International Society of the Learning Sciences

Examining the Effects of Teacher Instructional Approaches on Shifting Students' Experiences Towards Epistemic Practices During Scientific Modeling

Amanda M. Cottone, University of Pennsylvania Graduate School of Education, amandaco@upenn.edu
Susan A. Yoon, University of Pennsylvania Graduate School of Education, yoonsa@upenn.edu
Thomas Richman, University of Pennsylvania Graduate School of Education, trichman@upenn.edu
Clark Chinn, Rutgers University Graduate School of Education, clark.chinn@gse.rutgers.edu
Huma Hussain-Abidi, Rutgers University Graduate School of Education, hh429@rutgers.edu
Noora Noushad, University of Pennsylvania Graduate School of Education, noora@upenn.edu

Abstract: The spread of misinformation regarding socioscientific issues necessitates that science teachers shift focus towards promoting students' use of reliable scientific reasoning strategies and practices (in addition to content knowledge) during instruction. Our goals in this study are twofold: 1) to develop a trustworthy survey instrument that can effectively measure students' experiences engaging in the epistemic practices (EPs) of scientists in the classroom, and 2) to better understand the mechanistic link between instructional approaches and student outcomes. We present a mixed-methods analysis of the variation in classroom experiences that emerged between two teachers with contrasting sets of student outcomes. This study represents one pathway to creating a novel, contextualized survey instrument that can effectively measure the extent to which learners might engage with aspects of EPs in their science classroom. It also further supports student-centered argumentation practices as an effective approach for shifting instruction towards promoting EPs.

Introduction

Given the rampant spread of misinformation regarding important scientific issues (e,g, vaccine safety, climate change), there is a pressing need to promote scientific literacy and, thus, an understanding of scientific practices in classrooms (Gorman & Gorman, 2021; NRC, 2012). This focus foregrounds the development of students' abilities to critically evaluate evidence and distinguish well-justified and accurate information from false and misleading claims (Chinn et al., 2020). It can be contrasted with traditional modes of teaching and evaluation, which have favored emphasizing science content instead of such scientific practices (Windschitl et al., 2012). The shift to bring scientific practices and reasoning skills to the forefront can be framed through research on developing learners' *epistemic cognition*—or the ways of thinking and practices used to establish, critique, and use knowledge within disciplines (Greene et al., 2016). An individual's epistemic cognition informs how they evaluate the reliability of scientific claims and how they come to understand how scientific knowledge is generated. Given this connection, we refer to a set of scientific practices and reasoning skills discussed in this study as examples of the *epistemic practices (EPs)* of scientists.

However, steep challenges exist in promoting EPs during science instruction. First, the link between instructional design and teaching practices that can help develop students' knowledge and use of EPs is not well understood (Muis et al., 2016). Further, little is currently known about how to effectively measure student outcomes in relation to EPs and how they might change over time (Hofer, 2016). As such, teacher educators have been grappling with how to optimize the design of professional development (PD) opportunities needed to help teachers further develop their own knowledge regarding the EPs of scientists, as well as the skills needed to facilitate students' understanding and use of EPs (e.g., Park et al., 2022). This includes, for example, knowing the reliable strategies scientists use for systematically evaluating evidence, justifying why scientific knowledge is reliable, and then engaging students in discussions around these topics to further their understanding.

In this study, we apply the Apt-AIR model of epistemic cognition, which outlines explicit goals for a successful epistemic education (Barzilai & Chinn, 2018), as a framework to inform the design of a curricular intervention in high school biology that we hypothesize, in turn will lead to improved classroom experiences for students. We focus specifically on EPs related to the domain of scientific modeling with complex systems (described in more detail below). Classrooms that incorporate learning environments like this can likely foster students' skill in understanding and enacting reliable EPs to accurately evaluate claims and evidence in their everyday lives (Chinn et al., 2020). We were guided by the following research questions:

1. How can we measure the extent to which students engage in the EPs of scientists during class?

2. What, if any, relationships can be drawn from student classroom experiences regarding their engagement in scientific EPs and what we know about how teachers implemented?

Measuring student experiences related to promoting apt EPs of modeling

Measuring student learning outcomes in science class as they relate to teachers' PD is both difficult and rare (Hofer, 2016). Within the field of epistemic cognition, there is a dearth of studies that use survey instruments with enough statistical rigor (e.g., they incorporate an analysis of model fit and/or the reliability of latent constructs) that reliably capture students' learning of this construct (Sandoval et al., 2016). These instruments are also difficult to validate given the contextualized natures of one's epistemic cognition and thus in the moment observations, as well as metacognitive reflections, are needed to help clarify relationships that might emerge in survey data (Bricker & Bell, 2016).

To address the need to focus on student learning outcomes using robust methodologies, we used elements of the Apt-AIR framework to inform the design of a survey instrument aimed at measuring students' experiences and opportunities in engaging in the EPs of scientific modeling. This framework integrates two models of epistemic thinking: the AIR model (Chinn et al., 2014) and the multifaceted framework of epistemic thinking (Barzilai & Zohar, 2014). The AIR model establishes aims, ideals, and reliable processes as the three main components of epistemic cognition. *Epistemic aims* are the specific goals related to an inquiry process. *Epistemic ideals* comprise the criteria that people use to evaluate whether their aims have been successful. *Reliable processes* denote the strategies that people use to achieve their aims and enact ideals. For the purposes of this study, we used the framing of ideals and reliable processes to help elucidate the specific set of EPs germane to the curricular intervention we were investigating. This is because preliminary coding analysis uncovered infrequent articulation of students' aims. For example, one of the ideals and its associated process relevant to this learning environment included *Rigorous fit with evidence that is systematic and conclusive* (Ideal) and *Considering multiple hypotheses* (Reliable Process used to enact that Ideal). These findings were then used as a basis for item design in the survey.

Furthermore, the multifaceted framework can be used to help design for the varying aspects that promote a comprehensive epistemic education. It states that education should include five key aspects of epistemic performance. These include (*with a survey item example to demonstrate how we attempted to capture each facet*):

i. Cognitive engagement in epistemic performance (e.g., engaging in an educational task in accordance with learning goals); *I often have opportunities to hypothesize and predict scientific results.*

ii. Adapting epistemic performance (e.g., adjusting learning strategies across environments to fit new contexts); *In my science class, my teacher points out how a focus on fitting models to evidence has parallels in science that I can see in the news*.

iii. Regulating and understanding epistemic performance (e.g., considering and reflecting on the purpose and processes of educational tasks); *In my science class, my teacher encourages me to discuss how to use reasoning practices in science out of school.*

iv. Caring about and enjoying epistemic performance (e.g., expressing curiosity, interest and enjoyment while engaged in educational tasks); *I learn science for my own interest.*

v. Participating in epistemic performance together with others (e.g., engaging in collaborative and collective discourse to achieve learning goals). *In my science class, I discuss or share my ideas with people using computer technologies.*

It is important to note that these aspects are not mutually exclusive (i.e., the same item can often fall within multiple aspects). However, using Apt-AIR is theorized to promote learners' abilities to achieve success through competence in epistemic activities (e.g., in the ability to reliably use data to formulate accurate inferences), as well as understanding how to regulate their use of EPs through metacompetence (e.g., in evaluating whether certain inferences can be made given the data available). Thus, we invoked the Apt-AIR model here to measure the wide-ranging aspects of students' experiences related to engaging in what we describe as the *apt EPs of modeling*, which comprise the ideals and processes situated within the context of this intervention. The goal is to demonstrate how this instrument might be used to detect any shifts in students' classroom experience in engaging in apt EPs. Ultimately, constructing a survey like this could be used to assess global classroom experiences, which is important when considering how to scale-up interventions and optimize student impacts. In addition, it would help researchers uncover the links between instructional approach and the degree to which students engage in EPs. To this end, we incorporated findings from qualitative data to elucidate how teachers' instructional approaches might be linked to any shifts that occurred. We now briefly turn to select research on how teachers' instructional approach might promote students' development of epistemic cognition.

Instructional approaches that can promote apt EPs in the science classroom

The call for science instruction to shift emphasis from content to the real practices and reasoning strategies of scientists has been well-established (Duschl, 2008; NRC, 2012), yet the exact approaches teachers can implement to enact this shift are difficult to identify and vary across contexts (Windschitl et al., 2012). One type of generalized approach identified as a critical area for teachers to focus instruction on involves immersing students in argumentative practices, especially in science classrooms as it is through these practices that knowledge advances in the field (Hand et al., 2016). Focusing on argumentative practices requires an emphasis on language through active dialogue, collaboration, and various forms of communication (Hand et al., 2021).

For example, science classroom dialogue should focus students' attention on interpreting data to make evidence-based claims, evaluating opposing claims, and communicating these ideas through verbal, written and graphical forms. To generate knowledge in this way, the classroom environment needs to be student-centered and deemphasize the need to "get" to the right answer (Hand et al., 2016; Muis et al., 2016). In this study, we used qualitative data to understand how teachers' instructional approach when engaging their students in scientific argumentation may have contributed to the quantitative differences in student outcomes that emerged from the survey data. Therefore, this study represents an attempt to draw clear links between the instructional approaches that can lead students to further developing their knowledge and use of EPs during class.

Methods

This study builds upon a larger project that engages high school science students in data collection and analysis through complex systems modeling curricula (Yoon et al., 2017). These curricula consist of five stand-alone units (covering topics spanning biochemistry to ecology and evolution). In this exploratory investigation, the research team collaborated with 8 teachers considered expert in implementing these modeling units to help further theirs (and their students') understanding of apt EPs as they relate to scientific modeling. We also engaged teachers in co-design work in partnering with them to revise the curricula to better promote apt EPs. We present a brief description of the PD below.

The summer workshop portion of the PD occurred over 10 days (with up to 4 hours of synchronous and asynchronous work each day) from August 3-17, 2021 and consisted of introducing teachers to the terms *epistemic cognition* and the AIR model. Next, we explored the differences between cognition and metacognition, the notion of aptness when engaging in EPs, and how to avoid cognitive biases when evaluating evidence. Finally, we discussed instructional strategies for promoting apt EPs and how to extend the EPs of scientific modeling to reasoning about socioscientific issues. In addition, we held four synchronous, 1h meetups with teachers during the school year to analyze the progression of teachers' implementation towards promoting apt EPs. In implementing this sustained PD with teachers, our goal was to improve teachers' understanding and skill in supporting student development of EPs and to increase the infusion of apt EPs into their instruction. In addition, we followed the eight teachers who participated in the PD into their classrooms to understand how their teaching practice might have changed and how that might impact their students. Half of the teachers taught at private, college preparatory schools and the other half taught at large suburban public schools in the north or southeastern U.S. The eight teachers were selected for this study as they represented teacher experts with respect to the modeling units being studied (i.e., they had implemented the curricula multiple times already and helped to facilitate other teachers' uptake of it). These teachers had, on average, 10 years of teaching experience, with a range of 5–18 years at the time the summer workshop occurred.

Participation in student surveys was voluntary, so this study includes survey data from 215 self-selecting students across teachers. Sample sizes of students per teacher ranged from 10 – 52 students and averaged 28 pre- and post-responses per teacher. Despite these sampling limitations teachers indicated that the participating students represented a broad range with respect to performance level. Of students who chose to report their demographics, 56% identified as female, 41% as male, and 3% as non-binary or gender conforming; 1% identified as American Indian Pacific Islander, 17% as Asian, 6% as Black or African American, 2% as Hispanic, 60% as White, non-Hispanic, and 14% as multiple races.

To address the first research question, we created a Likert-scale survey administered to students pre- and post-intervention to measure aspects of their science classroom experience. The survey consisted of 50 statements and was designed to measure latent constructs related to apt EPs of modeling (e.g., engaging in inquiry, problem solving, and scientific reasoning). Students responded to these statements using a five-point agreement rating scale (1=strongly disagree to 5=strongly agree). We employed an exploratory factor analysis (EFA) to reduce the survey dimensions to a lesser number of constructs based on patterns in the respondent data. Construct validity emerged by running the EFA on pre-survey data and retaining items only if they produced factor loadings greater than 0.40 and retaining factors only if 3 or more items clustered under it. We used visual inspection of the parallel analysis scree plot to determine that a solution of 4-6 factors was likely. We then ran three EFAs using a 4-, 5-

and 6-factor solution and compared the outputs. We determined that a 6-factor solution was the most reliable and simplest with Cronbach alpha scores for each factor ranging 0.81-0.92, no items identified as too ambiguous (i.e., clustering significantly under more than one factor), and fewer items clustering under each factor (ranging from 4 to 9) when compared to the other solutions. Eleven questions were identified as too indeterminate (i.e., correlated with none of the factors >0.40) and these were dropped from the item pool. The research team came to consensus on the construct names based on the pattern of items that clustered together (see Table 1). We averaged the responses for all students within a given factor and compared them pre- and post-intervention using paired t-tests. We then calculated Cohen's *d* (interpreted as small-0.2, medium-0.5, and large-0.8) if significant differences from pre- to post-survey emerged.

Table 1
Six Factors and Corresponding Latent Constructs from Student Survey Responses

Factor (Alpha)	Construct	Example Items Comprising the Factor
Factor 1 (0.81)	Engaging in empirical investigations in the science classroom	• I have opportunities to collect data to test and modify hypotheses. • I often hypothesize and predict scientific results. • I often have opportunities to find out answers on my own.
Factor 2 (0.82)	Engaging in problem-solving and communicating using multiple sources with others	• I discuss or share my ideas with people using computer technologies. • I often practice these skills…solving real world problems. • …communicating in multiple ways (e.g., through graphs, email, Internet, web-based tools, writing).
Factor 3 (0.83)	Connecting scientific reasoning in class to everyday life	• In my science class, my teacher encourages me…to discuss how to use reasoning practices in science outside of school. • …to point out how a focus on fitting models to evidence has parallels in science that I can see in the news.
Factor 4 (0.92)	Positive attitudes towards learning science	• I am motivated to learn more science in the future. • I learn science for my own interest.
Factor 5 (0.85)	Using modeling and modeling practices to learn about and do science	• I use computer models to conduct experiments and produce evidence to help me reason and understand scientific ideas. • I use computer simulations, images or animations to collect and analyze data and to draw conclusions.
Factor 6 (0.83)	Using and discussing criteria for good models	• In my science class, my teacher encourages me…to discuss the common characteristics of good complex systems models. • …to discuss why good models should fit all the evidence.

Next, we addressed the second research question by testing whether the overall shifts in students' classroom experience varied by teacher. We ran paired t-tests *by teacher* to investigate any changes pre to post for the factors that showed significant growth in the overall sample paired t-tests described above. Of the 8 teachers, we chose to highlight two for the purposes of this study, where one teacher's (Rachel) students showed negative average growth, whereas the other teacher's (Catherine) students showed significantly positive average growth. In focusing on two teachers with contrasting student outcomes, we aimed to elucidate the teacher implementation strategies that may have led to the differences in student reported outcomes. To do this, we turned to our qualitative data (i.e., video-recorded classroom observations, teacher debriefs and student focus group interviews) collected in each classroom to uncover what differences in instructional approaches may have existed. This in-depth qualitative approach would help us discern if the survey instrument was sensitive enough to detect variation across classroom experiences.

Results

Below we first present any changes in classroom experience regarding students' engagement in the 6 apt EPs measured. Second, we compare the student outcomes of two teachers, Rachel and Catherine, and then use their observational and interview data to better understand the links between teachers' implementation and the differences in students' classroom experience that emerged in the quantitative analysis.

Students show significant growth in three classroom experience factors

Overall, students across the 8 teachers' classrooms made significant gains in three of the six factors measured (see Table 2). Namely, students expressed having more opportunities to *Engage in Empirical Investigations in Science Class*, *Connect Scientific Reasoning in Class to Everyday Life*, and *Use Models and Modeling Practices to Learn* *about and Do Science* after the intervention. Effect sizes ranged from small to medium and suggest that the modeling units were at least partially successful in increasing students' engagement in apt EPs.

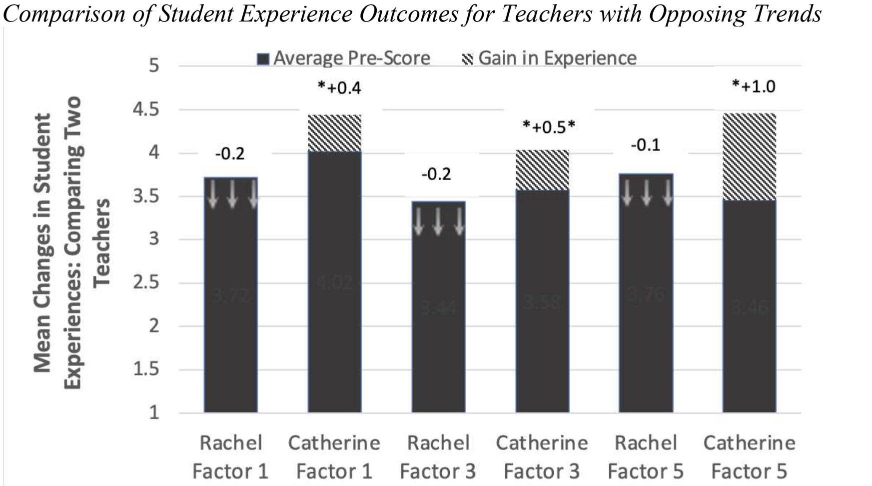
Table 2
Students' Overall Changes in Apt EP Experiences Before and After the Intervention

Factor	Pre Avg (SD)	Post Avg (SD)	Average Gain	Paired t-test results	Cohen's *d*
1-Engaging in empirical investigations in the science classroom	3.86 (0.96)	4.11 (0.90)	+0.24	t = -5.1, df = 214, p<0.0001*	**0.39**
2-Engaging in problem-solving and communicating using multiple sources with others	4.09 (0.88)	4.16 (0.83)	+0.07	t = -1.9, df = 214, p= 0.06	—
3-Connecting scientific reasoning in class to everyday life	3.64 (0.97)	3.85 (0.94)	+0.21	t = -4.0, df = 214, p<0.0001*	**0.28**
4-Positive attitudes towards learning science	3.90 (1.0)	3.91 (0.99)	+0.01	t = -0.13, df = 214, p= 0.89	—
5-Using modeling and modeling practices to learn about and do science	3.68 (0.99)	4.11 (0.86)	+0.43	t = -7.6, df = 214, p<0.0001*	**0.62**
6-Using and discussing criteria for good models	4.16 (0.84)	4.13 (0.81)	-0.03	t = 0.67, df = 214, p=0.50	—

Student classroom experiences vary by teacher

We next evaluated the three significantly different factors from above *by teacher* to uncover what, if any, variation in student outcomes existed across teachers' classrooms. For this smaller study, we present only the results from Rachel's and Catherine's implementation, as analysis at the teacher level showed these teachers' students demonstrated opposing trends in the survey data (see Figure 1 with * denoting the change from pre to post was significantly different). Paired t-test results showed that Rachel's students (n= 24) exhibited no significant measurable change across the three factors (Factor 1 t=1.3, p=0.2; Factor 2 t=1.7, p=0.1; Factor 5 t=0.8, p=0.5); however, Catherine's students (n=21) reported significant growth in all three (Factor 1 t=-2.3, p=0.03; Factor 2 t=-3.6, p=0.002; Factor 5 t=-7.0, p<0.0001). We now turn to qualitative data sources to better understand how teachers' instructional approach may have contributed to the differences in student outcomes that emerged.

Figure 1
Comparison of Student Experience Outcomes for Teachers with Opposing Trends



Selected differences in the instructional approaches of Rachel and Catherine

Here we present select details regarding the differences in how the two teachers implemented the project units to better understand how their instructional approach may have resulted in differing student experiences. However, we want to note that this study is in its early stage, and more complete coding of the full dataset is forthcoming.

The first noteworthy difference between Rachel and Catherine was in the number of units implemented. Rachel implemented only two of five units, and these were clustered towards the beginning of the school year (November and December). She had indicated her intention of implementing a third at the end of the school year

(June) but ultimately ran out of time. In contrast, Catherine implemented all five units and these were spread out over the course of the year (October, November, February, March, and May). This shows that the instructional time spent on curricula designed to emphasize the apt EPs of modeling (i.e., the "dosage" of the intervention) was much greater in Catherine's classroom.

Next, we compared how the two teachers facilitated the same group discussion, which was a Claim-Evidence-Reasoning (CER) prompt embedded within the Sugar Transport model. This model allows students to visualize the processes by which nutrients are transported from the intestines and into the bloodstream. In Rachel's class, students had worked in their groups for about 20 minutes to make observations about how the molecules were moving through the simulation. When facilitating the large group discussion about what mechanism was responsible for the spreading out of molecules, she called on each group in turn, asking "What did [your group] pick?" One student from each group would respond simply with the letter of the claim they chose without providing any evidence or reasoning behind their choice. Group 1 had chosen Claim B, while the other 7 groups called on after them had chosen Claim A. Rachel then exclaimed, "A! It is claim A. The random motion of the molecules is responsible for it spreading out. What evidence do we have of that?" A student from Group 1 discussed their reasoning behind choosing Claim B (i.e., the molecules repel because they "would spin a bit and then fling off"). Rachel seemed to avoid trying to unpack why Claim B did not fit the evidence in that moment when responding, "I guess you can say [that], but what evidence do we have of their random motion?" and shifted students' attention to justifying Claim A instead. She then called on 3 different students, solicited their ideas, listened to their responses, and then summarized back to them the take home message,

> We should have seen that the molecules moved in all different directions, which is random motion…so no particular pattern, it was random, they were moving in different directions. They may have bumped into each other and changed direction, and eventually they spread out evenly. Any questions? [Then Rachel moves on with the next activity in the packet].

In this episode, Rachel encouraged students to articulate the evidence they collected in the model from their various observations and confirmed that Claim A was correct. She geared the discussion towards primarily justifying the correct claim. In total she spent 2 minutes and 25 seconds on the class discussion of the CER.

Catherine, on the other hand, had tried out a new approach to facilitating this CER discussion with her students. In a debrief with her after the class, she mentioned how the PD workshop had impacted her decision to "do things a little a differently" this time.

> I was thinking about what we had talked [at the workshop] when I was planning for this. And just really leaning in to more like…I really feel like I'm guilty of…being like okay we need to get to the right answer. And then once we're there we need to immediately pick back up, or we're not going to finish. Because we're always just stressed for time...So that was me trying to kind of, let's actually take some time on this group discussion. And I modified the student handout so it only had this one group discussion so we could take the time. And I wanted them to think about alternate hypotheses and not just read it and be like, which is most right, and then not even consider the validity of the other two.

In her planning, Catherine anticipated that if she asked the students "what's right? [They're] gonna all say similar things", so she wanted to "explore" the validity of the other two claims a little bit. To do this, she assigned two groups each to Claim A, B and C and posed the question that if the claim they were assigned was correct, what evidence would students see in the simulation to prove that it is correct? She then had students talk in groups for about 5 minutes and told them to come to the board when they were ready to write and share the ideas they discussed. Going through the evidence on the board for each claim, she posed questions to students about whether they saw the evidence described. Students engaged in lively discussions, with one articulating in regard to Claim C (blood flow causes the molecules to spread out) that they "did not have enough evidence yet" from the simulation to determine if it plays a role, but they suspect that it did. Catherine replied, "That's totally fine to say" and validated this student's thinking even though Claim C was technically incorrect. This student was so engaged in the discussion that he stayed after class to work with the model to gather the evidence he needed to be sure.

After various students had the chance to articulate their ideas, she then ultimately asked them to rate each claim as to whether they thought it was the correct one via thumbs up (agree), thumbs side (partially agree), or thumbs down (disagree). She summarized her take home message,

> You always want to be looking for the *why* in biology. It's kind of lazy if we just say, "They just do!" We need to figure out why… It's like pool balls on an empty table in one area, if we shake it, they're going to bounce and spread out into the empty space. And I liked how we talked about cycle, tying back to our complex systems idea, this is not beginning or end, right? There is no end point. It's a cycle, and there would be more bouncing and more spreading out. That's what diffusion is. That's how things move out randomly without energy, it's a very interesting scientific concept we don't think about a lot…So take a moment with your partner, pick what claim you are most convinced by…And then talk about how you want to articulate your reasoning. *Why* does that evidence prove your claim is correct? [Then Catherine has students continue working on the packet]

In this episode, Catherine encourages students to predict what they would see, think about the reasoning behind it, and then decide which claim was correct, without indicating there was only one correct answer. She geared the discussion towards careful evaluation of each claim, making sure she engaged different students in articulating their thinking out loud. In total she spent 28 minutes and 28 seconds on the class discussion of the CER.

In comparing these instructional episodes across the two teachers, we can see elements of Factor 1 (*Engaging in Empirical Investigations)* emphasized in Catherine's practice, where she actively encouraged students to make predictions and find out the information on their own. Catherine's students also exhibited their largest gain in Factor 5 (*Using Modeling and Modeling Practices to Learn About and Do Science)*, and we can see clear demonstrations of this in her emphasis on using models to understand scientific ideas and to engage in scientific reasoning and inferencing. While Factor 3 (*Connecting Scientific Reasoning to Everyday Life)* was not necessarily on display in either episode, Catherine did later describe a fad diet project she implements with students in conjunction with this Sugar Transport model. She gives students a choice of researching 1 of 4 fad diets, all which make claims about how they increase athletic performance and energy. It is possible the connection to fad diets is something students found useful in their everyday lives, and similar to the episode presented here, Catherine tasks students with systematically evaluating whether these claims can be substantiated.

In contrast, Rachel mentioned in the debrief after her class that she's "usually pretty 'notes heavy' with biology just because [she] still feels like [she's] a new bio teacher." Her training was in chemistry-education and her course shift to biology occurred three years prior to this episode (compared to Catherine who was trained in biology and had six years of teaching experience). Rachel's students further elaborated on her practice in their focus group interview that in her biology class they would "write everything down, and that would be the entire class. It was just memorizing facts and definitions." From the comparison of these two episodes and approaches, we see that Catherine displayed a shift towards emphasizing apt EPs whereas Rachel's practice seemed largely focused on emphasizing content; furthermore, these differences were detected in the survey instrument.

Discussion

In this study, we used the Apt-AIR framework to inform the design of PD with high school biology teachers and administered a Likert-scale survey to the students in their classrooms to measure the extent to which teachers enacted shifts towards promoting apt EPs in their instruction. The survey was sensitive enough to detect significant shifts from pre to post in three constructs related to students' classroom experiences (*Engaging in Empirical Investigations, Connecting Scientific Reasoning to Everyday Life,* and *Using Modeling and Modeling Practices to Learn About and Do Science*). This study represents one pathway to creating a novel, contextualized survey instrument that can effectively measure the extent to which learners might engage with at least some aspects of apt EPs in their science classroom. In addition, we lend support to the findings from the survey data with observational and interview data that focused on the instructional approach of two teachers with contrasting student experience trends. This underscores the value in engaging with multiple assessment methods when measuring outcomes related to epistemic cognition (Bricker & Bell, 2016). Ultimately this research can help address the need to examine broad-level student impacts that extend beyond individualized case studies (Hand et al., 2016). Our future work building from this study will include a more systematic analysis of teacher variation and the qualitative data associated with it. For example, more analysis is needed to determine if the survey instrument is sensitive enough to detect differences among the other teachers in the cohort.

This study addresses another gap in the literature in that we aimed to draw relationships from student classroom experiences regarding their engagement in apt EPs and what we know about how teachers implemented. This can help uncover the connections between the instructional approaches that led students to further develop their knowledge and use of EPs (Muis et al., 2016). We can attribute at least some of this shift towards foregrounding EPs in Catherine's instruction to participation in the PD, as she cited it as part of her planning of the CER discussion that engaged her students in many apt EPs (e.g., engaging in multiple hypotheses, systematically evaluating evidence). She also facilitated a rich discussion with her students that prompted them to articulate their reasoning and evaluate opposing claims through various forms of communication (verbal and written), which has previously been identified as an effective approach in promoting EPs (Hand et al., 2021). However, Catherine admitted feeling "guilty" of focusing too much on performance goals (i.e., getting to the "right" answer) due to time pressure in the past. We saw this focus emerge in Rachel's approach, where she emphasized discussing the evidence for the correct claim only in her relatively brief CER discussion example. It is thus important for PD developers to focus more attention on developing teachers' instructional approaches while also considering the intense time constraints that teachers face in their everyday practice (Hand et al., 2021). Overall, this work is important in developing informed citizens that can transfer their use and understanding of EPs to their everyday decision-making regarding current socioscientific issues (Chinn et al., 2020).

References

Barzilai, S., & Chinn, C.A. (2018). On the goals of epistemic education: Promoting apt epistemic performance. *Journal of the Learning Sciences*, *27*(3), 353–389.

Barzilai, S., & Zohar, A. (2014). Reconsidering personal epistemology as metacognition: A multi-faceted approach to the analysis of epistemic thinking. *Educational Psychologist*, *49*(1), 13–35.

Chinn, C.A., Rinehart, R. W., & Buckland, L.A. (2014). Epistemic cognition and evaluating information: Applying the AIR model of epistemic cognition. In Rapp, D., & Braasch, J. (Eds.), *Processing inaccurate information* (pp. 425–454). MIT Press.

Bricker, L.A., & Bell, P. (2016). Exploring images of epistemic cognition across contexts and over time. In J. A. Greene, W. A. Sandoval, and I. Bråten (Eds.). *Handbook of epistemic cognition*. Routledge.

Greene, J. A., Sandoval, W. A., and Bråten, I. (2016). Introduction to epistemic cognition. In J. A. Greene, W. A. Sandoval, and I. Bråten (Eds.). *Handbook of epistemic cognition* (pp. 1–16). Routledge.

Chinn C.A., Barzilai, S. and Duncan, R.G. (2020). Education for a "post-truth" world: New directions for research and practice. *Educational Researcher,* 1–10.

Duschl, R. A. (2008). Science education in 3 part harmony: Balancing conceptual, epistemic and social goals. *Review of Research in Education, 32*, 268 – 291.

Gorman, S. E., and Gorman, J. M. (2021). *Denying to the grave. Why we ignore the facts that will save us*. Oxford University Press.

Hand, B., Cavagnetto, A., Chen, Y. C., & Park, S. (2016). Moving past curricula and strategies: Language and the development of adaptive pedagogy for immersive learning environments. *Research in Science Education*, 46(2), 223–241.

Hand, B., Chen, Y.C., & Suh, J. K. (2021). Does a knowledge generation approach to learning benefit students? A systematic review of research on the science writing heuristic approach. *Educational Psychology Review*, *33*(2), 535.

Hofer, B. K. (2016). Epistemic cognition as a psychological construct: Advancements and challenges. In J. A. Greene, W. A. Sandoval, and I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 19–38). Routledge.

Muis, K.R., Trevors, G., & Chevrier, M. (2016). Epistemic climate for epistemic change. In J. A. Greene, W. A. Sandoval, and I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 331–359). Routledge.

National Research Council (NRC). (2012). *A framework for K–12 science education: Practices, cross-cutting concepts, and core ideas*. Washington, DC: National Academies Press.

Park, S., Kite, V., Suh, J. K., Jung, J., & Rachmatullah, A. (2022). Investigation of the relationships among science teachers' epistemic orientations, epistemic understanding, and implementation of next generation science standards science practices. *Journal of Research in Science Teaching, 59*(4), 561-584.

Sandoval, W. A., Greene, J. A., & Bråten, I. (2016). Understanding and promoting thinking about knowledge: origins, issues, and future directions of research on epistemic cognition. *Review of Research in Education*, *40*, 457–496.

Windschitl, M., Thompson, J., Braaten, M., & Stroupe, D. (2012). Proposing a core set of instructional practices and tools for teachers of science. *Science Education, 96(5),* 878–903.

Yoon, S. A., Anderson, E., Koehler-Yom, J., Evans, C., Park, M., Sheldon, J., Schoenfeld, I., Wendel, D.,
Scheintaub, H., & Klopfer, E. (2017). Teaching about complex systems is no simple matter: Building effective professional development for computer-supported complex systems instruction. *Instructional Science, 45*(1), 99-121.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. 2009803.