JRST
Journal of Research in Science Teaching
Research Brief
VOL. 59, NO. 8, PP.1311-1353 (2022)

Assessing High-school Students' Modeling Performance on Newtonian Mechanics

Xiaoming Zhai

Tweet: This study applied Kane's validity framework to assess high-school students' scientific modeling competence in Newtonian mechanics and examined how students' performance depended on tasks.

Audience: Science Educators; Physics Teachers; Assessment Developers

Key Points

- Student modeling competence includes conceptualization, deploying scientific knowledge, and representation.
- Representation is most challenging for students compared to conceptualization and deploying scientific knowledge.
- Students were grouped into four levels: primitive non-model, intuitive model, qualitative model, and quantitative model.
- Significant score variance due to the variability of tasks, maybe due to depth of cognition, knowledge transfer, and construct irrelevant interest.

INTRODUCTION Assessing scientific modeling competence is challenging because of the multi-dimensionality of the construct and the potential variability of student performance across tasks. To deal with the challenges, this study applied Kane's validity framework to assess high-school students' scientific modeling competence in Newtonian mechanics and examined how students' performance depended on tasks. We first specified students' scientific modeling competence in three dimensions: *Conceptualization*, *Deploying Scientific Knowledge*, and *Representation*, incorporating 11 components (e.g., *conceptualization of objects*). We assessed 305 high-school students' modeling performance on two Newtonian mechanics tasks, taking task dependency into account. We applied the Many-facet Rasch measurement to examine students' ability and discriminate students' performance variability across tasks.

FINDINGS Analyses revealed that among the three dimensions of scientific modeling competence, representation is most challenging for students to achieve compared to conceptualization and deploying scientific knowledge, especially among higher-performing students. Five out of the 11 scientific modeling competence components were task-dependent, suggesting significant score variance due to the variability of tasks. After excluding task variance, we found that students were grouped into four levels (i.e., primitive non-model, intuitive model, qualitative model, and quantitative model), and the majority of students were distributed at the medium two levels. We conclude the study by introducing three putative theories that may account for task dependency: *depth of cognition*, *knowledge transfer*, and *construct irrelevant interest*.

TAKEAWAYS The study suggests that task dependency does not necessarily compromise the interpretation of scores, as long as assessors appropriately consider potential theories in their interpretive arguments when designing tasks and interpreting scores.