A Quasi-experimental Study of NGSS Curriculum and PD on Three-dimensional Learning Outcomes
Susan Kowalski, Jeffrey Snowden, Lisa Carey, Betty Stennett, and Lindsey Mohan (BSCS Science Learning)
Heather Young and Catherine Stimac (Oregon Public Broadcasting)

Next Generation Science Standards (NGSS)
- NGSS calls for integration of curriculum, professional development, and assessment.

Curriculum Designed for NGSS
- Focuses on a central phenomenon.
- Includes a coherent storyline from the student perspective (Reiser, 2014).
- Engages students in three dimensions simultaneously:
  - Science and engineering practices
  - Crosscutting concepts
  - Disciplinary core ideas

Online, Video-Based Analysis-of-Practice PD
- 50 hours of online PD
  - 2-hour weekly synchronous sessions
  - 3 hours of asynchronous work weekly in the summer
  - PD spread across 5 months (summer and fall 2018)
  - Teachers analyze their own and others' videos.
- Applied the Science Teachers Learning from Lesson Analysis, or StELLA PD model (Roth et al., 2011; Roth et al., 2019; Taylor et al., 2017).

Distal Three-dimensional Assessment
- Performance Expectation MS-LS1-3: Use argument supported by evidence for how the body is a system of interacting sub-systems composed of groups of cells.

In the unit, students develop a model explaining how body systems work together.

On the assessment, students use their model to and argue from evidence to explain why some athletic students on a hiking trip in the mountains have sore muscles and other athletic students don’t (Harris et al., 2016).

Phenomenon for Unit
- What’s wrong with M’Kenna, and how can symptoms in one part of her body lead to symptoms elsewhere?

Phenomenon for Assessment
- Why do some athletic students have sore muscles after hiking and others don’t?

Students develop model to explain M’Kenna’s symptoms.

Rasch person measures (scaled to 100 points)
- Person reliability = 0.80
- Person separation = 1.98
- Person separation is at low end of acceptability.
- Several items provide redundant information, particularly at low end of scale.

Media-enhanced digital materials for face-to-face classroom
- A Medical Mystery
  - Body Systems Unit for Middle School
  - Phenomenon-based evidence
  - Virtual tools & interactivities
  - Animations

Research Questions
- To what extent does the package of curriculum and PD
  1. enhance teacher instructional practice?
  2. enhance teacher three-dimensional science content knowledge?
  3. enhance student three-dimensional science achievement on a distal assessment?

Research Design
- Comparison group
  - Multiple classes/teacher
  - 1,652 students
  - Business-as-usual body systems unit
  - 2017-2018 school year
- Treatment group
  - Multiple classes/teacher
  - 1,592 students
  - A Medical Mystery body systems unit
  - 2018-2019 school year

Teacher PD
- 30 teachers
- Summer and Fall 2018
- Pretest/posttest

Student PD
- 3 hours of asynchronous work weekly in the summer
- 2-hour weekly synchronous sessions

Analytic Models
- Teacher Model
  - ANCOVA
  - Post scores predicted by pre scores and years of teaching experience
- Student Model
  - Three-level HLM
  - Students nested in classes
  - Classes nested in teacher
  - Random slopes for treatment (average treatment effect across teachers)

Results
- Sizable changes in teacher practice
- P<0.01
- Although intervention did not have large impact, change is significant.

- Modest changes in teacher content knowledge
- P<0.05
- Intervention may have had some positive effect on teacher content knowledge.

- Student impact is greatest on most challenging items
- P<0.001
- Intervention had a significant positive impact on student performance on challenging items.

- Student impact is mainly from the Teacher Model
- P<0.01
- Teacher Beliefs and Knowledge had a significant positive impact on student performance.

Proof of Concept
- This is one of the only studies that uses a quasi-experimental design to test the theory of action outlined in A Framework for K-12 Science Education (NRC, 2012).
- It is possible to “move the needle” on students’ three-dimensional learning with an integration of curriculum, PD, and assessment.

Challenges and Questions
- Effects were relatively small, particularly for students.
- Will a single unit ever show strong changes in three-dimensional learning?
- How will students perform after a full year of NGSS instruction? After multiple years?
- How are other researchers designing units, PD, and assessments? What improvements can be made on this model?

Limitations
- Teacher practice is measured with one video at pretest and one video at posttest. It is not clear if changes in practice are durable.
- Quasi-experimental design does not rule out all possible influences on teachers and students.
- Results may show improvement in outcomes that are to be expected on a year-to-year basis (teachers just getting better naturally over time).
- There may be issues with the sensitivity of the assessment in detecting student impacts.
- Is assessment over-aligned to instruction?
- Are we seeing an opportunity gap?
- Is it not sensitive enough to better reveal impacts on students?
- Assessment with higher person separation would provide greater sensitivity.

Implications
- We provide initial evidence in support of the theory of action outlined in A Framework for K-12 Science Education. More evidence is needed.
- We need additional models of high-quality assessments for NGSS.
- Much research remains:
  - Additional units from other developers
  - Alternative assessment structures
  - Head-to-head comparison of units designed to address the same NGSS performance expectations

Citations
- Reiser, B. (2014). Designing coherent storylines aligned with NGSS for the K-12 classroom. National Science Education Leadership Association, Boston, MA.

This project has been funded through a grant from the National Science Foundation (Award No 1528757).