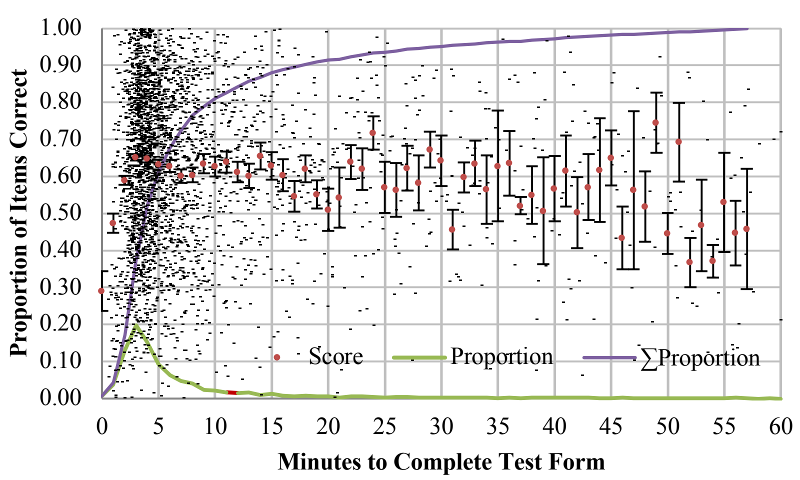
Project Goal
The goal of MOSART-HSPS is to develop rigorous assessment tools that aid in generating evidence-based measures of teacher and student understanding of high school-level physical science concepts. The project first developing multiple-choice items based on the grades 9-12 Next Generation Science Standards (NGSS) for Physical Science, divided into the domains of chemistry and physics using the published research on misconceptions related to that content. The items passed through several development stages including pilot- and field-testing with high school students. Those items yielding psychometrically impressive properties were formed into two instruments in chemistry and two in physics. Instruments were used to evaluate student and teacher understanding by running pre/post-tests in a nationwide sample of chemistry and physics courses in U.S. high schools. Instruments and research are available at no cost to teachers and researchers at the MOSART "Self-Service" website. Research findings have presented at meetings and through publication.

Major Accomplishments to Date:

1. Item Development of misconception-based multiple-choice items for grades 9-12 NGSS physical science: 395 in chemistry and 660 in physics. Draft items reviewed by external experts.
2. Pilot-Testing items using the Amazon Mechanical Turk crowd-sourcing platform to estimate preliminary item characteristics (used for selection of anchor items on field test forms): 3,186 subjects in chemistry and 1,769 in physics.

AMT Compliance: Scores and Times



3. Field-Testing was conducted to establish item characteristics using classical test theory (CTT: difficulty, discrimination, gender bias, misconception strength) and item response theory (IRT: difficulty, discrimination, guessing, item fit to 3P IRT):
   a. In chemistry with 395 items with the participation with 183 chemistry teachers and their 6,893 students and 361 physics teachers with their 6,841 students.
   b. In physics with 16,724 subjects covering 660 assessment items.
4. Instrument Construction of research and public tests for chemistry and physics from candidate items using these criteria drawing upon CTT and 3-parameter IRT characteristics:
   a. Distribution of items over all NGSS DCIs
   b. High IRT discrimination (≥0.70)
   c. Wide range of IRT difficulty (-2 to +3 theta) to have excellent Test Information Function (TIF) and low Standard Errors (SE) of theta over rang of ability (from students to teachers)
   d. High CTT misconception strength (MS>0.50) for half of items for later measurement of teacher knowledge of student misconceptions (KOSM or PCK-M)
   e. Low aggregate gender bias of instrument
5. Pre-Test/Post-Test Study involving recruitment, data collection at start and end of school year, data cleaning, and analysis:
   a. student demographic survey and content items (pre-test)
   b. student pedagogy inventory and content items (post-test)
   c. Teacher assessment of SMK and PCK-M

	Number of Teachers Participating in Study	Number of Student Pretests sent	returned	Number of Student Post tests sent	returned
Chemistry	137	7872	6761	6761	5301
Physics	149	7740	6680	6680	4281

6. Revision of Project Website. Updated MOSART Self-Service website with information regarding the NGSS; more extensive, easier-to-use, and up-to-date resources regarding student misconceptions; easier access to tests and project data sets; support for three categories of MOSART assessment users: classroom teachers, professional development programs, and science education researchers.
7. Advisory Board Meetings. Brought all advisors up to date with the project and solicited help with difficult issues, reviewed study findings, and guidance concerning additional avenues of analysis. The advisors will also recommend venues for conference presentations and article publication as well as other forms of dissemination based on the research findings. Advisory Board: Dr. Joel Mintzes, Chair, Science Learning Associates (Misconception Research); Prof. Angelica Stacy, UCal Berkeley (Chemistry); Prof. Craig Wells (Psychometrics), UMass, Amherst; Mr. Daniel Record, (Physics), Newington, CT.

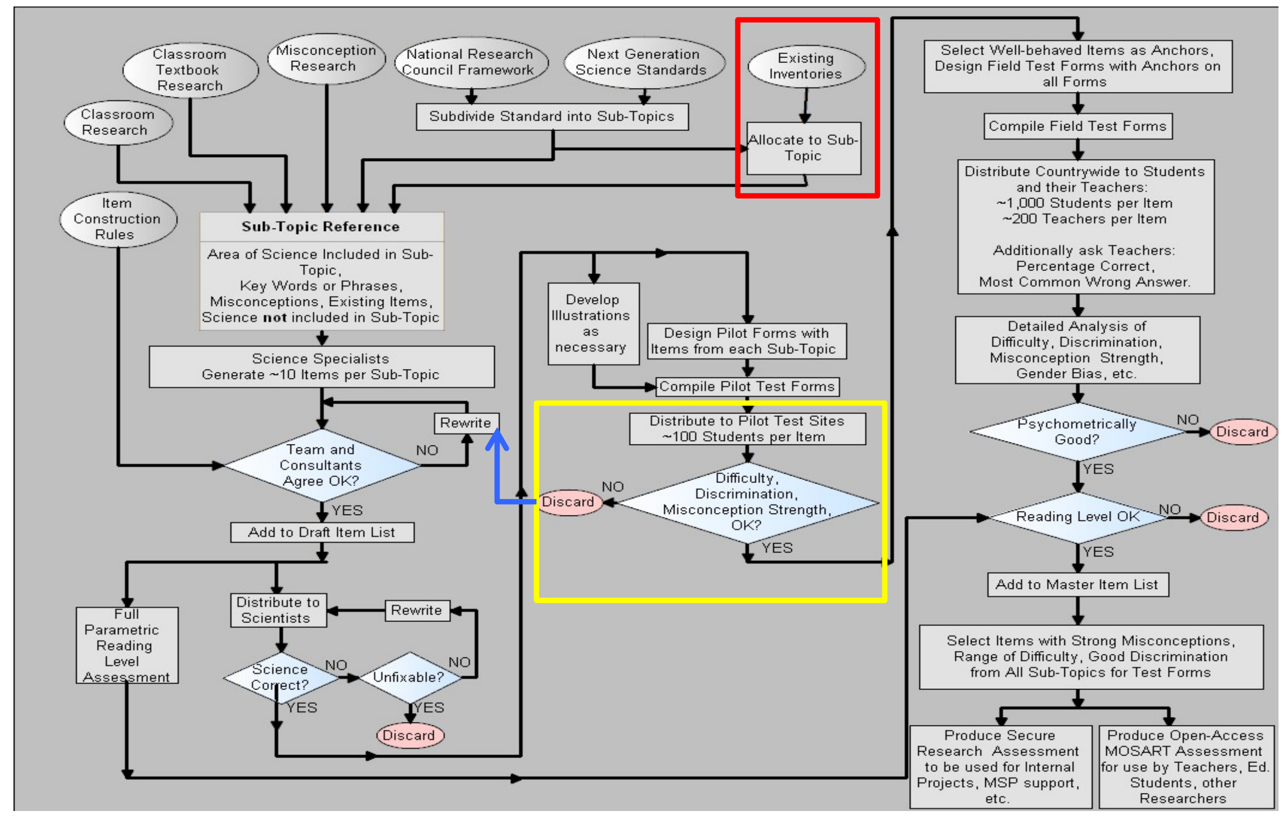
MOSART HS PS

Misconception Oriented Standards-based Assessment
Resource for Teachers of High School Life Science

DRK-12, 1621210; http://mosart.mspnet.org/

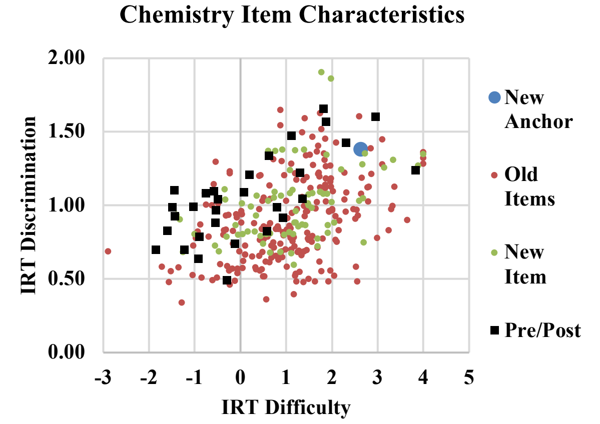
Harvard-Smithsonian Center for Astrophysics, Cambridge, MA

Philip M. Sadler, PI; Gerhard Sonnert, Co-I; Sue Sunbury, Project Manager
Cynthia Crockett, Science Education Specialist; Chen Chen, Post-Doctoral Fellow
John Murray, online test system associate; Annette Trenga, research assistant

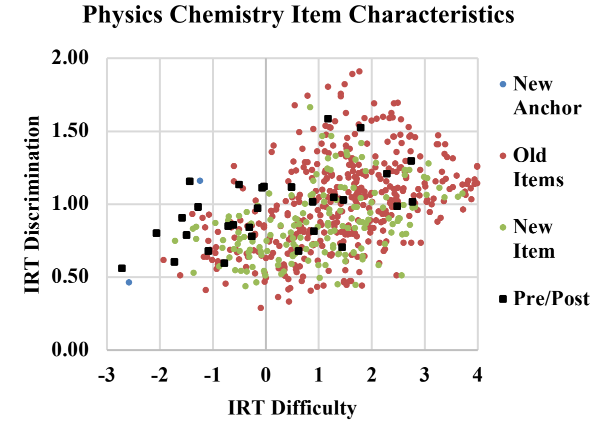


Item Inventory Development Process. Red box highlights harvesting of items from existing inventories. Yellow box highlights a key "bottleneck" of pilot-testing solved by using Amazon Mechanical Turk Crowdsourcing Platform.

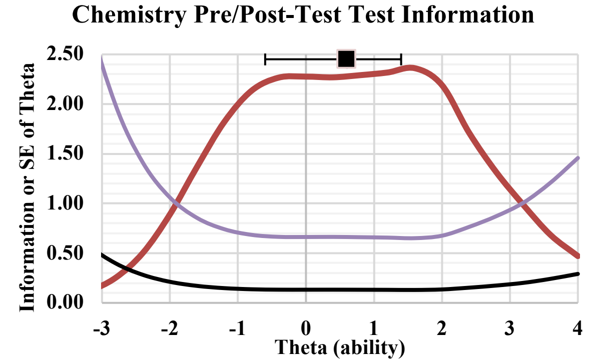
Chemistry Item Characteristics



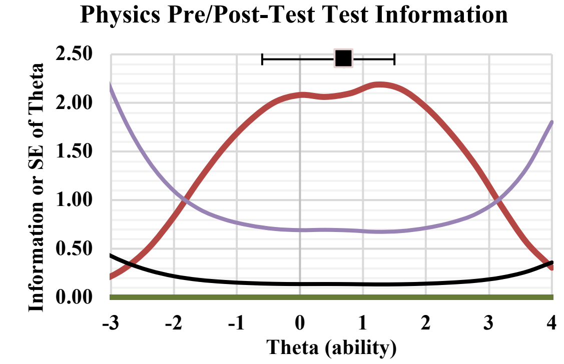
Physics Chemistry Item Characteristics



Chemistry Pre/Post-Test Test Information

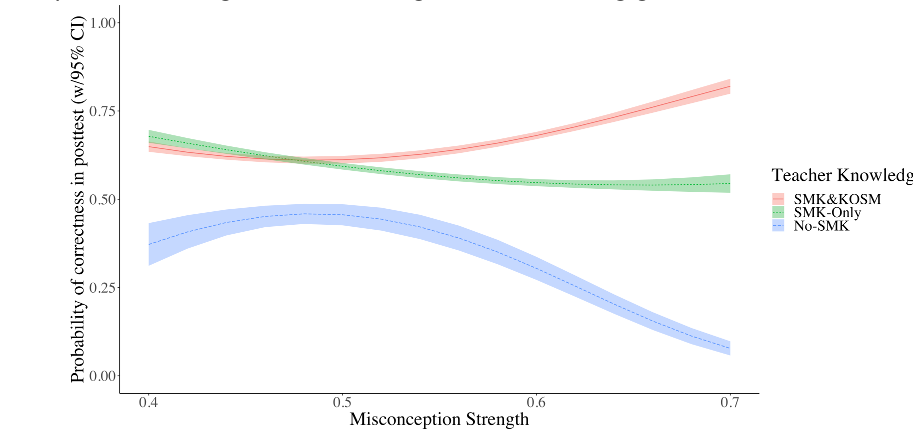


Physics Pre/Post-Test Test Information

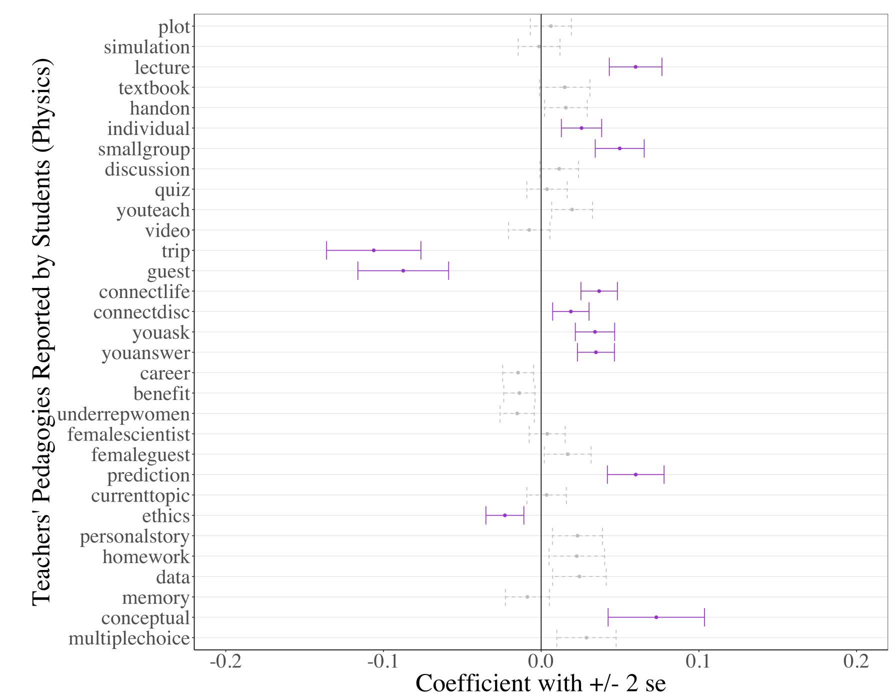


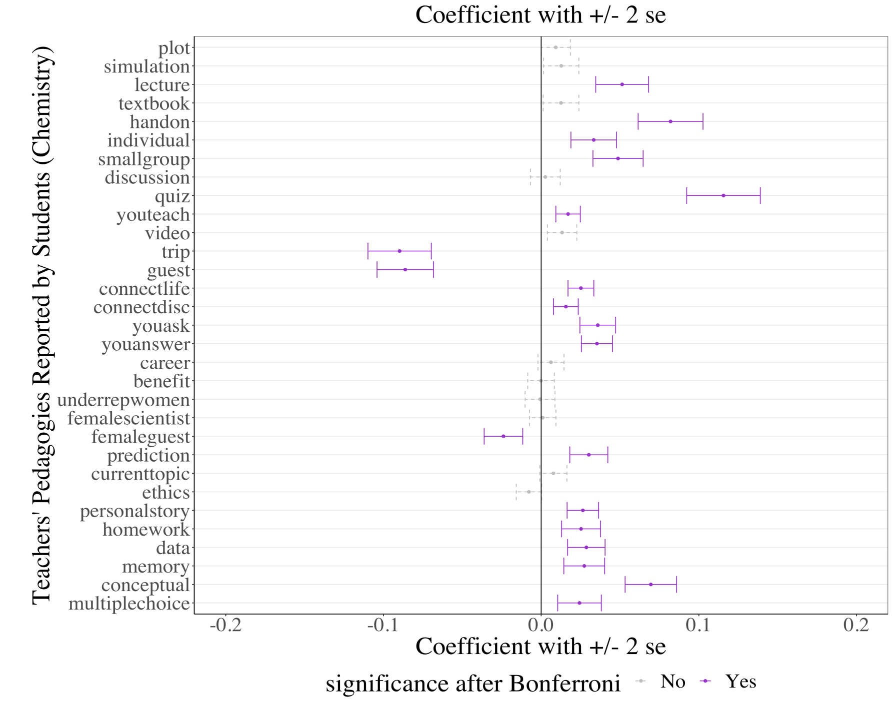
IRT Test Information Curves. Error bars mark ±1SD around HS sample mean. Red curve is Test Information Function (TIF). Blue line is SE of Measurement for an individual subject, dashes for classroom. Black line is estimated SE for classroom of 25 students.

Key Research Findings: Teacher Knowledge and Choice of Pedagogies



1. Predicted student HS physics posttest item correctness as a function of the misconception strength of the item and teachers' Subject Matter knowledge and Knowledge of Student Misconceptions, controlling for student pretest correctness by item.





2. Students gains associated with pedagogical practices. We examined the association between student gains and a. We estimated the effect of variety of classroom pedagogies on students gains while controlling for the background information such as gender and race/ethnicity. Field trips and guest speakers had negative effects, while giving lectures, making connections to life or other disciplines, having students ask or answer questions, and making predictions had positive effects. The most noteworthy difference is that hand-on activities, collecting data, doing homework and memorizing facts had positive effect on students' gain in high school chemistry, but not for physics.