

Exploring Plurality in Students' Ways of Knowing with Learning Progression-based Assessments of Computational Thinking

Beth A. Covitt, beth.covitt@umontana.edu, University of Montana

Carolyn Staudt, cstaudt@concord.org, Concord Consortium

Dale Cope, dalecope54@gmail.com, Independent Education Consultant

Joyce Massicotte, Concord Consortium

Nathan Kimball, nkimball@concord.org, Concord Consortium

Paper Presented at the Annual International Meeting of NARST (Online Conference)
April 8, 2021

This work is supported by NSF STEM +C (DRL-1640088) Integrating Meteorology, Mathematics, and Computational Thinking: Research on Students' Learning and Use of Data, Modeling, and Prediction Practices for Weather Forecasting. Any opinions, findings, and conclusions or recommendations expressed here are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Exploring Plurality in Students' Ways of Knowing with Learning Progression-based Assessments of Computational Thinking

Introduction

Science education's recently growing focus on computational thinking (CT) has occurred contemporaneously with growing focus in other areas of research and practice. These include calls to improve the science learning opportunities of students from historically nondominant communities (e.g., Bang et al., 2012) and critiques of assessment frameworks such as learning progressions (LPs) as defining learning targets in overly narrow ways (e.g., Pierson et al., 2017; Sikorski, 2019). This study lies at the confluence of these foci as they relate to assessing students' CT in a weather prediction context. Within a middle school level project aimed at teaching CT through weather prediction in Alaskan villages, we explored the use of pluralistic, LP-based embedded assessments to meet CT assessment goals in ways that are culturally responsive to Indigenous students and that acknowledge and value not just formal, but also lifeworld (vernacular) ways of sense making (Bang et al., 2018, Gee, 2005).

This exploratory assessment study emerged through the design-based research methodology (Cobb et al., 2003) that guides the *Precipitating Change* project. Initially, *Precipitating Change* project researchers sought to measure students' learning related to CT and weather phenomena separately. This approach yielded moderate success for assessing weather-related disciplinary learning but results from the separate CT practices assessment were less satisfying. Because the CT assessment was unrelated to weather, it did not provide a window into how students were learning CT practices in the context of their weather unit of study (i.e., assessment that is at least 2 if not 3-dimensional). Also, because assessment items were forced choice, they could be used to measure proficiency in relevant CT practices such as interpolation and data abstraction, but the items provided no information about *how* students were making sense of CT or computational modeling.

These issues are not surprising given that much work on assessing CT has focused on measuring whether or not or the extent to which students learn defined CT concepts and practices (Cutumisu et al., 2019). Because it focuses on absence/presence of understanding formal concepts and practices, this approach leads to several problems. For instance, this approach generally does not provide much information concerning students' ways of making sense that can be leveraged for purposes including (1) informing development of instructional design and curricular materials that are responsive to students' ideas, (2) providing opportunities for qualitative formative assessment to inform instruction, and (3) informing professional development that helps teachers understand the intellectual resources students are likely to bring. Also, when learning experiences (including assessment experiences) focus only on formal knowledge and practice, the message sent to students is that lifeworld ways of knowing do not have a valued place in the science classroom (Schwarz et al., 2020).

We report on an exploratory study within our project that responded to these issues through addressing the following design, research, and implications questions:

1. Design: How might assessments be developed to elicit ways students make sense of CT in a weather prediction context?
2. Research: What can plural-design, LP-based embedded assessments help us understand about *how* rural-Alaskan middle school students make sense of CT related to weather prediction?
3. Implications: What lessons are we learning from the design process and students' responses that can inform efforts in assessment, curriculum development, and professional development?

We addressed (design) question 1 in the Design/Procedure section, (research) question 2 in Findings/Analysis, and (implications) question 3 in the Discussion.

Design/Procedure

This study was part of a design-based research project aimed at teaching CT with middle school students from Alaskan Villages through studying weather prediction in a problem-based context. A core objective was to engage students in a relevant and meaningful learning experience that combines real Alaska data with local culture and knowledge of place. Participants included three middle school teachers and their students from three schools in rural Indigenous Alaska communities. In the unit, students adopt the role of local event planners who must decide if weather (and consequently, travel) conditions will allow the Alaskan Native Youth Olympics to proceed as planned. In a series of seven lessons, students interact with authentic weather data in ways that scaffold getting a sense of the predictive work of meteorologists. Students in each classroom work together as a team and apply computational approaches and practices (data abstraction, interpolation, extrapolation, rule abstraction, rule testing, and prediction) to make sense of data (e.g., radar maps and weather station readings), make a weather prediction, and present their decision for whether the event should proceed.

In academic year 2019-20, a set of LP-based assessments were integrated into the *Precipitating Change* unit. The assessments were co-designed by the team, including a new member with experience developing LPs related to Earth systems science and computational thinking and modeling. The assessments were designed using (1) understanding of Alaska Native culture of participating students (e.g., Guillory & Williams, 2014) and (2) insights from previous LP development associated with CT in Earth sciences contexts (Covitt et al., 2020; Gunckel et al., 2018; Podrasky et al., 2019). The team attended to the design question of: ***How might assessments be developed to elicit the ways students make sense of CT in a weather prediction context?*** The following design criteria and project goals were foregrounded.

Assessments should be designed with consideration of need for:

1. Accessibility to students and **plural design** – i.e., ability to elicit both lifeworld and formal ideas depending on how different students make sense of each assessment item.
2. Integration of assessment into the learning unit (i.e., **embedded assessments**) to minimize taking time away from the place and problem-based context of the unit.
3. Assessments to be appealing to and **culturally appropriate** for Alaska Native students.
4. Assessments to **elicit integrated sense making** concerning CT and weather phenomena.

We developed five assessments, each addressing a different CT practice. In each assessment, a fictional middle school teacher poses a weather modeling question and four fictional students offer different responses (Figure 1). Each assessment was embedded at both the beginning and end of the relevant lesson. On the pre-assessment, students discussed the question with peers and chose the answer they agreed with most. On the post-assessment, students discussed the question with peers again, chose a response, and then wrote an individual explanation for why they thought their choice was best. Here, we describe one assessment and the related rationale for and implementation of each design criterion. The full set of five assessments (including accompanying teacher materials) implemented in academic year 2019-2020 is provided in Appendix 1.

How Should We Estimate the Temperature?

Four students in Ms. Tevuk's class are figuring out how to estimate the temperature of a location. The students are looking at the map below and have different ideas for estimating the temperature at the location marked X using the recorded temperatures of 31° and 59°.



Here are the students' ideas:





 Alex We should use the number that is exactly halfway in between 31° and 59°. That's the most exact way to estimate.	 Delana We should mark a number line along the map with the estimate in each square going up the same amount between 31° and 59°. That'd be the best estimate.
 Kalin It doesn't make sense to estimate the temperature at the X. We'd need to go there and measure the temperature. You can't be confident unless you measure it.	 Sage The X is closer to the 59° than to the 31°, so let's choose a number closer to 59°. That's the best we can do because we don't know what's happening with the <u>weather</u> in those places.

Figure 1. *How Should We Estimate the Temperature?*

In Lesson 2, students have incomplete weather data and need to interpolate a value between known values. The *How Should We Estimate the Temperature?* assessment was embedded in the lesson. Responses from fictional students were crafted to represent ideas associated with different levels on a previously developed computational thinking learning progression (Covitt et al., 2020). A teacher guide for each embedded assessment provides short descriptions of theory and background concerning the addressed CT element. The guide also suggests ways to identify students' lifeworld ideas and scaffold students toward developing complementary, formal understanding (Table 1).

Table 1. Excerpt from teacher's guide

	Student response and reasoning	Scaffolds
Less Formal Ideas	Kalin's response (it doesn't make sense to estimate; you need to go there and measure) suggests a very concrete thinking approach consistent with the idea that "you've got to see it to believe it."	Students who choose Kalin's response could benefit from discussions about how while estimates and interpolations may not be perfect, they can still give us a sense of what's going on in a system when we have limited data. It might be helpful to discuss that we can never collect all the data in the world; there's too much! Because of this, we need a way to fill in missing values using the data that we do have available.
Middle Level	Both Alex's response (choose the value halfway between) and Sage's response (choose a number closer to 59) might be reasonable estimates in certain circumstances (e.g., for picking out what clothes to wear if you were at location X at that time). These approaches are both better than just choosing a random value as an estimate for X. However, for contexts in which more precision is called for (e.g., making a weather model), then there are ways that we can be more precise in our approaches to interpolation.	Students who choose Alex's and/or Sage's response could benefit from discussions of when and why we sometimes might need to be more precise in our estimation approaches.
More Formal Idea	Delana's response (mark a number line with the estimate in each square going up the same amount between the two measured values) suggests that she understands which mathematical interpolation approach will give the most precise value.	This level of precision will take more time and effort, and may not even be necessary in some circumstances. However, if our goal is to be precise in our interpolation estimate, then Delana's response is the most appropriate of the options.

This assessment approach allowed us to analyze responses with both a quantitative lens (i.e., examining if students chose more formal responses in the post versus the pre-implementation) and a qualitative lens (i.e., examining how students explain their choices).

Plural-design: We adopted a plural-design criterion in response to critiques of learning progressions as articulating overly narrow, fixed upper anchors as learning targets. Some scholars (e.g., Sikorski, 2019, p. 957) advocate for a, "more expansive 'upper reach' that acknowledges plurality and context-dependence in ways of knowing." In our unit, for example, while linear interpolation is a formal method sometimes used in computational modeling, other interpolation methods can be appropriate as well. One example is *in-between estimation*, which may be sufficient for contexts where less precision is needed. Plural-design can convey to students' that their lifeworld ideas are valued. It can also convey that science learning includes engaging with formal ways of sensemaking, considering how formal and lifeworld sensemaking are similar and different, and considering how each may be appropriate in some partially overlapping set of contexts. Our assessments meet the criterion of plural-design through eliciting and valuing both students' lifeworld ideas and their developing formal ideas.

Embedded assessments: Well-designed embedded assessments can provide evidence of student knowledge and practice, as well as evidence of student learning, with seamless integration of instruction and assessment (Wilson & Sloan, 2000). In other words, embedded assessments can provide student assessment opportunities without pausing or taking time away from instruction for testing. Integrating the embedded assessments met the design goal of providing an assessment opportunity within the instructional context of the unit.

Culturally appropriate: While seeking to avoid pejorative and racist stereotypes concerning Indigenous ways of knowing (McCarty & Watahomigie, 2004), we were interested in designing instruction with sensitivity to learning dispositions developed in the context of family, community, and Indigenous historical perspectives (e.g., of Western schooling). Examples of assessment strategies found to be culturally appropriate with Alaska Native students include

allowing students to help each other, avoiding targeting individuals for public response, and providing opportunities for modes other than paper and pencil – including modeling and demonstration (Trumbull et al., 2015). Following guidelines for culturally appropriate assessment, the *Precipitating Change* embedded assessments were designed such that students had opportunities to talk in pairs or small groups before responding. They did not write explanations on the pre-assessment. The fictional students' explanations offered for participating students to critique provided modeling of explanation. Students were asked for written explanations only on the post-assessment, after relevant learning experiences and peer-to-peer discussion.

Integrated sense-making: Goals for assessment in the NGSS era focus on integration of the three dimensions (disciplinary ideas, practices, and crosscutting concepts), close connection between assessment and classroom instruction, and connections to frameworks such as learning progressions (Pellegrino et al., 2014). We attended to these criteria, for example, through creating assessments that integrate disciplinary ideas (e.g., unequal heating of the Earth), practices (e.g., interpolation), and crosscutting concepts (e.g., systems and system models) (see Figure 1).

Findings and Analysis

In the above portions of this paper, we have discussed our initial response (i.e., during the 2019-20 academic year) to the assessment design question. And, in the context of a school year impacted by the pandemic, we were able to collect pre and post assessment data with students of two of the three participating teachers in Alaskan villages. These data provided initial evidence that the embedded assessments met the design criteria and, consequently, yielded both quantitative and qualitative data concerning how the students were making sense of CT and weather prediction in the context of the learning unit.

Here we provide and discuss results for one assessment (Figure 1) addressing: *What can plural-design, LP-based embedded assessments help us understand about how rural-Alaskan middle school students make sense of CT related to weather prediction?* Analyses for all assessments are complete and results are provided in Appendix 2. Due in part to the pandemic, there were challenges associated with implementing both instruction and data collection during academic year 2019-2020, leading to a less robust data set than we had hoped for. For each assessment, we completed a paired t-test to examine if students' ideas changed in ways consistent with our initial hypothesis concerning which response choices were less/more formal (see Table 1).

For *How should we estimate the temperature?* choosing Kalin was coded as less formal or 1, Sage or Alex as in the middle or 2, and Delana as more formal or 3. For the sample of students for whom we had matched pre/post data (n=41) for this embedded assessment, the averages were pre=2.00 and post=2.46 [$t(40)=3.98$, $p<0.001$]. We observed a positive change ($p=0.03$) in students' pre/post scores in only one of the other four embedded assessments (*Rules for fronts and precipitation*).

Plural (i.e., grouped, not ranked) qualitative analyses of post-assessment written explanations for the *How should we estimate the temperature?* embedded assessment are summarized in Table 2. Analysis categories were developed in an initial round of what will be an iterative cycle of assessment development, implementation, and analysis with combinations of deductive and inductive coding aimed at articulating empirically grounded reasoning categories (National Research Council, 2006; Black, et al., 2011). Categories representing >5% of responses are shown.

Table 2. Students' sensemaking categories from *How should we estimate the temperature?*

Category	Description	Example	%
Nearest neighbor	Response refers to nearest neighbor method.	Alex. Because it is closer to 59 and you have to do nearest neighbor and it is closer to 59.	22%
Agreement / makes sense	Responses suggest "that was best answer" or "I agree."	Sage. Because it's true.	14%
Linear (interpolation) method	Response refers to "linear method" or "linear interpolation method."	Delana. because she is using the linear method, that method works for predicting the weather.	14%
Number line	Response refers to "number line."	Sage because we did it on a number line and it was closer than any other ideas.	12%
Estimation	Response refers to estimation.	I agree with Alex cause he is estimate the number between 31 and 59.	8%
Focus on accuracy	Response suggests choice was more accurate.	Alex. Because it would probably be more accurate.	8%
Can't predict weather	Reasoning about difficulty of predicting weather.	Sage. Because you never know what the weather is doing.	8%

On the post assessment, many students mentioned estimation methods introduced in the lesson including linear interpolation. We also saw other ways to make sense (e.g., weather is unpredictable). Several students explained that you have to measure something directly. Generally, students were able to make sense of and respond to the assessment using either lifeworld or formal science sense making.

Discussion

What lessons are we learning from the design process and students' responses that can inform efforts in assessment, curriculum development, and professional development? We have gleaned useful insights from this work that are informing ongoing design processes in this and related projects. Example learnings that have potential to contribute to teaching and learning of science include the following.

First, we have found that plural design LP-based embedded assessments can meet design criteria and, consequently, yield useful data capturing both lifeworld and formal sensemaking. This approach may inform curriculum development, for example, through highlighting common lifeworld ideas that can serve as starting points for classroom experiences and discourse aimed at bridging between lifeworld and formal knowledge and practice. For example, students'

assessment explanations point to useful questions that can be considered in class: Is weather predictable? Do you have to measure something directly to get a useful measurement? Lifeworld ideas are not necessarily wrong – weather IS NOT completely predictable; weather models have constraints and limitations. Exploratory implementation showed that almost all students made sense of the assessments and provided post explanations reflecting either a lifeworld or a more formal idea. Further, students' lifeworld and formal explanations appear to have the potential to inform instruction.

We also encountered challenges we are considering how to respond to. For instance, across the assessments, a portion of students used what we call “agreement reasoning” in their explanations by providing answers such as, “Sage, because hers is the best answer.” We have discussed this issue and in response, made a revision to the embedded assessments for academic year 2020-21 implementation. In particular, we have changed the wording of the items asking for students' explanations to background which student they agreed with and to foreground the idea they agreed with and the reason they agreed with that idea. We will examine whether the occurrence of agreement reasoning is reduced as a result.

Similarly, we found that relatively few students provided explanations that explicitly drew on Indigenous ways of knowing; a few students referred to traditional activities such as trapping and hunting. Consequently, we decided to change the embedded assessments for 2020-21 in two ways. First, we decided to ask students for not just their choices but also their written explanations on the pre-assessments. We are interested to see if students provide more culturally situated reasons for their choices on the pre-assessments before they have experienced the relevant *Precipitating Change* instruction associated with the focus of each embedded assessments. This change was not a simple one to make given that the initial decision to NOT ask students for written explanations on pre-assessments emerged from a design choice meant to address a cultural consideration for students (i.e., not pressing students to express their initial ideas in writing). Given the fact that few students utilized explicitly cultural ideas on the post-assessment, we decided to make this change (i.e., asking for written explanations on the pre-assessment) to explore whether such a change could elicit additional lifeworld and cultural ideas that could serve as resources for discussion and engagement during classroom implementation.

We also made a second change in the 2020-21 unit design in response to the need to do a better job inviting culturally situated knowledge within the *Precipitating Change* learning experience. The second change to address this issue involved adding a beginning of unit assignment that asks students to interview an adult who has personal experience with predicting weather without relying on weather reports. The students will ask the adult how they learned to predict weather, what they observe when they predict weather, and how they use what they observe to make a forecast. We are interested to see how the inclusion of a home culture activity may bring students' family/community-based and Indigenous ways of thinking forward within the unit instruction.

One other issue we are grappling with is worth discussing here. That issue is the challenge of integrating plurality into learning progression-based assessments in ways that both respect and value students' lifeworld and culture-based ideas while also documenting learning gains with respect to canonical, Western science knowledge and practice. In some ways, where

we have landed so far with the analyses of the embedded assessments is not particularly satisfying. We have used a quantitative approach with students' choices only (associating different choices with different learning progression levels and examining learning gains from pre to post). And, we have used a qualitative approach to examine students' embedded assessment explanations, putting different categories of explanations into different bins, but not designating some categories as more formal or sophisticated than other categories. We do not believe that this breakdown of students' choices into quantitative, hierarchical findings and their explanations into qualitative, nominal findings does a great job achieving the purposes we aimed for in terms of providing strategies for assessment that can integrate cultural responsiveness and formal assessment of canonical ideas.

While we have not implemented it in the current project, we have been working on one strategy for how to address this dilemma that we have begun to develop in the context of a proposal for a next iteration of the *Precipitating Change* project. This strategy draws on several ideas from scholars Gregory Cajete (1999, 2008) and Glen Aikenhead (1997). One approach adapted from Cajete (1999) is the Creative Process Instructional Model, which is a science unit instructional that frames how Indigenous cultural content and Western science can be effectively integrated into units of science instruction. The second idea is a student-as-anthropologist approach, which Cajete describes as, “[students may act as anthropologists learning about another culture. Like cultural anthropologists, they would not need to accept the cultural ways of their “subjects” in order to understand or engage in some of those ways” (Cajete, 2008, p. 492).

We are intrigued by the notion that the student-as-anthropologist approach may provide a strategy for culturally congruent assessment that is welcoming of lifeworld, cultural, and Western science knowledge and practice. For example, within a unit of instruction, there may be times when a message is sent to students that either cultural or Western ideas are welcome. And, it is worth noting that there are areas of commonality between non-Western (e.g., Indigenous) science and Western science ideas and practices where lifeworld and formal science overlap. There may be times during instruction when a message is sent that cultural ideas in particular are being elicited and engaged.

There may also be times when students are asked to express their understanding of Western science ideas in particular. When Western science knowledge and practice are elicited, assessment or other instructional tasks can be framed with a student-as-anthropologist preface. For example, an assessment item might begin with, “If you asked a Western scientist this question, what do you think she would say was a good answer and why?” This approach seeks to place students in a culturally congruent epistemological stance that has the potential to query students' Western science knowledge and practice but that does not require students to abandon or devalue their home culture, or to even necessarily agree with Western science themselves. The student-as-anthropologist approach positions Western science itself as culture, which is consistent with ideas that have been voiced by Western scholars of science (e.g., Latour & Woolgar, 2013). While currently untested by ourselves, it is our conjecture that a student-as-anthropologist approach might provide one useful strategy for culturally congruent assessment of Western science knowledge and practice.

We are interested in continuing to explore ways to design and implement assessment of discipline-embedded computational thinking that addresses multiple aims and concerns including demonstrating cultural congruency and leveraging the potential of learning progressions. It is our intent to continue our iterative assessment and curriculum work in this and future years. With a nod toward Covid-19, however, we note that implementation of revisions to the *Precipitating Change* unit instruction will take place within the context of a very unusual 2020-21 academic year. Students in Alaska villages participating in the project are not attending school in person this year. They also do not have access to Internet. Thus, all of the *Precipitating Change* instructional materials needed to be formatted into a freestanding package that could be loaded onto each student's school laptop and completed by each student individually at home offline.

The *Precipitating Change* unit was originally designed as a highly collaborative and discourse-rich learning experience. This year, students will not be able to discuss the lessons and their associated ideas with other students and with their teacher. Teacher instruction within the unit will be provided via short videos that the teacher has recorded and that are embedded within each lesson that the students will complete individually on their computers. We anticipate that this change will lead to a very different *Precipitating Change* experience for students compared with previous years. While the isolation of students from each other and from their teacher is not ideal, this option was the best outcome that could be attained this year given that participating students do not have access to either in person or remote synchronous learning, or indeed to the Internet.

It is our aim that this study, which is still in progress, will contribute to growing collective understanding of how science education research and implementation efforts can integrate complementary goals of assessing students' CT, improving science learning experiences of students from nondominant communities, and expanding LP theory to consider plurality and context in learning targets.

References

- Aikenhead, G. S. (1997). Toward a First Nations cross-cultural science and technology curriculum. *Science Education*, 81, 217-238.
- Bang, M., Marin, A., & Medin, D. (2018). If Indigenous peoples stand with the sciences, will scientists stand with us? *Daedalus*, 147(2), 148-159.
- Bang, M., Warren, B., Rosebery, A. S., & Medin, D. (2012). Desettling expectations in science education. *Human Development*, 55(5-6), 302-318.
- Black, P., Wilson, M., & Yao, S. Y. (2011). Road maps for learning: A guide to the navigation of learning progressions. *Measurement: Interdisciplinary Research & Perspective*, 9(2-3), 71-123.
- Cajete, G. (1999). *Igniting the Sparkle: An Indigenous Science Education Model*. Kivaki Press.

- Cajete, G. (2008). Seven orientations for the development of Indigenous science education. In N.K. Denzin, Y. S. Lincoln, & L.T. Smith (Eds.), *Handbook of critical and indigenous methodologies* (pp. 487-496). Sage.
- Cobb, P., Confrey, J., DiSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational researcher*, 32(1), 9-13.
- Covitt, B. A., Berkowitz, A., Boone, R., Cooper-Wagoner, J., Gunckel, K. L., Jahnke, M., . . . Moore, J. (March, 2020). *Developing and validating a learning progression for computational thinking in Earth and Environmental Systems Sciences*. Paper presented at NARST, Portland, OR. <https://narst.org/conferences/2020-annual-conference> (Conference cancelled).
- Cutumisu, M., Adams, C., & Lu, C. (2019). A Scoping Review of Empirical Research on Recent Computational Thinking Assessments. *Journal of Science Education and Technology*, 28(6), 651-676.
- Gee, J. P. (2005). Language in the science classroom: Academic social languages as the heart of school-based literacy. In Yerrick, R. K., & Roth, W. M. (Eds.) *Establishing scientific classroom discourse communities: Multiple voices of teaching and learning research* (pp. 28-52). Routledge.
- Guillory, R. M., & Williams, G. L. (2014). Incorporating the culture of American Indian/Alaska Native students into the classroom. *Diaspora, Indigenous, and Minority Education*, 8(3), 155-169.
- Gunckel, K., Cooper-Wagoner, J., Covitt, B., Love, G., Boone, R., Berkowitz, R., & Moore, J. 2018 Students' Ideas About Computational Thinking Concepts When Learning About Modeling Hydrologic Systems. Annual Meeting of NARST: A Worldwide Organization for Improving Science Teaching and Learning Through Research, Atlanta, GA.
- Latour, B., & Woolgar, S. (2013). *Laboratory life: The construction of scientific facts*. Princeton University Press.
- McCarty, T. L., & Watahomigie, L. J. (2004). Language and literacy in American Indian and Alaska Native communities. In *Sociocultural contexts of language and literacy* (pp. 89-120). Routledge.
- National Research Council. (2006). Systems for state science assessment. Committee on Test Design for K-12 Science Achievement. M.R. Wilson and M.W. Bertenthal (Eds.). Board on Testing and Assessment, Center for Education, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.
- Pellegrino, J. W., Wilson, M. R., Koenig, J. A., & Beatty, A. S. (Eds.). (2014). *Developing assessments for the next generation science standards* (Vol. 500). Washington, DC: National Academies Press.

- Pierson, A. E., Clark, D. B., & Sherard, M. K. (2017). Learning progressions in context: Tensions and insights from a semester-long middle school modeling curriculum. *Science Education, 101*(6), 1061-1088.
- Podrasky, A., Covitt, B., Jahnke, M. 2019. High School Students' Developing Ideas About Computational Modeling of Earth and Environmental Systems. Annual Meeting of NARST: A Worldwide Organization for Improving Science Teaching and Learning Through Research, Baltimore, MD.
- Schwarz, C. V., Braaten, M., Haverly, C., & de los Santos, E. X. (2020). Using Sense-Making Moments to Understand How Elementary Teachers' Interactions Expand, Maintain, or Shut Down Sense-making in Science. *Cognition and Instruction, 1*-36.
- Sikorski, T. R. (2019). Context-Dependent "Upper Anchors" for Learning Progressions. *Science & Education, 28*(8), 957-981.
- Trumbull, E., Sexton, U., Nelson-Barber, S., & Johnson, Z. (2015). Assessment practices in schools serving American Indian and Alaska Native students. *Journal of American Indian Education, 54*(3), 5-30.
- Wilson, M., & Sloane, K. (2000). From principles to practice: An embedded assessment system. *Applied measurement in education, 13*(2), 181-208.

EXPLORING PLURALITY IN STUDENTS' WAYS OF KNOWING

Appendix 1: 2019-20 Embedded assessments with accompanying teacher materials



Activity: Lesson 2: Alaska Mainland 2 - Post - What Do You Think?

 1  

Lesson 2: Alaska Mainland 2 - Post - What Do You Think?

How Should We Estimate the Temperature?

Four students in Ms. Tevuk's class are figuring out how to estimate the temperature of a location. The students are looking at the map below and have different ideas for estimating the temperature at the location marked X using the recorded temperatures of 31° and 59° .



Here are the students' ideas:



Alex

We should use the number that is exactly halfway in between 31° and 59° . That's the most exact way to estimate.



Delana

We should mark a number line along the map with the estimate in each square going up the same amount between 31° and 59° . That'd be the best estimate.



Kalin

It doesn't make sense to estimate the temperature at the X. We'd need to go there and measure the temperature. You can't be confident unless you measure it.



Sage

The X is closer to the 59° than to the 31° , so let's choose a number closer to 59° . That's the best we can do because we don't know what's happening with the weather in those places.

Question #1

Which student's idea do you agree with the most?

- Alex
- Delana
- Kalin
- Sage

Question #2

Why do you agree with that student's idea the most? Please explain your reasoning.

**Teacher Tip**

Post-Assessment Implementation

Lesson 2, Conclusion: How Should we Estimate the Temperature? Post-Assessment During the post-assessment, each student will view the probe, make an individual choice, and enter an individual explanation. However, students can work in pairs to discuss their choices and explanations. Student pairs do not have to agree. After students submit their responses, you can lead a discussion and engage the class in talking about some of the issues described below. You should scaffold the class to come to a consensus about which is the best response. Ideally, the students should come to a consensus. If needed, though, you should help the class understand why some responses are better than others.



Theory & Background

Computational Thinking Concept and Practice:

- Interpolation

What is the Purpose of this Assessment Probe?

The following information provides background for **teachers** and is not necessarily representative of the learning goals for everything students should understand. Review the information and consider ways to scaffold students toward developing knowledge and practice as appropriate for their classes. We also recommend during the first implementation of the probe, you scaffold students to focus on student ideas about this problem, rather than on vocabulary terms such as **interpolation**. By discussing the different options, students can engage in productive talk related to data sense making and computational thinking without using technical terminology. During the ensuing lesson and the second implementation of the probe, the term interpolation can be used, as appropriate.

This assessment probe is designed to assess and scaffold students in thinking and talking about **interpolation**. **Interpolation** means estimating a value between ones that are known or tabulated using surrounding points or values (that is, estimating an unknown value based on surrounding information).

The assessment is situated within the unit context to provide an opportunity for students to make sense about interpolation of a specific, tangible weather question (i.e., temperatures along a map transect). There is always uncertainty involved when interpolating across data. Also, depending on circumstances, different methods may be appropriate. For example, if the students were trying to figure out how to dress for the day at location X, then simply choosing a number somewhere between the two might be sufficient. However, for developing a computer model, the students would likely want to choose a more robust approach. Using an indefinite problem provides space for students to think and talk about things like when and why it might make sense to use the different interpolation approaches.



Discussion Points

What Are Common Student Ideas and What Supports Could Help Students?

Below are some issues to consider in assessing students' ideas, and some suggestions of ways to support students in developing more formal reasoning.

Less Formal Ideas

- **Kalin's response (It doesn't make sense to estimate; you need to go there and measure.)** suggests a very concrete thinking approach consistent with the idea that "you've got to see it to believe it."

Students who choose Kalin's response could benefit from discussions about how while estimates and interpolations may not be perfect, they can still give us a sense of what's going on in a system when we have limited data. It might be helpful to discuss that we can never collect all the data in the world; there's too much! Because of this, we need a way to fill in missing values using the data that we do have available.

Middle Level

- **Both Alex's response (choose the value half way between) and Sage's response (choose a number closer to 59)** might be reasonable estimates in certain circumstances (e.g., for picking out what clothes to wear if you were at location X at that time). These approaches are both better than just choosing a random value as an estimate for X. However, for contexts in which more precision is called for (e.g., making a weather model), then there are ways that we can be more precise in our approaches to interpolation.

Students who choose Alex's and/or Sage's response could benefit from discussions of when and why we sometimes might need to be more precise in our estimation approaches.

More Formal Idea

- **Delana's response (mark a number line with the estimate in each square going up the same amount between the two measured values)** suggests that she understands which mathematical interpolation approach will give the most precise value. Again, this level of precision will take more time and effort, and may not even be necessary in some circumstances, as discussed above. However, if our goal is to be precise in our interpolation estimate, then Delana's response is the most appropriate of the options.

Generate a report



Precipitating Change



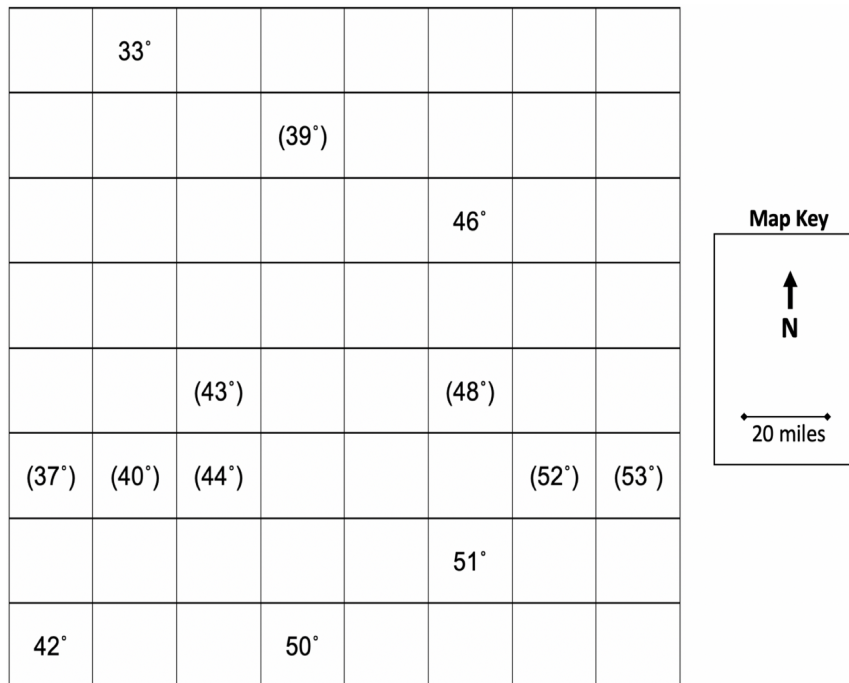
Activity: Lesson 3: Alaska Mainland 2 - Post - What Do You Think?



Lesson 3: Alaska Mainland 2 - Post - What Do You Think?

Which Estimates Should We Trust?

Ms. Tevuk's class is looking at a temperature map with five data points taken from weather station observations. All five weather station observations were made at the same time on the same day. The map also shows some temperatures that were estimated using the weather station observation data (estimated temperatures are in parentheses).



Ms. Tevuk asks the students which of the estimated temperatures they are LEAST CERTAIN (MOST UNSURE) about.

Here are the students' ideas. (Each student's estimated temperature choice is in **bold** to make it easier to find.)

Alex



I'm least certain about the estimate of **53°** because it's furthest outside of the temperatures that were actually measured at a weather station.

	33°						
			(39°)				

Delana



I'm least certain about the estimate of **43°** because it's far away from any temperatures measured at a weather station.

	33°						
			(39°)				
					46°		

				46°	
	(43°)		(48°)		
(37°)	(40°)	(44°)		(52°)	(53°)
			51°		
42°		50°			

Kalin



I'm least certain about the estimate of 39° because there aren't any other estimated temperatures near the 39° estimate.

	33°				
		(39°)			
			46°		
	(43°)		(48°)		
(37°)	(40°)	(44°)		(52°)	(53°)
			51°		
42°		50°			

		(43°)		(48°)	
(37°)	(40°)	(44°)		(52°)	(53°)
				51°	
42°		50°			

Sage



I'm least certain about the estimate of 37° because weather usually moves from west to east, so I'm unsure about estimates on the west side of the map.

	33°				
		(39°)			
			46°		
	(43°)		(48°)		
(37°)	(40°)	(44°)		(52°)	(53°)
			51°		
42°		50°			

Question #1

Which student's idea do you agree with the most?

- Alex
- Delana
- Kalin
- Sage

Question #2

Why do you agree with that student's idea the most? Please explain your reasoning.



Teacher Tip

Post-Assessment Implementation

Lesson 3, Conclusion: Which Estimates Should We Trust? Post-Assessment During the post-assessment, each student will view the probe, make an individual choice, and enter an individual explanation. However, students can work in pairs to discuss their choices and explanations. Student pairs do not have to agree. After students submit their responses, you can lead a discussion and engage the class in talking about some of the issues described below. You should lead the class to come to consensus about which is the best response. If needed you should help the class understand why some responses are better than others.



Theory & Background

Computational Thinking Concepts and Practices:

- Interpolation
- Extrapolation
- Pattern Recognition
- Data Aggregation

What is the Purpose of this Assessment Probe?

The following information provides background for teachers and is not necessarily representative of the learning goals for everything students should understand. Review the information below and think about ways to scaffold students toward developing knowledge and practice as appropriate for their classes. Also, we recommend that during the first implementation of the probe, that you scaffold students to focus on their own ideas about this problem, rather than on vocabulary terms such as **interpolation** and **extrapolation**. By discussing the different options, students can engage in productive talk related to these data sense making and computational thinking concepts without necessarily using technical terminology. During the ensuing lesson and the second implementation of the probe, the terms **interpolation** and **extrapolation** can be used, if appropriate.

This assessment probe is designed to assess and scaffold students in thinking and talking about **interpolation** and **extrapolation**. **Interpolation** means finding an estimation of a value between ones that are known or tabulated using surrounding points or values (that is, estimating an unknown value based on surrounding information). **Extrapolation** means finding or concluding something by assuming that existing trends will continue, or a current method will remain applicable beyond existing known values (that is, following a pattern beyond the known data points).

Interpolation and extrapolation are similar, but because extrapolation involves estimating beyond the known data points, this approach is subject to greater uncertainty and has a higher risk of producing more imprecise or potentially even meaningless results compared with interpolation.

The assessment is situated in the unit context to provide an opportunity for students to make sense about extrapolation and interpolation of a specific, tangible weather problem (i.e., temperatures in a two-dimensional area at one point in time). While extrapolation is associated with higher levels of uncertainty, it's also sometimes necessary. This is true, for example, for forecasting (prediction). Prediction of future weather will always involve extrapolation (i.e., extending the pattern of weather data into the future given that observed data cannot be obtained for the future until time passes and the future

becomes the present!). Uncertainty associated with extrapolation is why there's always some level of uncertainty in weather forecasts. Basically, this is why weather forecasts are sometimes wrong. There is always some uncertainty associated with interpolation too, though generally it is less so than with extrapolation.

This assessment also relates to **pattern recognition** and **data aggregation**. **Pattern recognition** involves identifying patterns and trends within and across groups of data/information as seen in the observable world. Students employ **pattern recognition** in this assessment as they view the distribution of temperatures in the map grid and make sense of patterns and trends across the grid (in both observed and estimated values) to think about the question. **Data aggregation** involves determining the appropriate data to collect or generate in order to study the identified phenomenon. Students should consider more and less appropriate ways to aggregate data in the map grid as they figure out which response they think is best. For example, students should think about which cells in the grid have enough data to estimate temperatures for from collected data, and which temperatures might be problematic to estimate – meaning we might need to collect more data to generate a good estimate for temperatures in those cells.

While one of the responses makes the most sense from a practical standpoint for this particular problem, there's no absolutely correct answer. Using an indefinite problem provides space for students to think and talk about how much confidence we can have in different interpolated and extrapolated values.



Discussion Points

What Are Common Student Ideas and What Supports Could Help Students?

Below are some issues to consider in assessing students' ideas, and some suggestions of ways to support students in developing more formal reasoning.

Less Formal Ideas

- **Kalin's response (39°)** suggests that he may not be differentiating between observed and estimated values. The values in parentheses are all estimated, so having other estimated values near the value that Kalin chooses would not necessarily make us more confident about Kalin's choice. If there were additional observed values near the cell that Kalin chose, that would help boost our confidence. We can feel more confident about interpolated values for data the more observed data values we have for a given area.

Students who choose 39° as the most uncertain value could benefit from discussions about the difference between observed data points and estimated data points. Students can think about which types of values are more trustworthy and why (e.g., local conditions like elevation changes or water bodies could affect local values, making estimated temperatures less reliable than observed values). Students may come up with potential problems with observed values too (e.g., a faulty thermometer), which is fine.

Middle Level

- **Sage's response (37°)** suggests that she may be thinking beyond interpolation to also consider weather processes (e.g., movement of air masses across Alaska). While that shows some good thinking, the values on the assessment probe map all represent one point in time, so we would be less concerned about air movement over time and issues like east versus west in Alaska for this particular problem.

Later in the unit, students who chose 37° may benefit from discussions that emphasize the difference between interpolating data from known values at one point in time versus doing things like making predictions into the future across a two-dimensional area based on known information such as wind speed, wind direction, etc.

Upper Middle Level

- **Delana's response (43°)** suggests that she understands that we can be more confident about estimates when we have

• **Delana's response (43°)** suggests that she understands that we can be more confident about estimates when we have observed values nearby. We should be less confident about interpolated data (there are many estimated values near the 43° cell) nearby in comparison with nearby observed data.

While students who choose Delana's response have useful ideas about confidence and interpolation, they could also benefit from some scaffolding to reason more deeply about interpolation versus extrapolation. Moving beyond the area of known data points (i.e., extrapolation) is generally associated with higher uncertainty compared with interpolation. For example, if there is a weather front just to the east of the 46° and 51° measured temperature values, the 53° estimated value could be quite different from what an actual observed measurement taken at that location would be.

More Formal Idea

• **Alex's response (53°)** and rationale recognizes the problems associated with extrapolation beyond known data points. As described in the discussion of Delana's response above, if there were a weather front just to the east of the 46° and 51° observed temperature values, the 53° estimated value could be quite different from what an actual observed measurement taken at that location would be. Data sense making about interpolation and extrapolation involves understanding the differences between these two estimation approaches and, in particular, the generally higher level of uncertainty associated with extrapolation versus interpolation.

Generate a report



Precipitating Change



Activity: Lesson 5: Alaska Mainland 2 - Post - What Do You Think?



Lesson 5: Alaska Mainland 2 - Post - What Do You Think?

How Should We Break It Down?

Ms. Tevuk's class is studying weather. The students will make a computer model to predict weather in Alaska. They need to break Alaska into smaller chunks of area so their model can use information (data) about each chunk to predict the weather. Four students have different ideas about what size the chunks should be. Alaska is approximately 663,000 square miles. The map below shows Alaska with a grid that divides the map into squares.



Here are the students' ideas:

Alex

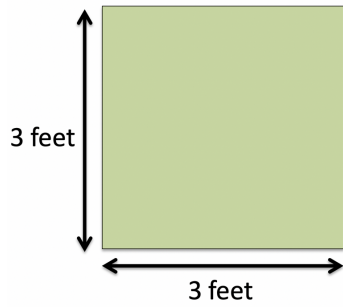


We need to include as much data as possible so our model will be accurate. Let's break Alaska down into chunks that are each one square yard in size (three feet on each side).

Delana



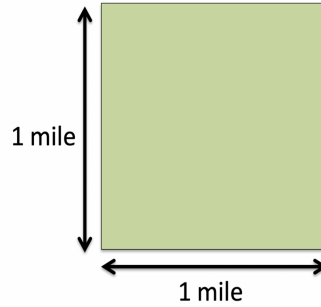
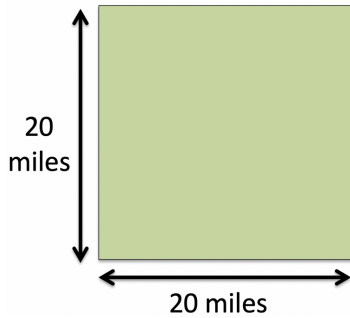
We want to predict weather for villages, so our chunks should be about the size of an Alaska village. Let's make chunks that are each a square mile (5,280 feet on each side).



Kalin



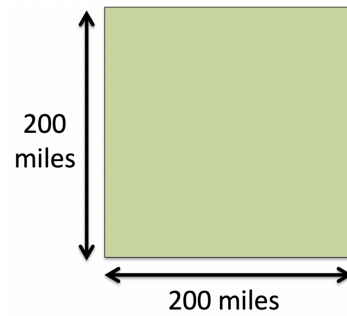
The chunks can be even bigger because there's space in between villages. Let's break Alaska down into chunks that are each 20 miles by 20 miles in size.



Sage



If we use 20 miles by 20 miles, we'll have over 1,500 chunks for Alaska. That's a lot. Let's use 200 miles by 200 miles so we'll only need about 16 chunks.



Question #1

Which student's idea do you agree with the most?

- Alex
- Delana
- Kalin
- Sage

Question #2

Why do you agree with that student's idea the most? Please explain your reasoning.



Teacher Tip

Post-Assessment Implementation

Lesson 5, Conclusion: How Should We Break It Down? Post-Assessment During the post-assessment, each student will view the scenario again and make a new individual choice, and enter an individual explanation. However, students can work in pairs to discuss their choices and explanations. Student pairs do not have to agree. After students submit their responses, you can lead a discussion and engage the class in talking about some of the issues described below. You should work to lead the class to come to a consensus about which is the best response. Ideally, the students should come to a consensus. If needed, though, you should help the class understand why some responses are better than others.



Theory & Background

Computational Thinking Concepts and Practices:

- Decomposition
- Data Abstraction

What is the Purpose of this Assessment Probe?

The following information provides background for **teachers** and is not necessarily representative of the learning goals for everything students should understand. We suggest you review the information below and think about ways to scaffold students toward developing **knowledge and practice** as appropriate for their classes. We also recommend emphasizing the underlying ideas and practices, rather than on vocabulary terms such as **decomposition** or **data abstraction**. By discussing the different options, students can engage in productive talk related to these computational thinking concepts without using technical terminology.

This assessment probe is designed to assess and scaffold students in thinking and talking about **decomposition**, which is *breaking down data, processes, or problems into smaller components and manageable parts so they can be more easily solved*.

In particular, the probe addresses physical decomposition of a system, which involves *dividing a continuous representation of a system into smaller, discrete parts*. In computational modeling, *physical decomposition makes it possible to apply values to various parameters in each small part (cell) of the model domain, which in turn makes it possible to apply mathematical approaches to the problem*.

At the broader level, this type of discretization is an example of the computational thinking concept of **data abstraction**. Abstraction *involves letting one object stand for many by “capturing essential properties common to a set of objects while hiding irrelevant distinctions among them”* (Wing, 2014, <http://socialissues.cs.toronto.edu/index.html%3Fp=279.html>). When we divide the Alaska map into cells, we are stripping away a lot of information about each cell (e.g., how many trees or houses are in that area) so that we can focus on just the pieces of information we need for addressing our particular problem (in this case, we will want to understand some other things about each cell at certain points in time to help us predict the weather, i.e., the temperature of the air, the wind speed and direction, the humidity level in each cell).

This assessment is situated within the unit context to provide an opportunity for students to make sense about decomposition of a specific, tangible weather problem. While one of the responses makes the most sense from a practical standpoint for this problem, there's no absolutely correct answer. Using an indefinite problem provides space for students to think and talk about the pros and cons of using different sized chunks for predicting weather. This is the kind of real-world problem that weather modelers need to think about as they do their work.



Discussion Points

What Are Common Student Ideas and What Supports Could Help Students?

Below are some issues to consider in assessing students' ideas, and some suggestions of ways to support students in developing more formal reasoning about decomposition.

Less Formal Ideas

- **Alex's response (3 feet by 3 feet chunks)** is probably too small for this problem. A model decomposed into cells of this size over the state of Alaska would require a very large amount of interpolation because, of course, there are not that many weather stations available to provide observed data for each cell (although satellites could potentially collect this data over land surfaces). Assuming data come from weather stations though, a model composed of one square yard sized cells would provide a false level of accuracy based on an insufficient amount of data for understanding weather at that scale. In addition, this model would require a great deal of computing power. It would take a lot of time and energy for the model to process all the data, and because so much of the data will need to be interpolated, the results will not really be trustworthy at the scale of one square yard.

Some students might find Alex's response appealing because they think it is important to be very accurate with a computer model and/or because they think computer models have the capacity to make very accurate predictions. These students could use support in understanding that computer model outputs are based on computer model inputs. With relatively few data inputs from weather stations compared to the outputs (predictions at the level of each square yard. Alex's model will not be able to perform very well.

- **Sage's response (200 miles by 200 miles chunks)** will not provide enough detail for Alaskans to make the decisions they need to make related to the weather. The map below shows an area of about 200 by 200 miles. If this area were used for the weather model, then it would not be possible to give people in Port Lay a separate weather prediction than people in Utqiagvik. While Alex and Delana's responses have problems with false accuracy, Sage's response has problems with insufficient precision for the real-world problem (predicting the weather at a scale that will be useful for people in different villages and towns).

Students who choose Sage's response may think there is a need to enter actual data (i.e., from a weather station) for every cell in a computer model. These students could benefit from discussions that connect this assessment probe to the idea of interpolation – we don't necessarily need to enter data for every single cell in a computer model. Much of the data can be interpolated. If we didn't have the capacity to interpolate data, then Sage would be right, - it would be a whole lot of work to make a computer model for Alaska, even at the scale of 20 by 20 mile cells. Interpolation helps solve this problem (within limits of course, because if we interpolate too much data we end up with the problems associated with Alex's and Delana's models).



Approximately 200x200 mile "chunk" of Alaska.

Middle Level

- **Delana's response (1 mile by 1 mile chunks)** is also probably too small for the problem. There would be approximately 663,000 cells in this model. Because of the number of cells, a lot of the data for this model would need to be interpolated. There is not necessarily a need for weather information at this fine-grained scale. While less so than Alex's response, this model would still have problems with too much interpolation leading to a false level of accuracy at the square mile scale and an excessive amount of computing power needed.

Delana's response shows that she's doing important thinking about the scale of information that is needed by the audience for the model (i.e., people living in villages and towns). Students often think about computer models and their outputs in terms of what can be shown or seen. While visualizations are an important output of computer models, students can also use support in developing deeper understandings about the need for sufficient data and appropriate processing algorithms for a computer model to show something. Students often think about computer models as black boxes that do something

for a computer model to show something. Students often think about computer models as black boxes that do something mysterious and then provide an answer. Over the course of the unit, students should develop deeper understanding of how data and rules (algorithms) are the guts of the computer model that help modelers use the model to explain and predict things like weather.

More Formal Ideas

- **Kalin's answer (20 miles by 20 miles chunks)** provides the ability to process data at a reasonable scale for an area as large as the state of Alaska. There would be approximately 1650 cells in this model. If the chunks were much bigger, it would not be possible to process and provide information at a small enough scale for what is needed in terms of weather information for individual villages, towns, and cities. If the chunks were much smaller, they would be subject to the same problems associated with Alex's and Delana's responses.

Generate a report



Precipitating Change



Activity: Lesson 4: Alaska Mainland 2 Post - What Do You Think?



Lesson 4: Alaska Mainland 2 Post - What Do You Think?

What Are the Rules for Fronts and Precipitation?

Ms. Tevuk's students need to write a rule for their computer model to predict areas of precipitation on a map showing weather fronts. They have different ideas for what the rule should be.

Here are the students' ideas: The rule should be that precipitation happens...

Alex



with a warm front coming in because warm air moves in and mixes with cold air, which leads to precipitation. You don't get precipitation with a cold front because cold air is dry.

Kalin



with a cold front coming in because it's usually rainy or snowy when it's cold out. You don't get precipitation with a warm front because it's warm when the sun is shining.

Delana



with both cold fronts and warm fronts coming in because in both cases warm air and cold air meet, leading to condensation and precipitation.

Sage



after a cold front has passed by because the cold air the front leaves behind is humid, which leads to precipitation.

Question #1

Which student's idea do you agree with the most?

- Alex
- Delana
- Kalin
- Sage

Question #2

Why do you agree with that student's idea the most? Please explain your reasoning.

**Teacher Tip****Post-Assessment Implementation**

Lesson 4, Conclusion: What Are The Rules For Fronts and Precipitation? Post-Assessment During the post-assessment, each student will view the probe, make an individual choice, and enter an individual explanation. However, students can work in pairs to discuss their choices and explanations. Student pairs do not have to agree. After students submit their responses, lead a discussion and engage the class in talking about some of the issues described below. Try to get the class to come to consensus about which is the best response. If needed, help the class understand why some responses are better than others.

**Theory & Background****Computational Thinking Concepts and Practices:**

- Pattern Recognition
- Data-based Prediction
- Rule Abstraction

What is the Purpose of this Assessment Probe?

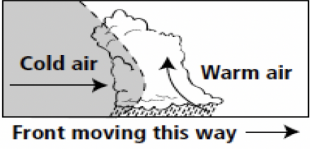
The following information provides background for teachers. We suggest teachers review the information below and think

about ways to scaffold students toward developing knowledge and practice as appropriate for their classes. Also, we recommend that during the first implementation of the probe, teachers scaffold students to focus on students' own ideas about this problem. By discussing the different options, students can engage in productive talk related to the different response patterns, priming them to look for productive patterns to associate with precipitation in the upcoming lessons.

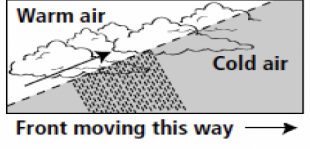
During the ensuing two lessons and the second implementation of the probe, students should be able to use the data from the virtual storm to identify the patterns of temperature, humidity, and movement that are associated with precipitation.

This assessment probe is designed to assess and scaffold students in thinking and talking about how patterns of data (i.e., temperature, humidity, and air movement) are associated with precipitation in a location. To keep things simple, we only focus on two types of fronts: cold fronts and warm fronts. Both types of fronts can be associated with precipitation. Identifying patterns associated with precipitation is an example of the computational thinking practice of **Pattern Recognition**. Once patterns associated with precipitation are identified, they can be used in conjunction with data to make a **Data-based Prediction**, which is what Ms. Tevuk's students are trying to do in the scenario for this assessment probe. **Data-based Prediction** involves identifying patterns and trends within and across groups of data/information as seen in the observable world. The patterns of data associated with precipitation in turn can be used to create rules for the computer model in a process of **Rule Abstraction**. **Rule abstraction** involves creating a general statement derived from exploring patterns in data in order to establish rule(s) or principles. In other words, if you know the patterns of variable values associated with precipitation, you can use those patterns to write rules for predicting precipitation.

As a **cold front** moves into an area, the heavier cool air pushes under the lighter warm air. The air behind a cold front is colder and usually drier than the warm air in front. The warm air in front becomes cooler as it is pushed upward by the entering cold air. If the warm air is humid enough, the water vapor it contains will condense into clouds as it rises, and precipitation may fall.

<p style="text-align: center;">Cold front</p>  <p style="text-align: center;">Front moving this way →</p>	<p>Forms when a cold air mass pushes under a warm air mass, forcing the warm air to rise.</p>	<p>Thunderheads can form as the moisture in the warm air mass rises, cools, and condenses. As the front moves through, cool, fair weather is likely to follow.</p>
--	--	---

When a **warm front** moves into an area, warm air moves above a cool air mass. As the warm air rises, it condenses, often forming clouds and precipitation.

<p style="text-align: center;">Warm front</p>  <p style="text-align: center;">Front moving this way →</p>	<p>Forms when a moist, warm air mass slides up and over a cold air mass.</p>	<p>As the warm air mass rises, it condenses into a broad area of clouds. A warm front brings gentle rain or light snow, followed by warmer, milder weather.</p>
---	---	--

The assessment is situated in the unit context to provide an opportunity for students to make sense about patterns when reasoning about air masses, fronts, and precipitation. In this assessment, one of the responses is scientifically correct while the other responses that relate to some informal ideas about weather that students may hold.



Discussion Points

What Are Common Student Ideas and What Supports Could Help Students?

Below are some issues to consider in assessing students' ideas, and some suggestions of ways to support students in developing more formal reasoning about patterns associated with fronts and precipitation.

Less Formal Ideas

- **Kalin's response (precipitation with a cold front coming in)** suggests that Kalin may just be thinking about patterns of familiar things going together (cold and wet, warm and sunny) rather than thinking about patterns associated with different types of air masses interacting with one another. Students may associate sun with warm, and rain and snow with cold. Based on that experience, they may think that a cold front is connected to precipitation.

Students who choose Kalin's response could benefit from having experiences that challenge their simple association idea. For example, looking at weather maps that show warmer air often has more moisture in it and colder air is often drier can help these students develop understanding that relationships might not be as simple as they think. Students who choose Kalin's response may not think about weather as something that happens with matter. Instead, they may just think of weather in terms of the actions of natural "actors" (e.g., the sun makes it warm, the wind makes it cold). These students could use help in learning that air masses are matter (they are made up of molecules and have mass). If we didn't have air masses made of matter on Earth, then we wouldn't have weather. You might talk with students about concrete examples that they can extend to think about air masses as entities made of stuff – for example, they could compare an empty balloon, a balloon filled with warm air, and a balloon filled with cold air. There is very little stuff (matter or molecules of air) in the empty balloon, while there is more stuff (matter) in the balloons filled with warm and cold air.

Middle Level

- **Both Alex's response (precipitation occurs with a warm front coming in) and Sage's response (precipitation occurs after a cold front has passed by an area)** suggest that they are aware that precipitation is associated with bodies of air moving and with the different conditions in those bodies of air (i.e., temperature, levels of humidity). However, both Sage's and Alex's ideas include some incorrect patterns that lead to problematic rules.

- **Alex** thinks that cold air and warm air mixing causes precipitation. However, warm and cold air masses don't mix very readily, and it is instead actually warm air rising, expanding, and condensing that generally leads to precipitation. In addition, while Alex knows that cold fronts can be drier than warm fronts, that doesn't mean that there won't be any precipitation with cold fronts. Precipitation can occur with a cold front when warm air is pushed up, expands, and condenses.

Students who choose Alex's response could benefit from working with data to observe the following patterns: (1) precipitation can occur with a cold front and (2) fronts do not necessarily mix.

- **Sage** thinks that cold air is humid, which leads to condensation and precipitation. However, warm air masses often have a higher level of humidity compared with cold air masses. It can be confusing that condensation occurs with rising, expanding, and cooling. The unit does not go into this in depth, however. For students who choose Sage's response, it is probably sufficient here to scaffold experiences with data and virtual storm showing the general pattern of warm air masses having higher levels of humidity than cool air masses.

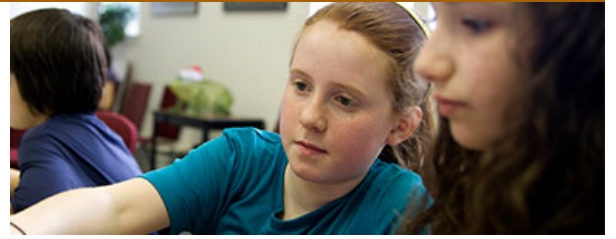
- Students who choose Alex or Sage's response can also benefit from discussion and emphasis on the importance of getting the rule right (enough) in order to create a computational model that works. The students working on their weather model will not be able to predict weather with any consistency or accuracy if there is a big problem with the rule (algorithm) they have written into their computer model.

More Formal Idea

- **Delana's response (precipitation occurs with both cold fronts and warm fronts coming in)** is the best answer. Delana's explanation does not include a more complete mechanistic sequence involving cold air pushing warm air up; warm air rising, expanding, and condensing; and precipitation occurring as a result. This is because the Precipitating Change curriculum simplifies weather phenomena by emphasizing important patterns and de-emphasizing complicated mechanisms. This focus allows the curriculum to scaffold students in connecting observed patterns with rules that can be used to create predictive weather forecasting computer models. Delana's response both recognizes that weather fronts are made of air (matter) and demonstrates recognition of the correct pattern of conditions in air masses (temperatures, levels of

humidity) associated with precipitation forming.

Generate a report



Activity: Lesson 6: Alaska Mainland 2 Post - What Do You Think?

 1  

Lesson 6: Alaska Mainland 2 Post - What Do You Think?

How Do We Make An Accurate Model to Predict Weather?

Ms. Tevuk's students wrote rules for their computer model to predict weather in Alaska. Then they tested their model to see how well it worked. They found their predictions were only right about half the time. The students had different ideas for improving the model.

Here are the students' ideas:



Alex

We need to run the model longer. The longer you run a computer model the better the prediction will be.



Delana

There's no way to make a computer model that can predict the weather. Computer models aren't the real world, so they can't predict what will happen in the real world.



Kalin

We should keep working on the rules based on the science of weather prediction. If we get the science right, the model will be accurate.



Sage

We should keep working on the rules and testing the model predictions using Alaska weather data. That way we'll know the model works.

Question #1

Which student's idea do you agree with the most?

- Alex
- Delana
- Kalin
- Sage

Question #2

Why do you agree with that student's idea the most? Please explain your reasoning.

**Teacher Tip****Pre-Assessment Implementation**

Lesson 6, Opening: How Do We Make An Accurate Model To Predict Weather? Pre-Assessment During the pre-assessment, each student will view the probe and make an individual choice. The request for explanation will only be on the post-assessment. After students make an individual choice for the pre-assessment, lead a discussion including, for example, a pair/share opportunity and/or full class discussion.

At this time, encourage students to share and discuss their ideas and reasons without indicating that any one answer is better or worse. Let the students know that they should keep thinking about this question and that the class will return to it at the end of the lesson.



Theory & Background

Computational Thinking Concepts and Practices:

- Rule Refinement
- Rule Testing

What is the Purpose of this Assessment Probe?

The following information provides background for teachers and is not necessarily representative of the learning goals for everything students should understand. We suggest teachers review the information and think about ways to scaffold students toward developing knowledge and practice as appropriate for their classes. We also recommend emphasizing the underlying ideas and practices, rather than on vocabulary terms. By discussing the different options, students can engage in productive talk related to these computational thinking concepts without using technical terminology.

This assessment probe is designed to assess and scaffold students in thinking and talking about **Rule Refinement** and, especially, **Rule Testing**. **Rule Refinement** involves refining a rule for use in a modeling environment using scientific principles and incorporating additional data, making the rule more precise, computational, and/or detecting and correcting errors. **Rule Testing** is a process of "calibration" between the modeling environment and the observable world. It involves systematically testing a rule to detect errors and limitations, and analyzing the efficiency of various solutions.

The assessment is situated in the unit context to provide an opportunity for students to make sense about rule refinement and rule testing in a specific, tangible weather problem. While one of the responses makes the most sense from a practical standpoint for this problem, other responses have implicit connections to facets of computational thinking (e.g., using understanding of systems to determine model rules and iteration in modeling). Using an indefinite problem provides space for students to think and talk about different approaches to refining and testing rules for a computer model. While having good data and using good science are important for refining rules, calibration (i.e., rule testing) of models of real-world environmental systems like the weather generally rely on comparisons between model outputs and real-world data as critical evidence of model accuracy and usefulness.



Discussion Points

What Are Common Student Ideas and What Supports Could Help Students?

Below are some issues to consider in assessing students' ideas, and some suggestions of ways to support students in developing more formal reasoning about rule refinement and rule testing.

Less Formal Ideas

- **Delana's response (computer models can't predict the weather)** suggests the perspective that if you can't predict something perfectly then your model really doesn't work or have any value. We sometimes see this type of reasoning when students and even adults talk about climate change models. Some people say things like, "they are just models, so they don't really tell us what's going to happen in the real world."

Students who choose Delana's response could use support in developing understanding that computer models don't need to be perfect in order to be useful. No model is perfectly accurate. Models just need to be good enough for a given purpose. Computer modelers set calibration targets that define how close to the actual event they want their prediction to be (e.g., maybe they want their model to predict the correct outcome of rain or no rain for the next day 95% of the time). Uncertainty

is also built into weather models for precipitation in other ways (e.g., a forecast of 70% chance of rain for a given day). Knowing that there is a 70% chance of rain tomorrow is better than not having any information about tomorrow's weather at all. Students who choose Delana's response (as well as all the other students) could benefit from conversations about and examples of what it means for a model to be "good enough" to serve a purpose.

Middle Level

- **Alex's response (run the model longer)** suggests awareness that models are not perfect and that they can improve over time. This response may also appeal to some students who know about the idea of computer iteration (i.e., a process in which a set of instructions is repeated in a sequence multiple times or until a condition is met). Sometimes computer models need to run over time through multiple iterations to reach a certain set of criteria for the output. However, in this case, if the computer model is only making a correct prediction about half the time, then the computer needing more time to process the data probably isn't the problem.

Students who choose Alex's response could benefit from a discussion about whether or not they think running the model longer would help in this case.

- **Kalin's response (keep working on rules using science)** has some merit. We do want to work on model rules based on scientific principles. However, when modeling a real-world system, it's not enough to just use the rules of science to make our models. We need to know that our model is actually going to work (i.e., be sufficiently accurate in outputs such as predictions) when it is applied in the real world. Real-world systems are complex. If we just go by the rules of science, we may be overlooking possible problems with either the rules of the model (e.g., omitting an important variable) or with input data (e.g., perhaps a temperature sensor is faulty or isn't calibrated correctly). In either case, just working on the rules of the model would not be enough to know that the model really works unless we check our weather model predictions against the weather that actually occurs.

More Formal Ideas

- **Sage's answer (refine and test model predictions using Alaska weather data)** is the only option that acknowledges the importance of using calibration to test model rules, which is essential for establishing confidence in a computer model that is designed to represent or simulate an event or phenomenon in the real world. Weather modelers need to compare their model outputs to real-world data to ascertain how well they are working. Calibration is a process of iteratively refining rules and testing with data until a target level of match between model output and real-world data is consistently achieved. Once this level of match (calibration) is achieved with sufficient consistency across multiple datasets, then the modeler can have an acceptable level of confidence in their model's performance. Weather modelers can and should use scientific principles to design and refine their models, but without real-world evidence there is no way to check one's model predictions against what really happens in the world.

Generate a report

Appendix 2: Embedded assessment analysis for 2019-20 academic year

Precipitating Change Embedded Assessment Analysis

Matched T-tests, 2 tailed 1

Qualitative Analyses..... 2

 Notes and possible implications from qualitative analyses 2

 How Should We Break It Down? 3

 How Should We Estimate the Temperature? 6

 Which Estimates Should We Trust? 9

 Rules for Fronts and Precipitation 11

 How Do We Make A Model That Is Accurate? 13

Matched T-tests, 2 tailed

Embedded Assessment	CT Component(s)	N	Pre Avg	Post Avg	P value
How should we break it down?	Data abstraction, Decomposition/discretization	41	1.90	1.98	0.68
How should we estimate the temperature?	Interpolation	41	2.00	2.46	<.001
Which estimates should we trust?	Data aggregation, Interpolation, Extrapolation	40	2.10	1.83	0.25
Rules for fronts and precipitation	Pattern recognition, Data-based prediction, Rule abstraction	42	2.21	2.48	0.03
How do we make a model that is accurate?	Rule refinement, Rule testing	37	2.41	2.16	0.08

Qualitative Analyses

Notes and possible implications from qualitative analyses

1. These analyses draw on written post assessment responses from all students who responses (i.e., not just matched).
2. Following goals of discourse-based learning progression research, an attempt has been made to characterize reasoning categories based on what students ARE doing and ARE thinking about, rather than what they are NOT doing or NOT getting. This represents an effort to focus on students' resources and funds of knowledge rather than characterizing reasoning in terms of deficits with respect to canonical ideas. In turn, knowing how students are making sense of the concepts, practices, phenomena, and systems they are engaging with can help us in various ways – e.g., using students' ideas and ways of making sense to inform responsive curriculum development and design and implementation of PD, to inform development of formative assessment resources that teachers can use to diagnose and respond to their students' ideas and sense making, and providing a way to measure how students' ideas may be changing as a result of educational experiences.
3. While percentages of responses in each category are provided, note that this is a small and unique sample and probably should not be used to draw any conclusions related to prevalence of categories in this or other populations. In addition, these responses emerged in the specific contexts in classrooms of just several teachers. It is likely that other sense making categories would emerge in other contexts and with greater numbers of students.
4. The categories represent a moderate amount of “lumping.” It could be possible to lump responses further into fewer categories or to pull them apart into more categories.
5. The categories have not been organized to designate reasoning more or less aligned with canonical science and computational thinking. What's more, responses in the same reasoning category often differ in levels of “canonical alignment” based on things like which fictional student the responses agree with.
6. In most cases, only categories represented by 2 or more student responses are listed below.
7. One observation based on these analyses is that it appears the embedded assessments are doing a decent job of providing opportunities for students to reason about problems that integrate aspects of computational thinking and the system/phenomenon of study (weather). I find it interesting to see examples of how students are (and sometimes aren't) reasoning about these two facets of the unit together. Problems are contextualized and even when their answers may not reflect canonical ideas, students who are answering in ways that integrate ideas related to CT and weather offer are clearly thinking about and sharing ideas in ways that provide more insights into thinking than would be possible with forced choice items alone.
8. What to do about prevalence of agreement reasoning? Are we okay with it? Could consider ways to refine the questions to decrease these types of answers (e.g., not provide the reasoning, just an answer associated with each fictional student). I'm still thinking it would be nice to have students write out answers on the pre assessment too – or maybe do some subset of clinical interviews? It's hard to get a sense of how their reasoning may have changed when we only have reasoning (i.e., qualitative) data from the post assessment.
9. Do we have any sense of what the teachers and students thought of the embedded assessments? Would be interesting to know how they were perceived.

How Should We Break It Down?

Response Categories (n=48 responses)

Sense Making Category	Description	Example (student ID)*	#(%)
Focus on data, accuracy, and/or precision	Responses in this category represent reasoning that the size of the chunks should be decided either to get a lot or the right amount of data or so that something (e.g., model, data, prediction) will be accurate or precise (in students' words).	Sage. Because he is able to get a lot of data within 200 miles (373156). Sage. The more ground we can cover, the more accurate and the more precise the map will be (373587). Kalin. Because if we want to be accurate we want the right amount of chunks not too many and lot too little (373524).	9(19%)
Agreement and makes sense reasoning	Responses in this category say things like "that answer was the best" or "I like her idea" etc. without also providing separate sense making connected to the assessment question itself that does more than repeat what the fictional student said. Responses with reasoning suggesting choice was based on "it makes sense" are also included in this sense making category.	Alex. I agree with Alex because Alex is saying lets break Alaska down into chunks that are each one square yard in size (358488). Kalin. I agree with Kalin because his idea made the most sense (353157). Sage. Because it just makes more sense that it's 200 miles by 200 miles (373595).	8(17%)
Focus on ease, easier	Responses in this category reason either that the choice they made will be easier for something (like calculating or gathering data) or else they just indicate ease or easier without specifying how or why.	Delana. I agree with Delana because 1 mile seems more easy to work with. If you used boxes bigger you would have to gather different weather things from different places (373172). Kalin. I agree with Kalin because itds easier 373165). Sage. I agree with sage because 200 by 200 miles seems easier to calculate over all of Alaska (358433).	8(17%)
Focus on scale (for prediction or of AK)	Responses in this category generally chose Kalin and we make inferences from the responses that students were reasoning about the appropriate scale for addressing the problem posed in the question. Some responses focused mostly on it being the appropriate scale because of the size of AK.	Kalin. Sage's suggestion is too big for me and the other students were too small. Alaska is a big state and if it was smaller than 20 miles I think it would be time consuming and the data would might be the same because some areas are too close to each other and they might have the same weather or temperature (373590).	6(13%)

		<p>Kalin. I agree with Kalin because his idea say that we should break down spaces between the villages so we know what the temperature is in the villages and we don't need to mess up our weather prediction (373599).</p> <p>Kalin. I agree with sis dude because their are only every other persons explanation was completely of of how big Alaska is and Kalin has a reasonable answer for for many square miles Alaska actually is. (373522).</p>	
Covering AK (Alex or Kalin) or Covering AK: AK is large (Sage)	Responses in the Covering AK (Alex or Kalin) category reason about making sure that all parts of the land in AK are included in a chunk. Responses in the Covering AK: AK is large (Sage) category seem to be reasoning that you need big chunks because AK is large.	<p>Alex. Cause only small parts of Alaska has been left out the squares (373592).</p> <p>Sage. i agree with sage because there's 663,000 square miles in Alaska. it needs a whole lot of ground to cover (376040).</p>	5(10%)
Number, length, or area reasoning	Responses focus on numbers, units of measurement, and/or math (without also discussing things like connections to weather, modeling, predictions). Responses often suggest that the chosen response is a number that is in the right place (e.g., close to or in between other numbers).	<p>Kalin. I agree the most with Kalin because his reasoning sounds right, 20 miles is also in between the other numbers (373167).</p> <p>Sage. Because the answer is 663.000 and 200 is the closets number to 663.000 (358484).</p> <p>Sage. I agree with this person because there are 66,300 square miles in Alaska. 200 by 200 is 40,000 (373593).</p>	5(10%)
Focus on villages or traditional activities	Responses that focused on villages reason that we are interested in the weather in different villages. Only one response related to traditional AK activities.	<p>Delana. I agree with Delana because it's reasonable if you would do chunks of the size of villages to figure out the weather of the specified village (373600).</p> <p>Sage. Because you can have more room for doing stuff like trapping and hunting (358481).</p>	4(8%)
Other	Other categories with fewer than 2 responses.		3(6%)

*Note. Example responses are copied verbatim.

Initial Considerations/Implications

(Note: Considerations and implications are just initial thoughts— I'm sure there are other considerations and implications that could be important that aren't included here.) Students shared lots of different ways of reasoning about this question (asking how to break down the map of Alaska to predict

weather). We see some ideas that are similar to ideas that we encountered in the Comp Hydro project (e.g., connecting with the idea of good answers making something about the process easier, focusing on modeling or predicting being accurate). Students who focused on scale (e.g., of Alaska or of villages) were engaging in some productive reasoning trying to connect the answer to this problem to the particular problem of predicting weather in the state of Alaska. The quantitative analysis did not show a significant change for this question from pre to post. If this is not considered an important issue in the unit (i.e., how to break up a physical space at an appropriate scale for computational modeling of a particular problem such as predicting the weather), then perhaps this embedded assessment could be removed from future phases of the project. If it is kept, consider how the ideas in this embedded assessment are or aren't or could be more explicitly addressed in the unit. Based on experience in Comp Hydro, I think this could potentially be a productive way to address the issue of decomposition with students because it's a very concrete and tangible aspect of decomposition (i.e., breaking down a physical space based on issues of scale of the system and scale at which it makes sense to produce outputs for things like weather). If the unit addresses decomposition in another way, however, again, it might make sense to replace this embedded assessment with a different one that more closely relates to how the unit addresses problem decomposition.

How Should We Estimate the Temperature?

Response Categories (n=49)

Sense Making Category	Description	Example (student ID)*	#(%)
Nearest neighbor reasoning (named or inferred)	Response either refers to the nearest neighbor method as reason for choice or we infer from the response that the student is describing something like the nearest neighbor method.	Alex. Because it is closer to 59 and you have to do nearest neighbor and it is closer to 59 (378658). Sage. Because the x is closer to 59 then 31 so the number close to 59 should be the answer (373162).	11(22%)
Agreement and makes sense reasoning	Responses in this category say things like “that answer was the best” or “I like her idea,” etc. without also providing separate sense making connected to the assessment question itself that does more than repeat what the fictional student said. Responses with reasoning suggesting choice was based on “it makes sense” are also included in this sense making category.	Sage. Because it’s true (358484). Sage. I agree with Sage because the way he explained it was like how i would explain it to another peer (376040).	7(14%)
Linear (interpolation) method reasoning	Response refers to the “linear method” or “linear interpolation method” as reason for choice.	Delana. I agree with Delana because her explanation of the linear method makes more sense rather than the other students (373166). Delana. because she is using the linear method, that method works for predicting the weather (373279).	7(14%)
Number line reasoning (named or inferred)	Response refers to “number line.” Inferred responses generally refer to things like counting. (Might consider lumping these with linear interpolation method as inferred).	Sage. because we did it on a number line and it was closer than any other ideas (373516). Delana. i counted by 2s i got 31 to 59 and the guess is 53 because it is close to 59 but few square away (373588). Delana. Because I counted and got hit e answer with weather (358480).	6(12%)
Estimation reasoning (named or inferred)	Response either refers to “estimation” as reason for choice or we infer from the response that the student is describing something like estimation.	Alex. I agree with Alex cause he is estimate the number between 31and 59 (373164). Sage. if you chose a number from 31-59 and the x was closer to the number 59 it would probably be higher than 31 because the temperature is	4(8%)

		not going to change from 31 and then right away to 51 there has to be some sort of blending of the temperatures (374759).	
Focus on accuracy	Response suggests choice was made because it is an accurate way to solve the problem.	Delana. I agree with Delana the most because X is still far from both temperatures. If we were to use stage's idea , then it wouldn't be accurate as to Delana's way of predicting (373158). Delana. I agree with Delana the most because the X is still far from both temperatures. If we were to use Sage's technique, then it wouldn't be so accurate as to Delana's way of predicting. All in all, his technique is still a popular one since I know for a fact that I use it a lot too (373159). Alex. Because it would probably be more accurate (373275).	4(8%)
You can't predict or estimate the weather reasoning	Reasoning associated with difficulty of or inability to predict the weather, estimate the temperature, etc.	Sage. Because we don't really know whats happening in those places (373156). Sage. Because you never know what the weather is doing (358481). Kalin. Because he is right, it doesn't make sense to estimate the temperature at the x (373258).	4(8%)
Direct measurement is better reasoning	Response agrees with Kalin, who indicates need to go to the place to take the measurement.	Kalin. Because he is measuring it (374630). Kalin. I agree with Kalin's idea because we don't really know what the answer is and cannot measure far away from 31 degrees. We also can use methods we think would lead to the right answer (373599).	2(4%)
Focus on ease, easier	Responses in this category reason either that the choice they made will be easier for something (like calculating or gathering data) or else they just indicate ease or easier without specifying how or why.	Delana. I agree with Delana this because it's easier to find the temperature (358483). Delana. I agree with Delana because guessing the temperature for x is too hard for me to try and guess what the temperature is (373165).	2(4%)
Other	Other categories with fewer than 2 responses.		2(4%)

*Note. Example responses are copied verbatim.

Initial Considerations/Implications

As with the first embedded assessment, we again see that students have shared lots of different ways to reason about this problem. This is one of the embedded assessments where we see some of the clearest evidence that the students are able to relate this problem directly to what they had experienced in a unit lesson. Many students are mentioning concepts and methods from the lessons including linear interpolation and nearest neighbor. We still see some other ways to make sense of this question (e.g., again, focus on accuracy; again, focus on ease; as well as some ideas such as that weather is unpredictable or that you have to measure something directly). These could also be good questions for students to discuss in class. Is weather predictable? Do you always have to measure something directly to have a good sense of the measurement? In Comp Hydro, we found that some students would just say that the world is unpredictable. The idea that we can use computational models to predict things like the weather seems like an important idea for some foundational understanding of computational modeling of Earth systems as an approach for developing explanations and predictions related to systems.

Which Estimates Should We Trust?

Response Categories (n=46)

Sense Making Category	Description	Example (student ID)*	#(%)
Need more or near numbers to interpolate reasoning (no weather station)	Responses in this category tend to agree with Kalin and indicate that there are no or not enough numbers nearby for estimating. These responses do not distinguish between estimated and observed values.	Kalin. I agree with Kalin's idea the most because since that temperature is so far away from any other temperature that it would be harder to interpolate (358478). Kalin. I Choosing Kalin because there's not many (373164). Kalin. I agree with Kalin because the temperatures would not be accurate without the temperatures nearby (373157).	20(43%)
Agreement and makes sense reasoning	Responses in this category say things like "that answer was the best" or "I like her idea," etc. without also providing separate sense making connected to the assessment question itself that does more than repeat what the fictional student said. Responses with reasoning suggesting choice was based on "it makes sense" are also included in this sense making category.	Kalin. I agree with Kalin because I would have the same answer (376040). Sage. Her seams easier to comprehend (358483). Alex. Because I just know (373286).	11(24%)
Nearest neighbor temperature difference reasoning	Some of these are inferred, but in responses in this category students seem to be deciding which estimate is least certain by using nearby numbers from other cells. If the difference between temperatures from nearby cells seems too great to students (e.g., compared with differences in temperature for other cells that are nearby each other in the grid), they indicate that is a reason to be less certain about the chosen estimate.	Sage. Sage is Sage, he is big brain. also it went from 40°F to 37°F out of nowhere (373523). Sage. because 37 is too close to 40 (373524). Kalin. I agree with Kalin because 39 degrees can be around the forties and it can use nearest neighbor since 39 is really close (373599).	6(13%)
Other	Other categories with fewer than 2 responses.		5(11%)
Nearest neighbor?	References nearest neighbor or proximity without mentioning temperature.	Delana. Because the number is close to that number and i would do nearest neighbor (378658).	2(4%)
Need more or near numbers to interpolate reasoning (with weather station)	This responses are similar to those in category above, but in these examples, the response refers to the weather station, perhaps indicating student is distinguishing between observed and estimated values.	Sage. because his estimate was kinda close actually weather station (373279). Delana. I agree with Delana's answer the most because the estimate of 43 is quite far away	2(4%)

		from actual weather station numbers so what method could they have used with it to be near exact? Therefore i am least certain about that and agree with Delana (373600).	
--	--	---	--

*Note. Example responses are copied verbatim.

Initial Considerations/Implications

We see a few prominent ways students made sense of this problem. The largest group of students focused on needing to have other data in close proximity in order to be more certain about a value for a certain location. In most cases, students considered any number shown (i.e., did not seem to distinguish between observed and estimated values). This could be an important idea to highlight with students related to interpolation and extrapolation. Some students also seemed to key into whether or not there was a big difference in temperature between two nearby or adjacent cells. This is interesting because on the one hand, when we interpolate, we take the average between known values. However, it is possible in observed data to see some adjacent cell values that have very similar temperatures and some adjacent cell values that have a temperature difference (e.g., across a front). While this clearly a difficult problem for many students, I would say that most students are making some sense of the problem (even if not canonical) and providing us with some interesting insights into their thinking. In other words, the answer suggest that students aren't just saying "this doesn't make any sense to me."

Rules for Fronts and Precipitation

Response Categories (n=43)

Sense Making Category	Description	Example (student ID)*	#(%)
Reasons that it rains when cold meets or mixes with warm	Responses choose Delana or Alex and indicate that it rains when cold meets or mixes with warm.	Alex. because when an warm front and a cold front mixes you get precipitation (373162). Delana. I agree with Delana the most because there needs to be both cold fronts and warm fronts (373158). Delana. I agree with Delana because precipitation occurs when warm and cold fronts meet (373157).	19(44%)
Agreement and makes sense reasoning	Responses in this category say things like “that answer was the best” or “I like her idea,” etc. without also providing separate sense making connected to the assessment question itself that does more than repeat what the fictional student said. Responses with reasoning suggesting choice was based on “it makes sense” are also included in this sense making category.	Delana. Because her idea relates to my rule and it works (358478). Alex. Because it is a rule of recepiatation (358484). Sage. I agree with sage because it already seems like she used my rule (358483). Kalin. i just picked it because it sounds right! (373274).	10(23%)
Reasons that it rains when cold meets or mixes with warm and adds that cold air is dry	Responses choose Alex and indicate that it rains when cold meets or mixes with warm as well as that cold air is dry.	Alex. I chose Alex because his thought about this seem more believable. He is correct about how cold air is dry and how when cold and warm meet to create precipitation (373253). Alex. Alex is right when cold air and warm air mix, it rains, but you don't get rain with cold air because it's dry (373279).	6(14%)
Other patterns described	Responses describe other patterns involving variables such as air temperature or air moisture.	Alex. You don't get precipitation cause the air is dry (373275). Alex. I agree with Alex because cold air is dry and warm air is moisturized or humid (373593).	6(14%)
Class lesson reasoning	Responses refer to having talked about or learned something in class.	Sage. Because we where just talking about that (358480). Delana. I agree with Delana because it learn that in 7 grade (373161).	2(5%)

*Note. Example responses are copied verbatim.

Initial Considerations/Implications

Students are pretty clearly referring to rules and patterns that they worked on in the lessons during this assessment (so they are making connections). It might be useful to look at students' rules and compare them with their answers to this embedded assessment question. I haven't done that yet. I think this embedded assessment provides students with another opportunity to explain their rules and reasoning related to patterns associated with precipitation, so it is perhaps a nice complement (opportunity to triangulate) with other data collected in the unit.

How Do We Make A Model That Is Accurate?

Response Categories (n=6, only data from one school with few students)

Sense Making Category	Description	Example (student ID)*	#(%)
Agreement and makes sense reasoning	Responses in this category say things like “that answer was the best” or “I like her idea,” etc. without also providing separate sense making connected to the assessment question itself that does more than repeat what the fictional student said. Responses with reasoning suggesting choice was based on “it makes sense” are also included in this sense making category.	Alex. I go with Alex because Alex has a better paragraph to me (358488).	3(50%)
Need for resilience	Response connects to Sage’s indication that she would keep working on it – relating to notion of resilience.	Sage. I agree with Sage because you need to be resilient (358483).	1(17%)
Need for good data	Response seems to connect Kalin’s answer (having to do with the rules of science) with data.	Kalin. Because we found that we needed true data (358482).	1(17%)
Use science to be accurate	Response suggests that using the rules of science will lead to an accurate model of weather prediction.	Kalin. Because he thinks to use the science of weather prediction which I think to be one of the most accurate ways (358478).	1(17%)

*Note. Example responses are copied verbatim.

Initial Considerations/Implications

There are too few responses to derive meaningful implications related to this embedded assessment.