Investigating the Development of Data Literacy Through Apt Epistemic Performance With Elementary School Students

Amanda Cottone, Graduate School of Education, University of Pennsylvania, amandaco@upenn.edu
Susan Yoon, Graduate School of Education, University of Pennsylvania, yoonsa@upenn.edu
Jooeun Shim, Graduate School of Education, University of Pennsylvania, jshim@upenn.edu
Bob Coulter, Missouri Botanical Gardens, bob@lrec.net
Stacey Carman, Missouri Botanical Gardens, stacey@lrec.net

Abstract: This research investigates how and in what ways elementary students' epistemic performance (EP) in working with data could be characterized. We used a framework that articulates the cognitive processes students use to evaluate data generated from investigations of a dynamic garden ecosystem simulation. The results illustrate greater abilities in some aspects of EP compared to others. We discuss their implications for teaching practice and future research to promote the development of students' EP.

Introduction and theoretical considerations

As the sheer quantity and availability of data steadily accumulates over time, it is important for educators to focus science learning on adapting to these changes and to promote the data literacy skills needed for informed meaning making of datasets (Jagadish et al., 2014). One way to address this is through an emphasis on teaching science and data literacy that promotes epistemic performance (EP) which involves both the beliefs and practices necessary to establish, critique, and use knowledge within disciplines (Chinn et al., 2014). However, a prominent challenge that remains for educators is how to effectively support the development of students' data literacy skills. Konold et al. (2015) noted that the aggregate perspective of data is more difficult to achieve in elementary school students when compared to middle and high school students. However, the ability to synthesize data and then articulate coherent inferences that are based in, and extend beyond, the data collected are crucial to introduce to students at the elementary-level (Makar & Rubin, 2009). Thus, the focus of this study is to understand the EP processes underlying the successful development of these skills. Here we use the newly derived *Apt-AIR* framework (Barzilai & Chinn, 2018) to analyze the enactment of students' EP.

The framework integrates two independently developed models of epistemic thinking. First is the AIR model which establishes three main components of epistemic cognition as being epistemic *Aims* which refers to the specific goal related to an inquiry process, *Ideals* which comprise an evaluation of the source to achieve new knowledge, and *Reliable processes* which denote the strategies that people use to successfully achieve knowledge aims (Chinn et al., 2014). Second is the multifaceted framework that emphasizes the importance of both cognitive and metacognitive experiences and knowledge in the development of a learner's ability to evaluate knowledge (Barzilai & Zohar, 2016). This framework states five key aspects of EP which are engaging in cognitive processes, adapting EP, regulating and understanding EP, caring about and enjoying EP, and participating in EP with others. When combined, the components of the AIR model can be applied to each of the aspects of the Multifaceted Framework to wholly identify the essential goals needed in epistemic education and can be used as a measure of a learner's "apt epistemic performance." Therefore, our goals for this study are (1) identifying the specific *Apt-AIR* components and demonstrating what these behaviors look like in practice, and (2) exploring ways that *Apt-AIR* can be used to guide curricula designs to promote epistemic growth.

Methodology

This study is part of a larger research program focused on developing curricula and instruction activities to support learning about complex systems with elementary school students conducted from 2016 to 2018. In the curriculum for this study, students used the StarLogo Nova simulation platform to understand the importance of rainfall on a garden ecosystem and how it affected the growth of different plant species. To collect the level of detail on how students enacted EP, we selected four highly engaged students from each class working in pairs giving us a sample population of 24 students (13 females, 11 males). We used 12 video recordings of each pair engaging in approximately 60 minutes of class time to identify aspects of the *Apt-AIR* model (Barzilai & Chinn, 2018). In order to analyze the data, we used a modified method of interaction analysis (Jordan & Henderson, 1995). We first coded the AIR categories and decided to exclude the category of Aims because the aim of the investigation was already provided to students. Then, we examined the videos for evidence of the five aspects of apt EP and excluded the category of adaptive EP which our task did not enable students to demonstrate. Eventually, six codes emerged in terms of the study goals and the coding scheme was vetted with the authors of Barzilai and Chinn (2018) at multiple times during its construction. Table 1 outlines details of the first AIR category.

Table 1: Coding scheme of *Apt-AIR* epistemic performances in the garden ecosystem unit

Category and Definition	Example and Explanation
Reliable Process + Cognitive: Students articulate that they notice trends in the data or assess the accuracy of their previous predictions using math, specific numbers, or more qualitative terms.	Student 2, responding to teacher prompt says, “I notice red, it kind of stayed in the same type of the area. Except for like the first round, it was zero and then went up to 45. I was actually kinda surprised about that” (Pair 10, 10/25/2018). *Student notices and articulates trends in the data by considering previous data runs in their observation of phenomenon.*

Findings and conclusions

Twelve hours of video analysis returned a total of 778 *Apt-AIR* codes (Table 2). We found that the majority of pairs involved cognitive processes (60%) and cognitive ideals (14%), as well as enjoyment/caring processes (18%); whereas, collaborative processes (5%) and metacognitive processes (1%), as well as ideals (1%) were present but to a much lesser degree and did not occur in all the pairs studied. Our findings revealed that students were able to make accurate inferences about plant species' ecology (cognitive ideals) but only to a certain extent.

Table 2: *Apt-AIR* coding frequency analysis

	Reliable + Cognitive	Ideal + Cognitive	Reliable + Enjoyment	Reliable + Collaborative	Reliable + Metacognitive	Ideal + Metacognitive
Mean (SD)	39 (9.8)	9 (7.0)	12 (2.8)	3.7 (3.4)	0.6 (0.8)	0.9 (1.2)
Min-Max	23-52	0-22	8-16	0-10	0-2	0-3
Total Sum	469	109	138	44	7	11

While this category comprised a relatively large frequency of the analysis, the variability in occurrence ranged widely depending on the pair being observed. This is somewhat surprising since using data as evidence to make predictions has been shown to successfully promote inferential reasoning in young students (Makar & Rubin, 2009) and this feature was intentionally built into the curriculum studied here. Still, some student pairs struggled to articulate accurate inferences when engaging with the learning activity as compared to others. One possible explanation for this may lie in the varying perspectives students employed in their cognitive processes as discussed above. In other words, as Konold et al. (2015) suggest, students' perspectives towards interpreting data (whether in aggregate or some other form) ultimately influence the conclusions they draw regarding the data. Thus, the high variability in code types for cognitive processes may translate to the observed high variability in students' ability to articulate cognitive ideals. Whether or not this correlation exists would require a deeper contextual analysis that explored the connections between the types of reliable processes expressed to the resulting cognitive ideals articulated, yet this is beyond the level of analysis presented here.

References

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). Cambridge, MA: MIT Press.

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.

Jordan, B. & Henderson, A. (1995). Interaction analysis: Foundations and practice. *The Journal of the Learning Sciences. 4,* 39–103.

Jagadish, H.V., Gehrke, J., Labrinidis, A., Papakonstantinou, Y., Patel, J.M., Ramakrishnan, R., & Shahabi, C. (2014). Big data and its technical challenges. *Communications of the ACM*, *57*(7), 86–94.

Konold, C., Higgins, T., Russell, S. J., & Khalil, K. (2015). Data seen through different lenses. *Educational Studies in Mathematics*, *88*(3), 305–325.

Makar, K., & Rubin, A. (2009). A framework for thinking about informal statistical inference. *Statistics Education Research Journal, 8*(1), 82–105.

Acknowledgments

This work was funded by a US National Science Foundation ITEST grant (#1513043).