MOSART HS LS

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

DRK-12, 1316645, http://mosart.mspnet.org

Harvard-Smithsonian Center for Astrophysics, Cambridge, MA

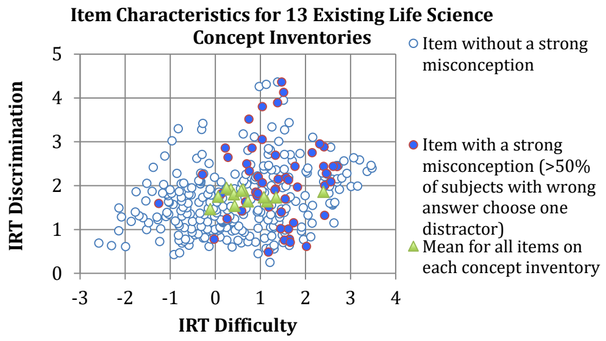
Project Goal

MOSART-HSLS developed assessment tools that will aid in generating rigorous, evidence-based measures of teacher and student understanding of high school-level life science concepts. The project first created multiple-choice items based on grades 9-12 NGSS for Life Science and the published research on misconceptions related to that content. The items pass through several development stages to yield psychometrically strong items that are used to evaluate student and teacher understanding in a nationwide sample of life science courses in U.S. high. In parallel with this effort, the project makes HSLS assessment instruments available on the MOSART Self-Service website for use at no cost by researchers and teachers.

Major Accomplishments to Date

1. Psychometric Analysis – 13 externally developed LS instruments.

Primarily developed for college students, reading level was high and many concepts were not part of the NGSS. Of the 138 items, 12% had weak IRT discrimination (<.65) and 34% were deemed too difficult for high school students. Strong misconceptions (>50% of wrong answers selected a single distractor) were identified for 34% items, making them discussion candidates for new item development.



1. Biology Concept Inventory (Garvin-Doxas *et al.*, 2007)
2. Breathing and Respiration assessment tool (Mann & Treagust, 1998)
3. Conceptual Inventory of Natural Selection (Anderson *et al.*, 2002)
4. Developmental Biology Assessment (Knight & Wood, 2005)
5. Diffusion and Osmosis Diagnostic Test (Odom & Barrow, 1995)
6. Flowering Plant Growth and Development assessment tool (Lin, 2004)
7. Genetics Concept Assessment (Smith *et al.*, 2008)
8. Genetics Literacy Assessment tool (Tsui & Treagust, 2010)
9. Genetics Literacy Assessment Instrument (Bowling *et al.*, 2008)
10. Introductory Molecular Biology Assessment (Shi *et al.*, 2010)
11. Measure of Understanding Macroevolution assessment (Nadelson & Southerland, 2010)
12. Photosynthesis and Respiration in Plants assessment tool (Haslam & Treagust, 1987)
13. Tree Thinking assessment tool (Baum *et al.*, 2005)

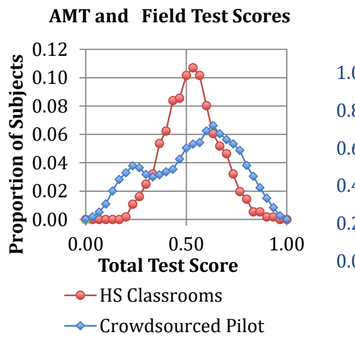
2. Development of 542 draft items

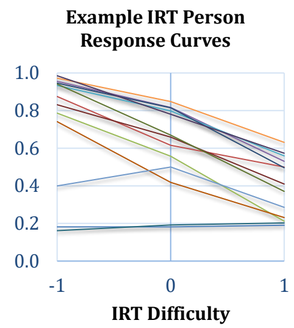
Written for the 15 Disciplinary Core Ideas (DCIs) in the HS LS NGSS.

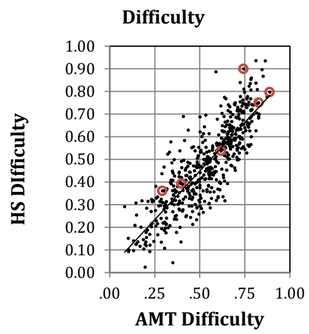
Flowchart of Item Inventory Development Process. We examine whether technological solutions can make the process more seamless and less costly.

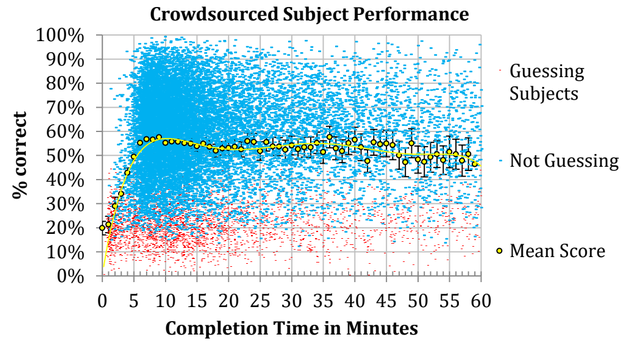
3. Pilot Testing of items using Amazon Mechanical Turk (AMT).

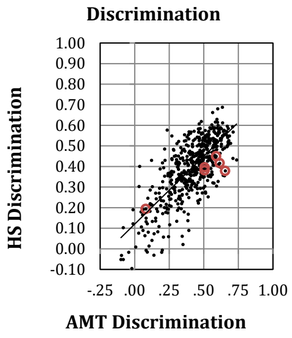
By piloting items using crowdsourcing, we were able to investigate the degree to which this sample of 2,392 subjects could be used to identify anchor items to be used on field tests with HS students, which item variants performed best, and the degree to which AMT subjects' responses match those of HS students. The mean number of tests taken was 8 with 19% taking a single test; 8% took all 18, allowing test forms to be equated and IRT item characteristics to be estimated. The 19% of subjects who primarily guessed at answers were identified by the time taken (<6minutes) and through a person response curve (PRC) for each subject.









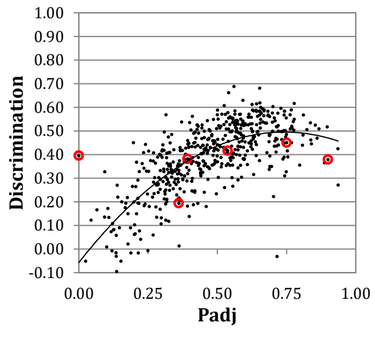


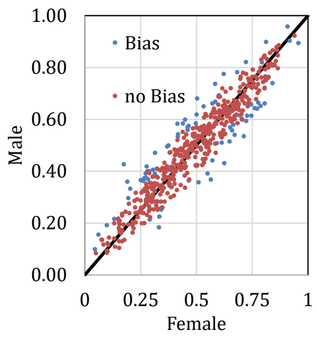
4. Field Testing With High School Life Science Students and Teachers

There is no substitute for administering test items to the target audience, HS LS students, to establish item difficulty, discrimination, misconception strength, and gender bias. Items were administered in sets of 30 with 7 anchors common between the 22 forms to 9,740 students in the classrooms of 187 teachers (averaging 52 students/teacher).

Sample Item: LS1_A_i_J008
The nucleus of a cell
a) is defined by protons and neutrons 9%
b) has a positive charge 2%
c) contains DNA 68%
d) is defined by electrons 3%
e) is located in the center 17%

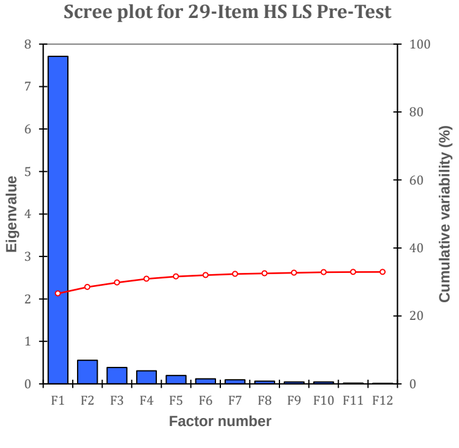
Item Characteristics:
Padj (easiness)=0.68
Discrimination=0.48
Misconception Strength=0.53

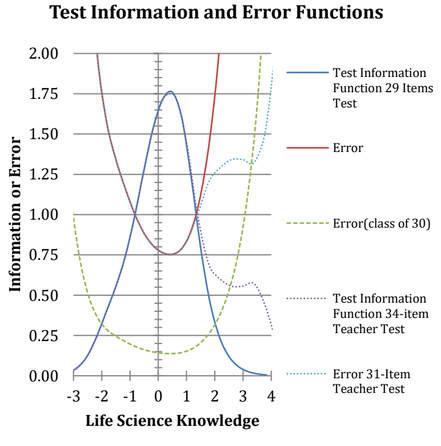




5. Instrument Construction

The resulting item parameters allowed the construction of unidimensional public and secure assessment instrument comprised of 29 items. Also, 6 additional items were added for extending item difficulty for teachers.

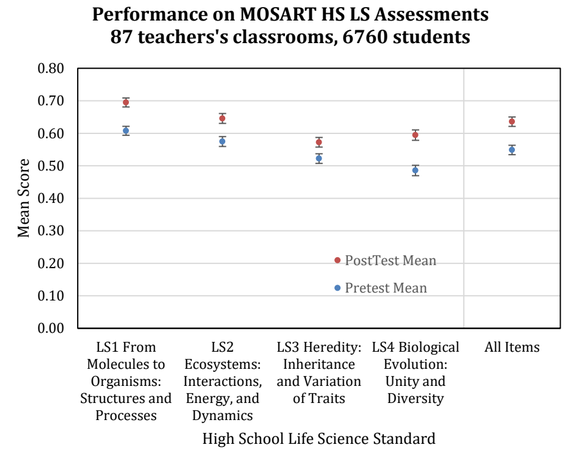




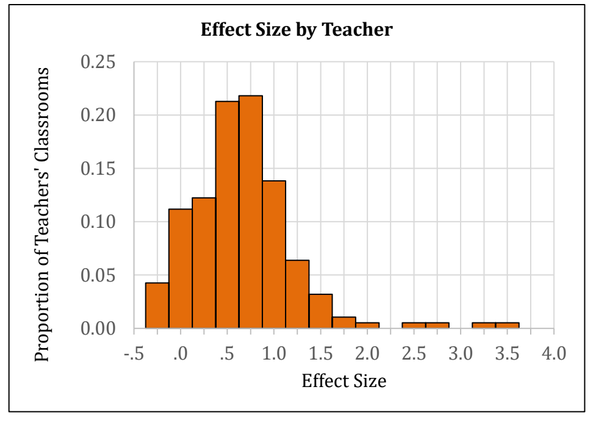
6. HS LS Gain Study

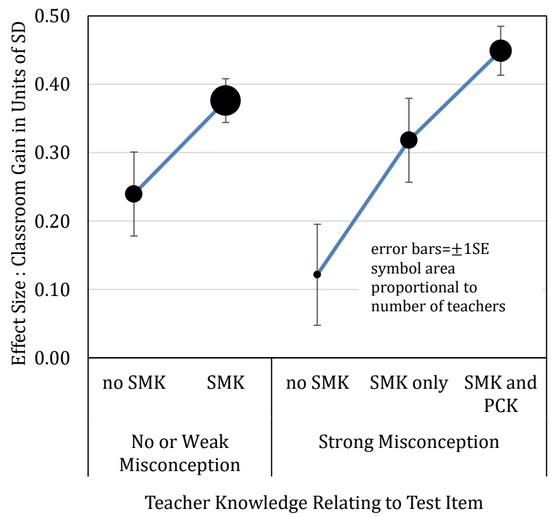
Administering pre-and post-test to high school LS students and their teachers, we were able to generate 6760 student gains for the classrooms of 87 teachers. The effect size (gain in units of SD of the pretest scores) were highest for evolution and lowest for heredity subscores.

		LS1 From Molecules to Organisms: Structures and Processes	LS2 Ecosystems: Interactions, Energy, and Dynamics	LS3 Heredity: Inheritance and Variation of Traits	LS4 Biological Evolution: Unity and Diversity	All Items
Pre-test	Mean	0.61	0.57	0.52	0.49	0.55
	SD	0.12	0.13	0.14	0.13	0.12
	SE	0.01	0.01	0.01	0.01	0.01
Post-test	Mean	0.69	0.65	0.57	0.59	0.64
	SD	0.13	0.14	0.14	0.15	0.13
	SE	0.01	0.02	0.01	0.02	0.01
Effect Size		0.71	0.54	0.36	0.84	0.71
	SE	0.07	0.07	0.06	0.07	0.06



Teachers showed a range in effectiveness as measured by the gains of their students.





Teachers with subject matter knowledge (SMK) and knowledge of student misconceptions (KOSM or PCK-M) had higher gains in their classrooms.

Philip M. Sadler, PI; Harold Coyle, Project Manager; Sue Sunbury, Content Expert; Gerhard Sonnert, Psychometrics; Joel Mintzes, Kimberly Tanner, Content Consultants; John Murray, Online Testing; Annette Trenga, compliance

Any opinions, findings, and conclusions in this article are the authors' and do not necessarily reflect the views of the National Science Foundation.