Improving Teacher Noticing of Students' Science Ideas with a Dashboard

Kelly Billings, Libby Gerard, Marcia C. Linn
kelly_billings@berkeley.edu, libbygerard@berkeley.edu, mclinn@berkeley.edu
Graduate School of Education, University of California, Berkeley

Abstract: We explore how a Teacher Action Planner (TAP) that synthesizes student ideas impacts teacher noticing. The TAP uses Natural Language Processing (NLP) to detect student ideas in written explanations. We compared teacher noticing while using the TAP to noticing when reviewing student explanations. The TAP helped teachers deepen their analysis of student ideas. We did not see any impact on immediate instructional practice. We propose redesigns to the TAP to better connect noticing to instruction.

Keywords: teacher noticing, natural language processing, NGSS, Knowledge Integration.

Introduction

Created in partnership with the Educational Testing Service, the Teacher Action Planner (TAP) uses Natural Language Processing (NLP) to summarize student ideas about science concepts. This summary and the recommended actions based on the student ideas are informed by Knowledge Integration and the NGSS. The TAP was implemented in the WISE Plate Tectonics unit for the first time to support teachers in reflecting on student ideas and planning their instruction (Zhai et al., 2020). This study focuses on investigating how the TAP supports teacher noticing of student's science ideas, as well as how recommendation actions might support teacher planning for distance learning.

The *Knowledge Integration Framework* outlines four processes that support students in developing scientific ideas. They include eliciting initial ideas, discovering new ideas, distinguishing between ideas, and revising their original perspective (Linn & Eylon, 2011). In this study, the TAP pretest report summarizes what ideas students are initially *articulating* at the beginning of the unit, which teachers can use to support their planning. The posttest report summarizes how students *revised* their initial ideas based on what they have learned from the unit and can help teachers to reflect on their teaching and student learning.

Methods

A TAP report was created for the item called Mt. Hood, which calls for students to link ideas about plate interactions and convection currents. We created rubrics to assess each student's explanation along three dimensions: *Knowledge Integration*, the DCI *ESS2.A Earth's Materials and Systems,* and the CCC *Energy and Matter* (Gerard et al, 2020). An NLP model was developed to automatically score student explanations the Mt. Hood Item (Riordan et al, 2020). The resulting model scored student responses with accuracy comparable to human-scorers (quadratic weighted kappa KI=.861, CCC= 0.836, DCI=0.763) and was then embedded into the Mt. Hood item to create a TAP. 'Recommended Actions' in the TAP were designed to help teachers use the summary of Mt. Hood scores to plan instruction and were based on an initial meeting with the lead teacher at the participating school.

We used a cross-case comparison study to examine how teachers used the dashboard. Four 7th grade teachers and their 504 students participated in this study. Two of the 4 teachers had access to TAP reports (Teachers 1 and 2) and two teachers had access to only logged student explanations (Teachers 3 and 4). This was the students' first unit of distance learning for Fall 2020. Teachers used a combination of WISE, Google slides, and Zoom.

Student data included student explanations on the Mt. Hood pre/post item. Each teacher was interviewed twice, after the pretest and again after the posttest. Interviews were analyzed using an emergent coding approach to find themes in teacher's noticing of student thinking, plans to adapt instruction, and their feedback on TAP design. Researchers observed synchronous Zoom classes to see how the teachers used the TAP or student explanations to adapt their instruction. Observation data was examined to discern similarities or differences between the teachers.

Results

The pretest interviews suggest that teachers in the TAP condition noticed more details about their students' ideas than those without the TAP. Teacher 1 noticed that students scored low on both NGSS standards. They wanted to look at the data in more detail, saying "*it would be nice if [the TAP] can figure out student misconceptions or incorrect ideas.*" They wanted more targeted information about individual students to identify which students to assist. Teacher 2 noticed differences in student ideas between their class periods. They noticed that one of their class periods scored low on both NGSS standards and remarked, "*This surprises me, because they talked about it last year."* They also noticed that the majority of students in this same period were struggling "*in that integrating piece.*" In contrast, their other periods displayed a stronger understanding. Both teachers noted that the TAP's recommended actions affirmed their original plans. Teacher 2 said "*[The TAP] gives me a little bit of confidence to do what I was planning to do.*" Teachers without the TAP did not have these insights from examining the student explanations and did not dive deeply into student ideas during their interviews. For instruction, the teachers followed a similar, agreed-upon plan to implement the unit. The alignment among teachers and with their original plan meant there were few opportunities to adapt instruction.

During the post-interview Teacher 1 identified a key idea that their students had trouble with. Students "*were confused about heat transfer and density*" as it applied to convection currents. The teacher elaborated, "*[Students] moved from level 1 to 2 and 3 in DCI. In CCC they mostly just moved to 2 and not a lot of them moved to 3*." The teacher's observation was reflected in the average student scores in the posttest across all classrooms. Teacher 1 wanted to explore this data more closely by looking at individual student data but was unable to due to the current limitations of the TAP. They suggested this as an important area of improvement. Teacher 2 noted their students' growth from pretest to posttest. "*I see a lot more 2s and 3s as opposed to when we first started*." They again noticed the difference in student understanding between periods. "*It's clear my period B is struggling. They may need a little more structure and a lot more discussion time.*" They found this helpful for planning how they can support their different periods in the future. Teachers 3 and 4 did not mention either of these observations when they looked at their posttest student data. Teacher 3 mentioned that they would take a closer look at the data later when they input grades. Teacher 4 said they noticed an improvement but did not look into student explanations in detail.

Conclusions and implications

While previous studies with the TAP focus on teacher implementation of recommended actions (King Chen, 2020), this study brings to light key insights about how teachers use the dashboard to analyze student data. Teachers who had the TAP were able to deepen their understanding of ideas students were still missing and were able to identify differences in performance between class periods. Though informative, the recommended actions provided by the TAP did not prompt teachers to change their instruction. The next steps for refining the dashboard include adding more specific data on individual student NGSS aligned scores to deepen teacher noticing and more closely aligning recommended actions with the student data to help teachers deepen their analysis of specific student's data.

References

Chen, King J. (2020). Impact of autoscored student data reports on teacher instructional customizations for students' science learning. *international conference of the National Association of Research in Science Teaching (NARST),* National Association for Research in Science Teaching, 2020

Gerard, L., Wiley, K., Bradford, A., King Chen, J., Breitbart, J., & Linn, M. C. (2020). Impact of a Teacher Action Planner that Captures Student Ideas on Instructional Customization Decisions. Paper presented at the Annual Meeting of the International Conference of the Learning Sciences, Nashville, TN.

Linn, M. C., & Eylon, B.-S. (2011). *Science Learning and Instruction: Taking Advantage of Technology to Promote Knowledge Integration.* New York: Routledge.

Riordan, B., Cahill, A., Chen, J. K., Wiley, K., Bradford, A., Gerard, L., & Linn, M. C. (2020). Identifying NGSS-Aligned Ideas in Student Science Explanations. Paper presented at the Thirty-Fourth AAAI Conference on Artificial Intelligence, New York, NY. Awarded (AI4EDU)

Zhai, X, C. Haudek, K, Shi, L, H. Nehm, R, Urban-Lurain, M. From substitution to redefinition: A framework of machine learning-based science assessment. J Res Sci Teach. 2020; 57: 1430– 1459. https://doi.org/10.1002/tea.21658

Acknowledgments:

Funding for this research was provided by the National Science Foundation award STRIDES: Supporting Teachers in Responsive Instruction For Developing Expertise in Science (NSF project # 1813713).