being misinterpreted, or irrelevant information is used. Unistructural arguments use a single feature of the data to make a claim, often resulting in incorrect conclusions if students neglect other pieces of the data that do not fit their claim. Multistructural arguments use multiple pieces of data but do not necessarily integrate them into a coherent claim. Relational arguments also use multiple pieces of data, but unlike in multistructural arguments, these data are integrated to make a claim that is consistent with all of the information. Extended-abstract arguments integrate multiple pieces of data into a claim and also relate this claim to relevant outside information. These arguments often critically evaluate the data and generate testable hypotheses and predictions to further support or refute the claims. We developed this rubric through iterations of testing and refinement on arguments written by students, using the rubric to determine the highest SOLO level achieved in each student’s argument. With the final form of the rubric, we achieved high reliability between two coders (86-95% agreement for all codes). We characterized the argumentation abilities of a stratified sample of upper-division genetics students (a total of 120 students, 60 from 2016 and 60 from 2017), selecting 20 students from each year who received an A (90-100%), B (80-89%) or C (70-79%) in the course, with two writing samples per student (240 total arguments). While nearly half of the arguments (48%) reached the multistructural level, only 28% of the arguments were relational, and less than 2% of the arguments were extended abstract. This indicates that while students were able to describe multiple pieces of data and make individual claims (multistructural), they struggled to relate multiple pieces of data together into a claim that was consistent with all data (relational). Hardly any students were able to make forward-looking hypotheses and predictions that would potentially explain the observed data (extended abstract). In our presentation, we will discuss these findings and implications of this rubric in terms of using it as a scaffold to develop targeted interventions and as an assessment tool for scientific argumentation.

**Development and Testing of Competencies for Experimentation in Biology** Stephanie Gardner\*, Purdue University; Nancy Pelaez, Purdue University; Dina Newman, Rochester Institute of Technology; Joel Abraham, CSU-Fullerton; Christopher Beck, Emory University; Trevor Anderson, Purdue; Yue Yin, University of Illinois at Chicago [abstract # 242]

Reforms to biology education at both the K-12 and undergraduate levels call for engaging students in the process of scientific inquiry to think creatively about real world issues that have a biological component, communicate these thoughts to others, and integrate these ideas to make decisions of benefit to science and society. Moreover, many contemporary world issues are directly related to the life sciences, so across the biology subdisciplines, educators are focusing on problems that can be addressed with biological experimentation. Although increasing numbers of biology undergraduates are expected to develop experimentation abilities, surprisingly little is known about what students actually learn from participating in biology research experiences, compared to what they ought to learn to gain a basic competence in biology experimentation. The Advancing Competence for Experimentation in Biology (ACE-Bio) Network was formed with funding from the National Science Foundation to tap the expertise of a cross-disciplinary group of faculty with the overarching goal of developing a set of competencies, assessments and recommendations for undergraduate education in the area of biological experimentation. The network brings together education specialists and research scientists representing a broad range of biology subfields with 46 active participants, some of whom will present examples of research addressing the following Research Question: How are we going to help undergraduate students along their path to learn the process of scientific inquiry by thinking creatively about experimentation in biology? Collaboratively, a core group of scientists and educators developed a set of Basic Competencies of Biological Experimentation about the concepts underlying an experiment and the nature of representations applied in reasoning about experiments and visualizing data. These are published as a document that is easy to modify at http://docs.lib.purdue.edu/pibergiim/4. This presentation will integrate multiple smaller talks about what students should know about experimental research and be able to do according to a set of competencies specific to the area of experimentation in biology. Specifically we will present: (1) the literature base, as well as the research design and methods used to define these areas of competence for experimentation in biology; (2) validation of the competencies according to data from interviews with life scientists in various disciplines; (3) validation of the competencies based on published student difficulties with biological experimentation, (4) results from new instructional and mentoring approaches some have adopted targeting the different aspects of experimentation that undergraduate students should practice, and (5) holes in the literature and places where work is still needed to address the gaps. As a key result, one finding from this work is that breaking the areas of competence for experimentation into component pieces may target related achievements not focused on experimentation. Another key result has raised questions about how well different parts of the competencies translate from one discipline to another and how to adapt what exists to a different discipline or context or a different student demographic. The speakers come from a variety of biology sub-disciplines, are at different career levels, and come from a wide range of institutions. By providing a framework of the competencies, the project aims to support faculty from liberal arts colleges, masters-degree-granting institutions, and major research institutions, working together, to provide input to the development of assessments that can positively impact their students. As a key outcome, the common list of competencies and shared assessments are being used across multiple institutions and disciplines of biology engaged with examples of experimentation that will be presented to illustrate these issues.

**Student Thinking- Sunday morning**

**If I can make it in O-Chem, I make it anywhere! Factors affecting mindset development** Lisa Limeri\*, University of Georgia; Hannah Harper, University of Georgia; Jun Choe, University of Georgia; Erin Dolan, University of Georgia [abstract # 21]

How students think about intelligence affects their academic performance, but what determines these beliefs? Implicit theories of intelligence, commonly called mindsets, refer to whether a person believes that intelligence is an innate, unchangeable trait (fixed mindset), or a malleable trait that can be improved with work (growth mindset). Extensive research has demonstrated that mindsets substantially affect students' academic achievement and persistence. Most research on the origin and development of mindsets has highlighted the influence of messages young students hear from parents and teachers and the effects of interventions designed to shift students' mindsets. However, more recent studies tracking college students in introductory STEM courses have shown that students' mindsets can change over a semester without intervention. The idea that college students are still evaluating and modifying their mindsets opens new questions about factors driving mindset change. The purpose of this study is to identify factors influencing development of mindsets throughout a challenging, upper-level course. Our overarching hypothesis is that academic struggle or failure will shape students' mindsets. Here we present the results of the qualitative component of a larger mixed-methods study. We surveyed and interviewed students in organic chemistry II at a large public research university (n=410) about their mindsets and course experiences throughout the semester. We analyzed responses from a phenomenological perspective using both a deductive and inductive coding process. We drew on Bandura's sources of self-efficacy framework as a starting point for our analyses. Participants reported certain factors that influenced their mindsets that related to Bandura's framework, such as previous personal experiences (mastery experiences) or direct observations of peers' experiences (vicarious experiences). Many of these experiences were related to academic performance, supporting our overarching hypothesis that academic struggle will shape students' mindsets. Participants also cited factors that were unrelated to Bandura's framework, such as using logic and reasoning to conclude that intelligence must be malleable or fixed. For example, some reasoned that, because people are not born with knowledge but acquire it throughout their lives, intelligence must be malleable. We conclude that Bandura's framework serves as a useful starting point but is insufficient to describe the development of mindsets in our sample. This study will be among the first to examine the factors that influence undergraduate biology students' mindsets. These results could be leveraged to maximize the effectiveness of mindset interventions, ultimately improving student persistence and success.

**Student performance and ability increases following a novel neurophysiology simulation** Jack Cerchiara\*, University of Washington; Emily Scott, Univ. Washington; Kerry Kim, SimBiotic Software; Eli Meir, SimBiotic Software; Mary Pat Wenderoth, University of Washington; Jennifer Doherty, University of Washington [abstract # 190]

Undergraduate physiology students find learning neurophysiology challenging. Students in introductory courses understand ion concentration gradients but struggle to simultaneously account for electrical charge gradients. This concept is further complicated as ion flux and electrical potential both change temporally and spatially. There is some evidence that laboratory simulations help students better understand complex physiology topics. We sought to better understand these learning challenges and assess the impact of a new simulation tool, the Action Potentials Extended tutorial by SimBio. We created a pre- and post-intervention assessment comprised of 17 multiple-choice items, which we gave to students in an introductory plant and animal physiology course (n=496). We administered all pre-tests prior to instruction, while post-tests were administered immediately following laboratory sections using the tutorial. We determined student ability by calculating weighted likelihood estimates (WLE, Θ) through item response theory (IRT) analysis (reliability: WLE = 0.730; EAP = 0.762; n=496). We investigated the internal validity of the assessment by analyzing item characteristic curves of score probability and student ability for each question, as well as a Wright Map of student scores and ability. We used linear mixed-effect regression to test for the impact of the tutorial on student performance and ability and learning gains. Our assessment discriminated students’ with average ability well (WLE -2 to 1.5 Θ), however, did not discriminate well students of the highest ability (WLE > 2 Θ). This suggests some students had more sophisticated knowledge than could be measured by the current assessment. We found significant improvements in both student raw scores (model r2 = 0.5064, p<0.001, n=447) and student ability (model r2= 0.4747, p<0.001, n=447) after the tutorial and course instruction. Student scores, as well as ability, were correlated with cumulative GPA (p<0.001, n=447). We also found that males scored higher (p=0.014) and had higher ability scores (p=0.003) on both pre- and post-assessments. Caucasian students of both genders were positively correlated with score (p<0.001) and ability (p<0.001). The number of lectures a student had the opportunity to attend between pre- and post- assessment was also positively correlated with both score (p=0.011) and ability (p=0.009). Our results contribute to the field by demonstrating the efficacy of a new simulation tutorial in assisting undergraduates in understanding challenging neurophysiology concepts. We also developed, aligned and validated a new assessment for testing not only the Action Potentials Extended tutorial, but also student understanding and learning of electrophysiological principles.

**Assessemnt- Sunday morning**

**Can Mixed Assessment Methods Make Biology Classes More Equitable?** Sehoya Cotner\*, University of Minnesota; Cissy Ballen, University of Minnesota [abstract # 20]

In many countries, women who enter college in any of the STEM (science, technology, engineering and mathematics) disciplines exhibit greater attrition than do their male peers, a gap that continues throughout most STEM professions. Some explanations for this phenomenon relate to student preparation or academic abilities, which is collectively known as the student deficit model. We have proposed the course deficit model, whereby instructional decisions exacerbate or minimize gaps in performance and retention. We offer new evidence in support of the course deficit model, broadening our focus to include first-generation college students in introductory biology, chemistry, math, and physics courses. This presentation combines findings from three recent publications as well as emerging work on a large (N > 5,000) sample of students across STEM at one large research institution in the United States. Among the factors proposed to explain attrition in science is lower performance in introductory courses. Thus, we focus on the "introductory course problem," whereby these gateway courses effectively "weed out" students-disproportionately women, minorities, and first-generation college students. First, we demonstrated that for female students only, test anxiety negatively affected exam performance. We hypothesized that mixed assessment methods, whereby we evaluate students on a number of low-stakes assessments rather than a few high-stakes exams, would reduce gender gaps in performance. Specifically, we analyzed gender-based performance trends in nine large (N ~ 1000 students) introductory biology courses. Females underperformed on exams compared to their male counterparts, a difference that does not exist with other methods of assessment. Next, we analyzed three case studies of courses that shifted grading schemes to either de-emphasize or emphasize exams as a proportion of total course grade. We found that the shift away from an exam emphasis benefits female students, closing gaps in overall performance. These data confirm our previous findings from performance- and survey-based studies, which revealed the negative impact of test anxiety on female performance in introductory biology. New findings, from a study of several thousand students in introductory STEM courses, indicate that first-generation college students exhibit similar trends-specifically, lower confidence and higher test anxiety relative to their continuing-generation counterparts. Here we broaden our understanding of the inclusive classroom by testing a number of ways to create assessments that are relevant to actual professional skills for developing scientists. We conclude by challenging the student deficit model, and suggest the course deficit model as explanatory of these performance gaps, whereby the microclimate of the classroom can either raise or lower barriers to success for underrepresented groups in STEM.

**Is it what you know or who you know? Exploring instructor assessment viewpoints in relation to social context** Melody McConnell\*, North Dakota State University; Lisa Montplaisir, North Dakota State University; Erika Offerdahl, Washington State University [abstract # 119]

In-class assessment and instructor-generated feedback are important aspects of evidence-based undergraduate science instruction. To help students learn and excel, it is imperative that we support college instructors in adopting such practices and learn more about the factors that may inspire changes in attitudes and practices relating to assessment. In this study, we sought to understand the role of social context in regards to instructor assessment practices. We investigated the following research questions: 1) What is the interaction between the informal academic social environment of an instructor and their viewpoints on assessment practices? 2) How, if at all, do social factors influence those viewpoints? 3) What qualities lead to some instructors being seen as a teaching resource by their peers? To address these questions, we used a measure of faculty self-reported assessment expertise (the Faculty Self-Reported Assessment Survey (FRAS); Hanauer & Bauerle 2015), a social network survey, and interviews. Once per semester over 2.5 years, we administered the FRAS and a social network survey to instructors within a biology department with an accepting culture of teaching innovation and rich in professional development opportunities. The FRAS allowed us to measure the self-reported assessment expertise of faculty over time within such a context, including dimensions of knowledge, confidence, attitude, and practices. The social network survey asked faculty who they talk to informally about teaching, allowing us to explore those connections. Overall, the average reported assessment expertise values in the department increased over time (although this may be due to participation bias) and are significantly correlated with the number of people who report teaching conversations with each participant. We chose individuals with high or low assessment expertise scores, and whose scores changed over time, and followed up with them by conducting interviews to further probe both social connections and assessment ideas. Hypothesized drivers of changes in assessment ideas and practices, both positive and negative, will be discussed based on these interviews. In two semesters, we also asked instructors to name anyone in the department whom they see as a teaching resource. Those faculty who were named as a resource tended to have higher self-reported assessment expertise on average and included faculty involved in DBER; however, there was a wide range. Instructors named as a resource varied widely in attitudes and reported practices, from instructors who lead the department in implementing active learning to instructors who advocate traditional lecture. Taking into account these findings could be beneficial for universities and departments, as they could find ways to encourage productive social connections and target professional development to best support instructors in implementing evidence-based teaching techniques.

**SUNDAY afternoon**

**Assessing student Practices- Sunday afternoon**

**The Classroom Discourse Observation Protocol (CDOP) for Undergraduate STEM Classrooms: A New Instrument to Characterize Teacher Discourse Moves** Petra Kranzfelder\*, University of Minnesota; Jennifer Bankers-Fulbright, Augsburg University; Marcos García-Ojeda, University of California, Merced; Sagal Mohammed, University of Minnesota; Vinit Vaghani, University of Minnesota; Lindsey Walker, University of Minnesota [abstract # 11]

Student-instructor interactions increase when instructors teach using active learning strategies, creating opportunities for productive classroom discussions that guide student thinking and learning. Currently, most observational protocols used in undergraduate STEM classrooms focus on the instructor and student behaviors. For example, Smith et al. (2013) developed COPUS (Classroom Observation Protocol for Undergraduate STEM) to characterize how instructors and their students spend time in undergraduate STEM classrooms. However, while COPUS and other extant protocols can measure the prevalence of student behavioral engagement, they are not designed to measure the nature of teacher discourse moves. Teacher discourse moves (TDMs) are a type of classroom discourse where instructors use conversational strategies to foster the development of content knowledge. Previous studies have been conducted on TDMs, especially in the K-12 space, yet no instrument currently exist that measures TDMs from observational data. Therefore, we developed a new instrument, the Classroom Discourse Observation Protocol (CDOP), to reliably measure and quantify TDMs from observational data in undergraduate STEM classrooms. The specific aims of our research were to 1) generate TDM codes for CDOP based on the literature and course transcripts; 2) determine the reliability, validity, and ease of use of CDOP; and 3) use CDOP to characterize the nature of TDMs in undergraduate STEM classrooms. We videotaped 13 faculty teaching undergraduate biology courses administered by the Department of Biology Teaching and Learning in the College of Biological Sciences at the University of Minnesota. Six class sessions out of the 37 video-recorded class sessions were selected for the transcription and development of TDM codes. These classrooms were chosen because they had the highest active-engagement instruction (i.e. average percent instructor guiding and student talking to class) based on COPUS results. Preliminary results suggest that courses with high active-engagement instruction tend to promote productive discourse, but there is variation in the nature of the TDMs between instructors. We used extant literature on TDMs and an iterative process of coding the six class sessions between multiple coders to generate a total of 22 initial TDM codes. We will describe the development and the validation of the CDOP based on these TDM codes and discuss how discourse observation data can be used to increase our scientific understanding of STEM teaching and improve instructional practices in undergraduate STEM learning environments.

**Student, instructor, and observer agreement regarding frequencies of Scientific Teaching practices using the Measurement Instrument for Scientific Teaching-Observable (MISTO)** Mary Durham\*, University of Nebraska-Lincoln; Jenny Knight, University of Colorado, Boulder; Alex Paine, University of Colorado; Brian Couch, University of Nebraska-Lincoln [abstract # 106]

The Scientific Teaching (ST) pedagogical framework encompasses many of the best practices highlighted in the literature and national reports. Understanding the growth and impact of ST requires instruments to accurately measure the extent to which ST is used in science courses. Researchers typically rely on students, instructors, or observers to document practices, but each of these perspectives has benefits and limitations. Students represent a large sample size, but are criticized for infusing personal biases. Instructors have pedagogical expertise, but may over-report implementation. Observers are less prone to biases, but often evaluate small samples of class sessions. Few studies have compared these viewpoints, and they typically use different instruments for each perspective, so it remains unclear whether and how these perspectives differ. Addressing this problem was our rationale in developing an instrument to measure teaching practices from all three perspectives. We modified our previously published instrument to create the Measurement Instrument for Scientific Teaching-Observable (MISTO), which estimates ST practice frequencies in undergraduate science courses and can be completed by students, instructors, and observers. We used MISTO to investigate these research questions: To what degree do student mean, instructor, and observer MISTO scores correlate with each other across subcategories? How closely do the perspectives match on individual items and across subcategories? Our research design included collecting MISTO scores from students, instructors, and observers in 70 undergraduate science courses at 7 U.S. institutions. Students and instructors completed the survey online outside of class using Qualtrics. We developed a video scoring workbook for observers to analyze a week of video-recorded class sessions from each course. Embedded workbook formulas calculate frequencies and proportions of ST practices and generate automated observer responses. Responses were used to calculate full MISTO and subcategory scores. Pearson’s correlations gauged the similarity of responses among perspectives and a match score (1 – ((|score1 – score2|) / maximum score)) estimated the level of agreement between perspectives. We tested for differences in match scores using ANOVAs with pair-wise Tukey tests. Overall, the level and type of agreement between perspectives varied across the full MISTO and its subcategories. Full MISTO and Active Learning (AL) showed moderate to high correlations ((r = 0.59 – 0.91, p < 0.001) and high match scores (> 0.75), suggesting that each viewpoint can produce similar estimations on these scales. The Experimental Design, Data Analysis, and Cognitive Skills subcategories had moderate to low correlations (r = 0.2 - 0.7) and high matches (> 0.75). Practices in these categories were less common than AL, which partially explains the lower correlations despite high matches. Inclusivity (IN) and Responsiveness to Students (RtS) showed low correlations (r ≤ 0.3) and low matches (most < 0.75), which was likely an artifact of agree-disagree response scales. Student-instructor match scores were significantly higher than observer matches in the full MISTO, AL, IN, and RtS subcategories (p<0.001). This likely stems from intrinsic differences in course access and scoring criteria of the perspectives. MISTO is a broadly usable instrument that can be used to gauge ST levels in science courses and make direct comparisons between student, instructor, and observer perspectives. To our knowledge, this is the first study to compare these perspectives using a single instrument. Building on the nuanced nature of our results, we recommend MISTO users consider research goals, available resources, and potential artifacts when deciding which perspective best fits their needs in measuring classroom teaching practices.

**COPUS profiles 2.0: Characterization of instructional practices in STEM courses offered in North American universities**  Marilyne Stains\*, University of Nebraska-Lincoln; Jordan Harshman, Auburn University; Marc Levis-Fitzgerald, University of California Los Angeles; Michelle Smith, University of Maine; MacKenzie Stetzer, University of Maine; Nicole Michelotti, University of Michigan; Jenny Knight, University of Colorado, Boulder; Patricia Schulte, University of British Columbia; Timothy McKay, University of Michigan; Blaire Van Valkenburgh, University of California Los Angeles; Megan K. Barker, Simon Fraser University; Stephanie Chasteen, University of Colorado Boulder; Renee Cole, University of Iowa; Sue Ellen DeChenne-Peters, Georgia Southern University Armstrong Campus; Kathryn M. Plank, Otterbein University; M. Kevin Eagan, University of California Los Angeles; Joan M. Esson, Otterbein University; Frank A. Laski, University of California Los Angeles; Christopher Lee, University of California Los Angeles; Stanley Lo, University of California San Diego; Lisa McDonnell, University of California San Diego; Michael Palmer, University of Virginia; Tamara M. Rodela, University of British Columbia; Erin Sanders, UCLA; Natalie G. Schimpf, University of British Columbia; Erin Vinson, University of Maine; Laura Weir, University of British Columbia; Paul Wendel, Otterbein University; Lindsay Wheeler, University of Virginia; Anna Young, Otterbein University; Amanda Musgrove, University of California Los Angeles [abstract # 198]

As evidence-based instructional practices (EBIPs) have gained more and more momentum, the interest in assessing the effect of their implementation has grown. To best assess this, we need a national baseline that characterizes how instructors currently teach their courses so in the future we are able to assess the changes in teaching as EBIPs continue to be disseminated, encouraged, and ultimately implemented in undergraduate science courses. We aimed to make progress toward this goal by addressing the following research questions using a large data set: 1) What are the different instructional styles that faculty employ in Science, Technology, Engineering, and Mathematics (STEM) undergraduate courses? Do these styles vary by STEM discipline, type of course (lower-level versus upper level undergraduate courses), class size, and classroom physical layout? To date, characterizations of the instructional landscape in STEM courses in higher education are mostly based on faculty-based, self-reported surveys, which are prone to validity threats. We thus leveraged a successful observation-based method previously developed to characterize instructional practices and implemented it on a large scale. In particular, we collaborated with eleven discipline-based education research teams across North America and characterized a snapshot of the instructional landscape in undergraduate STEM courses by observing 2,028 classes taught by 545 STEM faculty from over 25 institutions. Each class was codified using the Classroom Observation Protocol for Undergraduate Science (COPUS) that measures 13 student and 12 instructor behaviors every two minutes of each class. A total of seven instructional profiles were determined using mixture model clustering. The most prominent behaviors were lecturing for instructors and listening for students, resembling a traditional didactic instructional style (~55%). Another ~27% of instructors were classified into interactive lecture categories while the remaining 18% demonstrated predominately student-centered strategies. We found that although student-centered teaching practices were more prominent than expected in flexible classroom layout and small courses, didactic teaching still accounted for about 50% of the observations in these environments. Didactic teaching also dominated throughout the undergraduate curriculum. A comparison of instructional practices by discipline indicated that mathematics and geology have more student-centered styles than expected, biology has more interactive styles than expected, and chemistry has more didactic styles than expected. Additionally, we discovered that an instructor’s practices can vary significantly within a course, indicating that multiple observations are required to gain a complete picture of that faculty’s instruction. Based on the instructional profiles determined, an online app was developed that allows users to import and analyze COPUS data from individual or a group of instructors (www.copusprofiles.org). Additionally the app allows users to classify their observations into one of the seven instructional profiles determined. We will briefly showcase the app during this talk.

**The Effects of Instructor Support in Active Learning Biology Classrooms on Student Motivation and Performance**  Kim Pigford\*, NC State University; Miriam Ferzli, NC State University; Margaret Blanchard, NC State University [abstract # 223]

According to self-determination theory, learning environments with instructors that actively support students’ basic psychological needs for competence, relatedness, and autonomy result in intrinsically motivated learners that exhibit higher performance and overall wellbeing. Courses using active learning, due to how they are taught, should actively support and provide for students’ needs by giving them the opportunity to actively engage in course material in a meaningful manner under the guidance of a supportive instructor. Previous research has shown that undergraduate STEM active learning environments positively impact students’ motivation and performance for learning. This study explains these gains by examining levels of instructor support in three sections of an introductory biology course each utilizing differing amounts of active learning: a traditional lecture section with no active learning (240 seats), a modified lecture conducted in the same auditorium setting as the traditional lecture but utilizing various active learning methodologies (240 seats), and SCALE-UP (96 seats; Student-Centered Active Learning Environments with Upside-down Pedagogies) a specialized, collaborative active learning environment. It also explores the impact of instructor support on student motivation, performance, and other affective measures. To examine the effects of active learning environments on students’ perceived levels of instructor support, students filled out the Learning Climate Questionnaire, a validated Likert-style survey that assesses how well the instructor of a course provides for the students’ basic psychological needs. Differences between the responses of the three sections were analyzed using a one-way ANOVA analysis. Results showed that levels of instructor autonomy support were higher in the two sections utilizing active learning than the traditional lecture section (p< .001), however, there were no differences between the two active learning sections despite instructor support levels being slightly higher in the SCALE-UP section. To examine the impacts of instructor support on student motivation, performance, and other affective measures, students completed the Intrinsic Motivation Inventory, a set of validated Likert-style surveys that measure intrinsic motivation and other related affective variables (effort, value, competence, perceived choice, tension). Performance data for each student was collected in the form of the final course grade. Additionally, students’ high school GPA and SAT scores were collected to measure and control for student ability. To examine the impact of instructor support, each measure was individually regressed onto perceived instructor support after first removing variance attributable to student ability. The individual step-wise linear regression models showed that levels of instructor support significantly predicted all measures (p< .001). Instructor support positively predicted students’ levels of intrinsic motivation (R-sq.=.343), effort (R-sq.= .096), value (R-sq.= .269), competence (R-sq.= .292), perceived choice (R-sq.= .046), and performance (R-sq.= .266), while negatively predicting levels of student tension (R-sq.= .191). This study suggests that due to the nature of student-centered learning environments, students enrolled perceive their instructor as being more supportive of their basic psychological needs than instructors in traditionally taught classrooms. This additional support then nurtures students’ development of volitional, intrinsic motivation towards learning and results in higher performance and affective gains. The study provides further evidence of the benefits of active learning environments and suggests that the manner in which these courses are taught results in a more supportive learning experience for the student leading to higher performance and affective gains as noted in earlier studies.

**Instructor Training- Sunday afternoon**

**Pedagogical knowledge for active-learning instruction in large undergraduate biology courses: A large-scale qualitative investigation of instructor thinking** Tessa Andrews\*, University of Georgia [abstract # 18]

Though active-learning instruction has the potential to positively impact the preparation and diversity of STEM graduates, not all instructors are able to achieve this potential. One important factor is the teacher knowledge that instructors possess, including their pedagogical knowledge. Pedagogical knowledge is knowledge about teaching and learning that is not topic-specific, such as knowledge of learning theory, classroom management, and student motivation. Pedagogical knowledge has been studied much less extensively than other teacher knowledge bases (e.g., pedagogical content knowledge), even among K12 instructors, and there have been no studies of college instructors. Yet a key difference between active-learning instruction and a more traditional lecture approach is the pedagogy. Thus, there is a need for work to define and investigate pedagogical knowledge for active-learning instruction. We investigated the pedagogical knowledge that 77 active-learning instructors used as they analyzed video clips of lessons in large active-learning biology courses. We used qualitative content analysis, and drew on cognitive and sociocultural perspectives of learning, to identify and characterize the pedagogical knowledge instructors employed. Our qualitative analyses were collaborative and iterative and used a constant comparative approach to ensure our findings are trustworthy and solidly grounded in our data. We used the collective thinking of these instructors to generate a framework of pedagogical knowledge for active-learning instruction in large undergraduate biology courses. We identified seven distinct components of pedagogical knowledge, as well as connections among these components. At the core of their thinking, participants evaluated whether instruction provided opportunities for students to generate ideas beyond what was presented to them and to engage in scientific practices. They also commonly considered student motivation to engage in this work and how instruction maximized equity among students. Participants noticed whether instructors monitored and responded to student thinking in real-time, how instruction prompted metacognition, and how links were built between learning tasks. Participants also thought carefully about managing the logistics of active-learning lessons. Active-learning instructors displayed knowledge of principles of how people learn, practical knowledge of teaching strategies and behaviors, and knowledge related to classroom management. Their deep knowledge of pedagogy suggests that active-learning instruction requires much more than content knowledge built through training in the discipline, yet many college STEM instructors have little or no training in teaching. Further research should test this framework of pedagogical knowledge in different instruction contexts, including different STEM disciplines. Additional research is needed to understand what teacher knowledge is critical to effective active-learning instruction and how the development of this knowledge is best facilitated. Achieving widespread improvement in undergraduate STEM education will likely require transforming our approach to preparing and supporting undergraduate instructors.

**High School Teacher Conceptions and Lessons about Meiosis Reveal a Critical gap in Molecular level Knowledge** Kate Wright\*, Rochester Institute of Technology; Dina Newman, Rochester Institute of Technology [abstract # 200]

Meiosis, which is intimately linked with concepts of chromosomal behavior and patterns of inheritance, is the foundation on which ideas in genetics and evolution are built. While instruction on meiosis often begins in the middle school years many students continue to struggle with these concepts throughout their undergraduate training. The DNA triangle framework (Wright et al, 2018), which links the chromosomal (structural), molecular (sequence of DNA) and informational (genes and alleles) levels of DNA, can be used to conceptualize three important ideas of meiosis; ploidy, homology and mechanism of homologous pairings. Previous work demonstrated experts are able to appropriately integrate the molecular and informational levels of DNA to explain homology, integrate the chromosomal and informational levels to explain ploidy and integrate the chromosomal and molecular levels to explain mechanism of homologous pairing. Undergraduates, however, struggle to integrate multiple levels of DNA knowledge, especially at the molecular level. To better understand why this gap in molecular knowledge is so widespread, we interviewed 14 High School Biology teachers (experience ranging from 2 to 26 years) about their own understanding and approach to teaching meiosis. Teachers were asked to explain how they conceptualized ploidy, homology and homologous pairing and how they explained these concepts to their own students. We analyzed the qualitative data through the lens of the DNA triangle and found that the majority of teachers (11 of 14) did not incorporate any molecular level knowledge into their own explanations of meiosis and none emphasized DNA structure during meiosis instruction. Homology, for example, was explained at the chromosomal (same size and shape) and informational (same genes) levels, but not the molecular level (nearly the same sequence of DNA nucleotides). Analysis of 20 high school lessons about meiosis yielded similar results; instruction focused on the chromosomal and informational levels but not on the molecular level. During interviews, several teachers explicitly acknowledged the separation between instruction about DNA (i.e. structure and complementary base-pairing) from lessons on meiosis. While this work suggests a critical gap in the high school biology curriculum it also highlights an opportunity to offer new professional development opportunities for high school teachers and curriculum developers. This work also partially explains why many undergraduate biology students have difficulty thinking about complex processes, like meiosis, at the molecular level.

**Using feedback to student misconceptions as a tool for assessing and developing teacher pedagogical content knowledge** Revati Masilamani\*, Tufts University; Carol Bascom-Slack, Tufts University School of Medicine; Berri Jacque, Tufts University [abstract # 268]

Teaching multi-concept topics like natural selection and evolution to students is challenging. For biology teachers, effective teaching of these topics involves the ability to identify and address student misconceptions. However many teachers hold misconceptions themselves, and even when aware of student misconceptions, often focus on providing accurate content without diagnosis and correction of the nature of the misconception. It is therefore essential to provide teachers with the content knowledge (CK) as well as the pedagogical content knowledge (PCK), i.e. the interface of their pedagogical knowledge with the content domain, to be able to teach a topic like evolution successfully. It is generally agreed that experience is the main source of PCK and that adequate CK must precede development of PCK. However it has also been suggested that PCK can be developed through professional development (PD) courses. We asked whether CK and PCK could be built simultaneously via short duration PD and whether the gains were dependent on prior teacher experience. For the purpose of this study, we defined PCK as the ability to identify and correct student misconceptions. The study was carried out with 29 teachers in a 13 week-long in-person course meeting once a week, on the topic of infectious diseases To assess both CK and PCK, we designed a pre-post test. Part 1 of the test contained content questions of multiple choice and short answer formats to establish teacher CK on evolution. Part 2 of the test contained versions of those same questions with student answers that teachers were asked to grade, to demonstrate their PCK. In their grading teachers were given specific and detailed instructions to identify the student misconception and address them. We used the pre-test as a formative assessment to design an in-class intervention, Teachers were assigned to small groups and given a couple of questions similar to those in the PCK section of the test, and were asked to first write down their individual feedback to the student, followed by a group discussion. After the discussion, each teacher was asked to write their perspective on the group’s discussion and if/how it had transformed the feedback they would give the student. Apart from pre-post test responses and in-class discussion artifacts, we collected teacher responses to pre-post Likert style attitudinal surveys to assess changes in confidence towards teaching the topic, and identifying student misconceptions. The data was analyzed by a mixed-methods approach. Numerical scores were derived from the Likert-style survey questions and the teacher CK responses in the tests that were graded on an accuracy rubric. The PCK responses from the tests and in-class discussion were analyzed using a rubric to assess accuracy and level of detail provided. The key findings of this study were that 1) Unless trained to, most teachers with high CK tend to write out the correct answer, instead of parsing through student errors in thinking. 2) Teachers who begin with high CK but low PCK in the pre-test, improved their PCK through the in-class collaborative discussion activity, as reflected by higher PCK scores in the post-test. 3) Teachers who began with both low CK and PCK improved their content understanding through the in-class group discussion activity, as well as their ability to address misconceptions, as reflected by a higher CK and PCK on the post-test. 4) Teachers with more experience began the course at a higher level of PCK than teachers with less experience, but both cohorts improved PCK through the course. These data suggest that teacher PCK can be improved in conceptually challenging content areas of biology, and that using student responses for teacher feedback practice in a collaborative setting not only improves PCK but can also be a tool to provide CK. Also, even short duration PD can transform teacher ability and confidence to address student misconceptions, and should do so.

**Through the eyes of faculty: using personas as a tool for user-centered professional development** Patricia Zagallo\*, University of Georgia; Jill McCourt, University of Puget Sound; Bo Idsardi, University of Georgia; Kevin Haudek, Michigan State University; Jenny Knight, University of Colorado, Boulder; John Merrill, Michigan State University; Ross Nehm, Stony Brook University; Luanna Prevost, University of South Florida; Michelle Smith, University of Maine; Mark Urban-Lurain, Michigan State University; Paula Lemons, University of Georgia [abstract # 148]

This work aims to identify the best ways to support biology faculty amidst the national call to implement student-centered approaches in undergraduate STEM education. Our research goal is to design teaching professional development (TPD) opportunities that take into account individuals and their perceived situational affordances and barriers. Our rationale is that doing so will enable faculty to adopt and sustain the changes to which they've been called. The theoretical framework for our work is situated learning, which asserts that knowledge emerges within a dynamic social context as learners participate. We view college faculty as learners, and we aim to equip ourselves and other change agents with insights about faculty that can be used to customize TPD experiences. To that end, we investigated the following research questions for a cohort of nineteen biology faculty: What are their knowledge and values regarding students and the educational context? What are their attitudes toward instructional innovation? What outcomes do they seek from teaching professional development? We investigated our research questions through persona construction, a user-centered approach common in marketing research. Persona construction utilizes ethnographic methodologies to create fictional characters that convey robust data about real users in engaging ways. Data on faculty were collected through two semi-structured interviews and two or more classroom observations per participant using COPUS, as part of a five-year project with Automated Analysis of Constructed Response (AACR). AACR provides faculty with computer-scored, free-response items for timely formative assessment in large-enrollment classes and supports interpretation of reports through faculty learning communities. Interview transcripts were analyzed by developing an emergent coding scheme that addressed our research questions. From clustering similar themes in our coding analysis, we developed several candidate personas to describe the variation among participants. For example, Wendy is an example of one persona capturing blended coding themes. Wendy earnestly prepares lectures that align with the textbook because she knows students take comfort in in their notes and textbooks. She cares about understanding what students struggle with conceptually, thus she values getting to know students and helping them during office hours. She feels she cannot scale-up the personal interactions she experiences in office hours to large-enrollment classes. Given that the majority of her student body are pre-medical students, she sees it as her job to sort students for medical school admission committees and believes the more content she can cover, the more prepared her students will be. Thus, she struggles with the instructional choice of including active-learning activities because she feels guilty covering less content. What she seeks most out of TPD is the chance to talk with her colleagues about teaching because she wants to get better at what she does. Further analysis includes prioritizing and validating each persona. This study adds to the literature base in biology education by providing a robust, narrative approach to understanding faculty as individuals who are themselves consolidating their educational views and belief systems with innovative practices. The study provides implications for teaching and learning by providing insights to change agents on how best to meet faculty in their change process. Thus, we expect these findings to be of interest to SABER attendees who research faculty in the change process or those who provide TPD.

**Data Mining-Sunday afternoon**

**Improving Inter-rater Reliability in Qualitative and Mixed Methods Content Analysis: Lessons from Computer Analysis of Student Writing** Mark Urban-Lurain\*, Michigan State University; Kevin Haudek, Michigan State University; John Merrill, Michigan State University; Brian Donovan, BSCS; Molly Stuhlsatz, BSCS; Christopher Wilson, BSCS [abstract # 170]

Many biology education researchers use qualitative or mixed methods. Qualitative data can provide insight beyond what is happening into the why, the meanings that participants bring to the experience. Qualitative analyses can also be used to triangulate with quantitative data to provide richer insight into a research problem. Much of qualitative analysis entails the iterative development of a coding scheme, which is applied to the data and refined. To address validity and reliability, qualitative data analysts often have multiple coders code data independently, then measure inter-rater reliability (IRR) among coders. Cohen’s Kappa (K) and other measures of IRR are easily calculated, but the guidelines for interpreting them are arbitrary. Landis and Koch (1988) recommend 0.0 – 0.2 as slight agreement, 0.21 – 0.4 as fair, 0.41-0.60 as moderate, 0.61-0.8 as substantial and > 0.8 as almost perfect agreement. Many researchers are satisfied achieving values above 0.6. These values are often arrived at after extensive coding discussion and revision among raters. Our research group creates computerized models that predict expert ratings of student writing, which depend on qualitative data analyses to score a large corpus of student writing. We create rubrics, have multiple raters score data, calculate IRR among raters, refine the rubrics, rescore, and then resolve scoring differences. Then we use natural language processing, statistical, and machine learning techniques to create computer models that predict expert scores of student responses. We refine these computer models by reviewing expert-computer disagreements and use that as feedback for rubric and scoring refinement. RQ1) What are the implications of varying IRR values for the meaning of the results? To explore the impact of varying IRR among expert scorers on the accuracy of the computer models, we looked at the relationship between the trained expert IRR (TE-IRR) and the computer-to- expert consensus IRR (CC-IRR) on 25 analytic rubrics applied to student responses across multiple questions. RQ2) How can we improve expert IRR? To improve TE-IRR for holistic rubrics, we created a series of binary analytic rubrics to be scored by experts, which were then combined using Boolean operators to determine a holistic score for each student response. We evaluated the accuracy of the CC-IRR by training computer models on the resulting holistic scores. Results: RQ1) We computed pairwise TE-IRR among scorers after they were trained. We used consensus expert scores, after resolving disagreements, to create computer models and calculated the CC-IRR. The Pearson correlation between the TE-IRR and CC-IRR was r = 0.90. For rubrics where the TE-IRR < .6, the CC-IRR tends to underperform TE-IRR. The CC-IRR tends to be slightly better than TE-IRR when the TE-IRR is > .6. This provides evidence that K > .6 should be the minimum acceptable expert IRR for research purposes and that resolving differences with a third scorer does not reduce ambiguity in the rubric. RQ2) Performance on holistic scores varied based on the TE-IRRs on the analytic rubrics for each item. While the CC-IRRs for analytic rubrics track the TE-IRRs, the holistic CC-IRRs are better than the averages of TE-IRRs on the individual analytic rubrics. We also combined the machine predictions for analytic rubric scores using the same Boolean rules to calculate holistic scores. Both methods to calculate holistic scores perform similarly (r=0.95). This suggests that creating binary analytic rubrics and combining them for holistic scores can provide better expert IRR results than having experts apply holistic rubrics directly. We will present further details and discuss how to apply these guidelines in qualitative rubric development and analysis. The talk should be of interest to qualitative and mixed methods biology education researchers by providing a framework for improving the validity and reliability of research results.

**How to Get More Out of Conceptual Assessment Data** Dina Newman\*, RIT; Kate Wright, Rochester Institute of Technology [abstract # 250]

Robust conceptual assessment tools are crucial for testing new activities, assessing pedagogical strategies and tracking impacts of curricular changes. While the vast majority of instruments use the forced choice response format, some researchers are using non-traditional questions and analysis methods to measure student learning. Multiple select and multiple true/false questions may be preferable to forced choice formats because they are less susceptible to guessing. Students are far less likely to be able to preferentially accept or eliminate choices based on non-content-based clues among the responses, as they must consider each option independently of the others. In addition, the random guess rate is much lower (e.g., 3% for a 5-option multiple select question, compared to 20% for a 5-option forced choice question). Scoring multiple select or multiple true/false instruments is more complex, but can yield deeper insights into student learning. Using our large collection of multiple response data from the Central Dogma Concept Inventory instrument (n=2164) we have begun to explore multiple ways to analyze and represent the gains that students make when learning new concepts. While learning gain data is valuable, it does not always reveal the true incremental (and oftentimes incomplete) gains that students make when attempting to learn new things. Analysis of partial responses can reveal overly “stubborn” misconceptions while also highlighting where instructional strategies are most successful. There are multiple ways to look at these more complex data, such as tracking the percent of correct and incorrect choices, calculating the ratio of correct to incorrect choices compared to each question’s guess rate, and comparing the slope of change over time. Latent class analysis can identify patterns in student responses that can provide insight into student thinking. Our analysis of multiple select data aligns with theories of learning; students are faster to accept new correct ideas than they are to “let go” of incorrect (often deeply ingrained) ideas (e.g. slope of 0.11 for changes in correct response selection over two years compared to -0.03 for changes in incorrect response selection). Breaking up the questions into conceptual categories for analysis indicates that some concepts are particularly strengthened in an advanced course (e.g. the nature of genetic information) while others were strengthened in the introductory course and then lost ground—perhaps due to not reviewing the concepts—in later courses (e.g. building blocks of macromolecules). We feel that this work is important to share with the Biology Education Research community so researchers and educators can get the most out of their assessment data to move the field forward and improve education for biology learners.

**Using data mining methods to quantify the contributions of institutional and course-specific data to student success predictions** Yaqi Xue\*, Stony Brook University; Stephen Finch, Stony Brook University; Ross Nehm, Stony Brook University [abstract # 156]

Learning Analytics (LA) seeks to harness big data to answer questions and to test hypotheses about student performance outcomes. Data mining is transforming many aspects of daily life, but has rarely been used in biology education research. LA frameworks and data mining methods have the potential for improving student learning outcomes and mitigating failure. One important goal of LA is to predict which students will need help before it is too late. Predicting failure in introductory biology courses would be valuable for: (1) social engineering (e.g. forming heterogeneous groups within active-learning classrooms (2) targeting supplemental instruction (e.g. delivering tailored instructional materials, prerequisite concepts, and learning scaffolds), and (3) aligning students with course pathways that promote success (e.g. placement decisions). Currently, most research on predictive analytics has utilized institutional data. Our study explores (1) the degree to which biology assessments and course-specific data enhance predictive accuracy above and beyond standard institutional research data, (2) the most effective Imputation Strategies (IS), and (3) the most effective Data Mining Methods (DMM) for prediction. We used data from 3145 total students from six semesters. Two experiments were run to assess variability of model performance. The first experiment used 2077 students from the first four semesters to build predictive models for 503 students in the fifth semester; and the second experiment used 2021 students the second to fifth semesters to build predictive models for the 565 students in the sixth semester to test the models. The data included institutional (e.g. demographic, pre-college, and college predictors) and course-specific predictors (e.g. pre-course diagnostic assessments, clicker attendance, and blackboard logins) from four time points: pre-course, 3, 6, and 9 weeks into the semester. The students and course data were used to build and test models predicting final course performance (final grade <= D). Methodologically, we studied the comparative efficacy of seven DMMs (generalized linear model, generalized linear model with elastic net, decision tree, K-nearest neighbors, 1-hidden layer neural network, stochastic gradient boosting, random forest) and nine ISs (most common, concept most common, event covering, K means, fuzzy K means, K nearest neighbors, weighted K nearest neighbors, Local least square, support vector machine) to build and test the predictive models. Five widely-used measures evaluated the efficacy of the predictions: sensitivity (our target measure--number of predicted and actual failures over number of actual failures), specificity, accuracy, precision and AUC (area under the receiver operating characteristic curve). Our analyses indicated that sensitivity averaged over DMMs, ISs, time points, and 2 semesters was 52.4% (university only) and 57.4% (university+course). The k-nearest neighbor DMM using the IS strategy of fuzzy k-means clustering had sensitivities that substantially increased over time: 52.6% at pre-course, 71.2% at 3 weeks, 73.9% at 6 weeks, and 80.3% at 9 weeks. The k-nearest neighbors method using both university and course variables had the highest sensitivity averaged over IS method, time points, and semesters (65.4%). In summary (1) there were no significant differences among IS strategies; (2) no DMM was universally best among the five evaluative measures; and (3) adding course-specific data improved predictive accuracy overall and over time (weeks 3, 6 and 9). In conclusion, our work advances understanding of both the data sources and analytic methods best suited to predictive accuracy in UG biology. Our sensitivity rates exceed some prior work, and have the potential to add fresh data sources to evidence-based instructional decisions seeking to improve student success and degree pathway outcomes. Funding provided by the Howard Hughes Medical Institute.

**Developing a Model of Retention In the Biological Sciences: A Precursor to Creating Effective Interventions** Daniel Baltz\*, University of Minnesota; Meaghan Stein, University of Minnesota; Laurie Parker, University of Minnesota; Robin Wright, National Science Foundation; Anita Schuchardt, University of Minnesota [abstract # 17]

In spite of significant investments, considerable progress still needs to be made in increasing the number and diversity of students who earn STEM baccalaureate degrees. In order to design effective interventions to accomplish this goal, it is necessary to develop a model to understand the factors that enhance or inhibit the retention of biological sciences students. Prior research has identified performance on standardized tests and demographic factors as major predictors of retention and a meta-analysis has identified that psychosocial factors contribute to both college retention and GPA. However, few studies have identified the specific psychosocial factors that interact with academic performance to affect retention in the biological sciences. We were interested in developing a model that specifies how psychosocial factors interact with academic performance and student demographic variables to impact retention of biology majors at a large midwestern university. During the years 2015 and 2016, incoming Freshman cohorts in biology majors at a large midwestern university (n = 1022) were surveyed on psychosocial factors shown in the literature to impact academic performance, including Persistence on Task, Academic Self-Efficacy, and Science Confidence. Using multiple regression analysis, we had shown that the biggest predictor of student retention in the biological sciences is first semester GPA. One drawback of multiple regression analysis is that it is difficult to evaluate the interactions of multiple factors on an outcome. Thus, in our current study, we used structural equation modeling (SEM) to test hypothesized models of the combined effect of demographics, psychosocial factors, and academic performance on first semester GPA. Our results reveal that students’ self-reported tendency to persist at a task (e.g., “I finish whatever I begin.”) is the most direct psychosocial predictor of first semester GPA ( =.14, p<.001, CFI =.99, RMSEA=.03). This tendency is influenced by their sense of academic self-efficacy (e.g., “I expect to do well in all my courses.”) ( =.25, p<.001) and correlated with their confidence in performing science processes (e.g., “Please rate your confidence in your ability to design a well-controlled experiment to test a hypothesis.”) ( =.12, p<.001). Students’ first generation and underrepresented minority status interact with these psychosocial factors, suggesting that intervening to increase persistence on task will have a greater relative impact on these students’ academic performance. While the reported findings are specific to this context, the methodology demonstrates how SEM can reveal interactions and possible intervention targets that could be missed by regression analyses. This research provides an effective model for evaluating predictors of persistence in the biological sciences. In addition, the model provides