Abstract Submitted
for the MAR10 Meeting of
The American Physical Society

Do large-scale assessments measure students' ability to integrate scientific knowledge?[1] HEE-SUN LEE, Tufts University — Large-scale assessments are used as means to diagnose the current status of student achievement in science and compare students across schools, states, and countries. For efficiency, multiple-choice items and dichotomously-scored open-ended items are pervasively used in large-scale assessments such as Trends in International Math and Science Study (TIMSS). This study investigated how well these items measure secondary school students' ability to integrate scientific knowledge. This study collected responses of 8400 students to 116 multiple-choice and 84 open-ended items and applied an Item Response Theory analysis based on the Rasch Partial Credit Model. Results indicate that most multiple-choice items and dichotomously-scored open-ended items can be used to determine whether students have normative ideas about science topics, but cannot measure whether students integrate multiple pieces of relevant science ideas. Only when the scoring rubric is redesigned to capture subtle nuances of student open-ended responses, open-ended items become a valid and reliable tool to assess students' knowledge integration ability.

[1]This material is based upon work supported by the National Science Foundation under grants No. ESI-0334199, and ESI-0455877.

Hee-Sun Lee
Tufts University

Date submitted: 20 Nov 2009 Electronic form version 1.4