horizon
RESEARCH, INC.

Assessing the Impact of the MSPs: K–8 Science

Classroom Observation Protocol: User Guide

December 2014
Updated August 2015

Horizon Research, Inc.
326 Cloister Court
Chapel Hill, NC 27514

Disclaimer

The *Assessing the Impact of the MSPs: K–8 Science Classroom Observation Protocol: User Guide* was prepared with support from the National Science Foundation under grant number DUE-0928177. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

Suggested Citation

Horizon Research, Inc. (2014). *Assessing the impact of the MSPs: K–8 Science; Classroom observation protocol: User guide.* Chapel Hill, NC: Author.

Table of Contents

Page

Introduction

Researchers have identified principles and practices of effective science instruction that can inform how teachers support student learning.[1] Specifically, there is considerable evidence that shows that instruction is most effective when it elicits students' initial ideas, provides them with opportunities to build on and/or confront those ideas, helps them formulate new ideas based on evidence, and encourages them to reflect upon how their ideas have evolved.[2]

The Assessing the Impact of the MSPs: K–8 Science (AIM) Classroom Observation Protocol (COP) was developed in accordance with these elements of effective science, and provides a structure for examining instruction in order to gauge student opportunity to learn science ideas. The protocol is not intended to advocate a particular set of instructional strategies, but rather focuses on opportunities for fostering conceptual change among students regardless of pedagogy.

The AIM COP is intended to be used to analyze all of the relevant instruction for a particular targeted idea, whether that idea is taught in a single lesson or over multiple lessons. Targeted ideas are the specific scientific concepts that students should understand as a result of instruction, often informed by state and/or national science standards. For the purpose of this protocol, the targeted idea may be stated by the teacher (either verbally or in writing) or may be found in the instructional materials.

The protocol is divided into three main sections: (1) Description of the Instruction, (2) Ratings for Student Opportunity to Learn, and (3) Ratings for Classroom Culture. This user guide provides a brief overview of section of the COP, along with annotation that highlights the intention of each component of the form. The entire unannotated AIM COP is provided at the end of the user guide.

---

[1] National Research Council. (2011). Successful K-12 STEM education: Identifying effective approaches in science, technology, engineering, and mathematics. Washington, DC: The National Academies Press.

National Research Council. (2005). How students learn: Science in the classroom. Committee on How people learn, A targeted report for teachers. M. S. Donovan & J. D. Bransford, (Eds.) Washington, DC: The National Academies Press.

National Research Council. (2000). How people learn: Brain, mind, experience, and school (Expanded Edition). J. D. Bransford, A. L. Brown, & R. R. Cocking (Eds.). Washington, DC: National Academy Press.

[2] Banilower, E., Cohen, K., Pasley, J., & Weiss, I. (2010). Effective science instruction: What does research tell us? (2nd ed.). Portsmouth, NH: RMC Research Corporation, Center on Instruction.

Section One: Description of the Instruction

The first section of the COP, the Description of the Instruction, is where basic details of the instruction are documented, including dates of observations, lengths of lessons, and instructional arrangements/activities used. Additionally, space is provided for observers to write a brief narrative describing the instruction observed so that readers can get a sense for the instruction related to the targeted idea. This narrative describes the structure and flow of the activities, as well as the nature of student intellectual engagement during the lessons, without any judgment of the quality of the instruction.

An annotated version of Section One (Description of the Instruction) is provided below. The annotations are meant to clarify points about completing the section and to more explicitly define some of the terms used in the protocol. Each comment is keyed to the item to which it refers.

SECTION ONE: DESCRIPTION OF THE INSTRUCTION

I. Overview of Instruction

Teacher ID: _____
Researcher completing this COP: _____
Targeted Idea: _____
Number of lessons observed that addressed the targeted idea: _____

Date of Observed Lesson about Idea						
Name of Observer						
1. For each observed lesson that addressed the targeted idea, please give the length of the lesson, teacher's stated purpose for the lesson,[1] approximate length of time spent on the targeted idea, the type of instructional materials used for teaching the targeted idea (textbook, science kit, teacher created, materials from the PD, etc.), and the title (if commercially published).						
Length of lesson						
Teacher's stated purpose for the lesson						
Approximate length of time spent on targeted idea						
Instructional materials used for teaching the targeted idea						
2. Of the time spent on the targeted idea in each observed lesson, approximately how long did students spend working in the following arrangements (in minutes)?						
Individually						
Small Groups						
Whole Class						
3. For the time spent on the targeted idea in each observed lesson, please note if each of the following occurred. (Write "yes" or "no" in each box.)						
Teacher lecture[2]						
Teacher demonstration						
Class discussion[3]						
Writing in journal, recording answers to open-ended questions, etc.						
Completing worksheets, reading textbooks, doing homework, etc.						
Hands-on activities						
Other						

[1] Observers may ask the teacher for the purpose of the lesson, find it posted in the room, or the teacher may state to the class the purpose for the lesson.

[2] A lecture is when the instructor is providing information (facts, definitions, etc.) to the learners as opposed to this information coming from hands-on activities or a written source. Delivering instructions for an activity should NOT be considered lecture.

[3] Discussion is when learners are talking with each other or with the teacher. Discussion can occur as part of a lecture, demonstration, or hands-on activity.

I. Overview of Instruction

Researcher completing this COP: Enter one name. If there were multiple observers for a specific targeted idea, they should be listed in the table.

Targeted Idea: Enter the science idea that was the focus of instruction being analyzed.

Number of lessons observed that addressed the targeted idea: Enter only the number of lessons that addressed the targeted idea. If zero, the remainder of the protocol should not be completed.

1. **Length of lesson:** Enter the length of the entire lesson in minutes, regardless of whether only a portion of the time was spent addressing the targeted idea.

Teacher's stated purpose for the lesson: The teacher may state the purpose for the lesson during the observed lesson, either verbally or in writing. However, if the purpose of the lesson is not explicitly stated/written, observers may ask the teacher about the intended purpose.

2. Arrangements: Enter the number of minutes spent in each arrangement (i.e., time as individuals, small groups, and whole class). Only instructional time spent on the targeted idea should be considered in these time estimates.

3. Instructional activities: Indicate whether or not each instructional activity occurred during the observed lesson. Only instructional time spent on the targeted idea should be considered when deciding whether or not these activities occurred.

A lecture refers to when the instructor is providing information (facts, definitions, etc.) to the learners as opposed to this information coming from hands-on activities or a written source. Delivering instructions for an activity should not be considered a lecture.

A class discussion refers to learners talking with each other or with the teacher for the purpose of the instructional activity. Discussion can occur as part of a lecture, demonstration, or hands-on activity.

II. Comments

Your ratings in Section Two will focus on the instruction about the targeted idea. Use this space for any information on the lesson and/or classroom that you consider so salient that it needs to get "on the table" right away to help explain your ratings.

III. Description of Instruction about Targeted Idea

Describe what happened in the observed instruction. The description should include enough rich detail that readers have a sense of what happened during the instruction, but should not include a judgment of quality. Include:

- A description of the structure and flow of the instruction related to the targeted idea (i.e., the nature and progression of the activities).
- A description of the level of intellectual engagement (i.e., who is doing the intellectual work – the teacher, some students, most students).

II. Comments

Describe anything unusual that may have occurred during the observed instruction. For example, this would be a good place to note that a large number of students were absent or called out of class.

III. Description of Instruction about Targeted Idea

Provide readers with a brief overview of the instruction for the targeted idea (e.g., 1–2 paragraphs describing each observed class session). The activities that addressed the content should be mentioned, but the details of each activity should be saved for Section Two. For example:

> *The teacher began the lesson on by having pairs of students apply pushes to a toy car and discuss whether a force was acting on the car when it was moving. Students were also instructed to come up with other ways to get the car to move and to consider whether a force was acting under those conditions. The teacher gave each pair of students a piece of foam and explained that a change in the shape of the foam indicated the presence of a force. Students attached the foam to the top of their toy car and pushed on it, observing the ensuing change in the car's motion and considering whether a force was acting on the car. The teacher concluded the lesson by facilitating a whole class discussion about forces.*

Section Two: Ratings for Student Opportunity to Learn

The Ratings for Student Opportunity to Learn section is divided into five subsections:

I. Science Content,
II. Opportunities to Surface Prior Knowledge,
III. Engaging with Examples/Phenomena,
IV. Using Evidence to Draw Conclusions and Make Claims about the Examples/Phenomena, and
V. Sense Making of the Targeted Idea.

The Science Content subsection focuses on the accuracy, appropriateness, and completeness of the content. The remaining four subsections are directly aligned to the elements of effective instruction: Opportunities to Surface Prior Knowledge, Engaging with Examples/Phenomena, Using Evidence to Draw Conclusions and Make Claims about the Examples/Phenomena, and Sense Making of the Targeted Idea. For each of these four subsections, a checkbox is provided to indicate if the observed instruction did not include deliberate opportunities for students in a particular subsection. For example, a teacher may start instruction for a targeted idea without asking students to consider their initial ideas. Therefore, the observer would check the box to indicate that the observed instruction did not include deliberate opportunities to surface students' prior knowledge and that subsection would not be rated. Similarly, instruction on the targeted idea may not include deliberate opportunities to help students connect what they did to the targeted idea, their prior ideas, or other science ideas. Therefore, the observer would check the box indicating that the observed instruction did not include opportunities for sense making.

Rating the Instruction

Each of the five subsections consists of three sets of ratings:

A.1. Extent to which key indicators were present,
A.2. Alignment of the relevant instruction to the targeted science idea, and
B. Sufficiency of the relevant instruction for learning the targeted idea (also called the synthesis rating).

In each subsection, observers should:

1. Provide a brief description of what happened in the instruction related to the subsection (e.g., what examples/phenomena students engaged with, what evidence was used to make claims).

2. Rate the key indicators on a four-point scale ranging from "not at all" to "to a great extent." The extent to which the key indicator was present in the instruction and the level of student engagement (i.e., whether most students are participating or just a few) are both taken into consideration when assigning ratings. The observer should also support each rating by providing a statement of the quality of the instruction, citing specific pieces of evidence from the observed instruction that were used to assign the rating (e.g., particular activities, student/teacher interactions). These descriptions are essential for documenting how the ratings were generated.

3. Rate the extent to which the instruction is aligned with the targeted idea. As with the key indicators, the observer should provide a rationale and detailed description of the evidence that was used to determine the rating.

4. Provide a final (synthesis) rating focused on whether the science content related to the targeted idea was appropriate for students' learning of the targeted idea (science content subsection only) or the extent to which opportunities in the instruction were likely to be sufficient for student learning of the targeted idea (all other subsections). This holistic rating should be supported by a written justification that takes the individual indicator ratings into account (i.e., key features and alignment), but is not intended to be an arithmetic average of ratings of the key indicators. It is important to note that synthesis ratings are specific to each subsection and are not representative of the instruction as a whole.

Taken together, these ratings represent the deliberate opportunities that were made available to students in order to further their conceptual understanding. The five subsections that comprise Section Two (Ratings for Deepening Student Content Knowledge) are explained in detail below and illustrative examples of each are provided.

I. Science Content

In order to gauge student opportunity to learn the targeted science idea, observers must begin by considering the actual science content addressed in the instruction. This subsection of the protocol asks observers to rate the extent to which the science content addressed in the instruction was both accurate and developmentally appropriate. All of the science content that is addressed during instruction, *whether or not it is related to the targeted idea*, is considered for the accuracy and developmental appropriateness ratings. For example, a targeted idea for a science lesson may be related to the phases of the moon. During instruction on this idea, a teacher may inaccurately state that the moon does not have a gravitational pull. In this case, the accuracy rating would be lowered, even though the inaccuracy was not directly related to the targeted idea. However, because the inaccuracy is not about the targeted idea, it would not affect the synthesis rating.

An annotated version of the Science Content subsection is provided below. Each comment is keyed to the area or item to which it refers.

SECTION TWO: RATINGS FOR OPPORTUNITY TO LEARN

I. Science Content[4]

A. Ratings of Key Indicators

The first part of this section asks you to rate the extent to which several key features of the science content were present, regardless of how well aligned the content was to the targeted idea. The second part asks for your judgment of that alignment.

	Not at all			To a great extent
1. The science content addressed in this instruction was:				
a. accurate.	1	2	3	4
b. developmentally appropriate.	1	2	3	4

	Not at all			To a great extent
2. To what extent was the science content addressed in this instruction *aligned* with the targeted idea?	1	2	3	4

B. Synthesis Rating: Science Content

The synthesis rating is not intended to be an arithmetic average of the individual indicator ratings. It is the observer's responsibility, with knowledge of the purposes of the science instruction and the experience of being there, to come up with a holistic rating. For the science content synthesis rating, if any indicator is rated a "1," the synthesis rating must a "1"; a synthesis rating of "4" can be given only if each key indicator is a "4."

For the purposes of this lesson and considering this group of students, the science content **related to the targeted idea** was:

1	2	3	4
Inappropriate (i.e., unaligned, inaccurate, and/or developmentally inappropriate)	Minimally appropriate	Moderately appropriate	Very appropriate (i.e., aligned, accurate, developmentally appropriate, and *covers the entire targeted idea*)

[4] Science Content can include disciplinary science concepts and/or ways of knowing in science.

I. Science Content

A. Ratings of Key Indicators

1.a. Observers should consider whether the science content addressed was accurate. The accuracy rating should start at a "4." If there were minor content inaccuracies the rating should be lowered to a "3." The rating should be lowered even further if the content inaccuracies were more critical.

1.b. Observers should consider if the science content addressed was appropriate for the students in the class (e.g., ability level, prior understanding of the students), and are encouraged to consult relevant curriculum documents (e.g., state standards, district curriculum guides) to determine if the content was developmentally appropriate for the particular grade level. Again, whether the instruction addressed all or part of the idea should not factor into this assessment. If all of the addressed science content was appropriate, the rating should be a "4." If aspects of the addressed science content were not developmentally appropriate, the rating should be lowered.

2. Observers should consider the extent to which the science content addressed was aligned with the targeted idea. Given that this protocol is used to rate student opportunity to learn a specific idea, this rating will generally be a "4." Further, it is not possible to give this indicator a rating of "1." If observers are inclined to rate this indicator as a "1," it is likely that the protocol should not be used for the targeted idea at all.

The quality of instruction and amount of time devoted to teaching this idea should not factor into this rating. Similarly, the alignment rating should not be adjusted if only part of the targeted idea is addressed in instruction.

B. Synthesis Rating: Science Content

Observers should consider the content accuracy, developmental appropriateness, and alignment when assigning a synthesis rating. The synthesis rating should be lowered if instruction did not address the entire targeted idea, regardless of whether the content that was addressed was accurate and developmentally appropriate. For example, if the targeted idea is to "develop an understanding of the structure and function of cells," but the instruction focuses only on cell structures, the synthesis rating should be lowered to reflect the lack of emphasis on cell function.

If any key indicator is rated a "1," the synthesis must also be rated a "1."

C. Supporting Evidence for Synthesis Rating

Description of Science Content Related to the Targeted Idea

1.a. The science content addressed in this instruction was accurate.

1.b. The science content addressed in this instruction was developmentally appropriate.

2. To what extent was the science content addressed in this instruction aligned with the targeted idea?

Synthesis

If any key indicator is rated a "2" or "3," the synthesis can still be a 4 as long as the lowered ratings reflect science content that is not aligned to the targeted idea.

C. Supporting Evidence for Synthesis Rating

Observers should thoroughly describe what happened in the instruction related to this subsection, as well as when (i.e., what day, which lesson) it occurred. In addition, for each key indicator listed in this subsection, observers should provide a detailed description of the evidence used to determine the ratings. For example:

Description of Science Content Related to the Targeted Idea: *The science content addressed in the instruction was related to the following targeted idea:*

A force is a push or pull interaction between two objects. Forces can vary in strength and direction.

The students engaged with this content by examining the changes in shape of a foam "force indicator" as well as changes in motion of objects as the result of pushes and pulls.

1.a. *The content addressed in this instruction was accurate to a great extent. No misconceptions were introduced, and the language used by the teacher and students was appropriate.*

1.b. *The science content addressed is part of the state standards for elementary students, and, thus, is considered developmentally appropriate for this group of 5th graders.*

2. *The science content in this instruction was greatly aligned to the targeted idea. The students observed both pushes and pulls were forces acting on a toy car.*

Synthesis

The instruction was accurate, developmentally appropriate, and well-aligned to the targeted idea. However, the instruction did not cover the entire targeted idea, specifically that a force is an interaction between two objects. Thus, the science content observed was only moderately appropriate for the students' learning of the targeted idea.

II. Opportunities to Surface Prior Knowledge

Students come to school with pre-existing ideas about science, and learning theory suggests that instruction is more effective when it takes students' initial ideas into account. When students hold naïve conceptions/misconceptions that do not align with current scientific thinking, these ideas need to be addressed. Similarly, when students hold ideas that are consistent with scientists' views, making these ideas explicit allows students to connect new knowledge with pre-existing knowledge, increasing the likelihood of learning and retention. Therefore, effective science instruction surfaces students' prior ideas about a targeted idea, and their reasons for how they are thinking, so that subsequent instruction can provide experiences that facilitate a deeper level of understanding. In this subsection, observers rate and provide rationales for five key features of opportunities to surface prior knowledge. Observers rate the extent to which these deliberate opportunities:

- Are structured/implemented so that students would be aware of their own prior knowledge,
- Surface students' reasons for how they are thinking,
- Have students record aspects of their prior knowledge,
- Have students make public aspects of their prior knowledge, and
- Allow for students' ideas to be surfaced without judgment.

For the purpose of completing this protocol, prior knowledge is defined as what students think about the science idea prior to instruction on it, and does not include a review of the previous instruction. The following scenarios show two different ways that students' prior knowledge may be surfaced during instruction. In both scenarios, the instruction is meant to address the targeted idea that ***in a contact push/pull interaction, the force ceases to exist as soon as contact between the interacting objects is lost.***

Scenario A: Surfacing Prior Knowledge

A teacher asks her students to answer the following question:

Imagine a soccer player taking a shot on goal. She runs up and kicks the ball which flies toward the goal, where the goalkeeper catches it.

Which of the choices below is closest to when you think the force of the kick stopped acting on the ball?

a) Before the ball lost contact with the foot
b) At the moment the ball lost contact with the foot.
c) After the ball lost contact with the foot, but before it got to the goalkeeper.
d) When the goalkeeper stopped the ball moving.

The teacher instructs the students to record their answers in their notebooks and then move to one of four different locations in the classroom depending on which answer they chose. The students in each of the four groups discuss their reasoning for their response, and one student is selected to share each group's ideas with the whole class.

Scenario B: Surfacing Prior Knowledge

A teacher asks her students the following question:

What are some examples of forces that you saw on your way to school this morning?

The teacher instructs the students to record their answers in their notebooks and share their ideas with a partner. She then calls on several students to share their ideas with the whole class.

The instruction in both scenarios provides opportunities for most students to discuss their prior knowledge, record aspects of their prior knowledge, and make public aspects of their prior knowledge (i.e., share their ideas with others). However, only the instruction in Scenario A allows students to consider the specific science idea that is the focus of the lesson and surfaces students' reasons for how they are thinking. Therefore, the instruction in Scenario A would receive a synthesis rating of a "4" on the protocol and the instruction in Scenario B would be rated lower.

An annotated version of the Opportunities to Surface Prior Knowledge subsection is provided below. Each comment is keyed to the area or item to which it refers.

II. Opportunities to Surface Prior Knowledge

A. Ratings of Key Indicators

☐ Check here if the observed instruction addressing the targeted idea did not include opportunities to surface students' prior knowledge and skip to III. Engaging with Examples/Phenomena.

Of the observed lessons that addressed the targeted idea, how many included opportunities to surface prior knowledge? _____

The first part of this section asks you to rate the extent to which several key features of surfacing students' prior knowledge engaged students, regardless of how well aligned the surfacing was to the targeted idea. The second part asks for your judgment of that alignment.

	Not at all			To a great extent
1. Deliberate opportunities provided to surface students' prior knowledge:				
a. were structured/implemented so that students would be aware of their own prior knowledge.	1	2	3	4
b. surfaced students' reasons for how they were thinking.	1	2	3	4
c. had students record aspects of their prior knowledge.	1	2	3	4
d. had students make public aspects of their prior knowledge.	1	2	3	4
e. allowed students' ideas to be surfaced without judgment.	1	2	3	4

II. Opportunities to Surface Prior Knowledge

A. Ratings of Key Indicators

NOTE: The ratings in the first part of this subsection (1.a.–e.) do not take into consideration how well aligned opportunities to surface students' prior knowledge were to the targeted idea.

1.a. Observers should consider whether students had opportunities to become aware of their own prior knowledge. If most students had opportunities to consider their prior knowledge, the rating should be a "4." This rating should be lowered if fewer students had this opportunity.

This indicator is focused only on opportunities for students to become aware of their prior knowledge. The rating should not be lowered if there were not opportunities for students to consider the reasons for their ideas (which is addressed in 1.b.).

1.b. Observers should consider whether students had opportunities to consider the reasons for how they were thinking about science concepts. Students can provide reasons for how they were thinking verbally, in small or large group discussions, or in writing. If most students had opportunities to consider reasons for how they were thinking, the rating should be a "4." This rating should be lowered if fewer students had this opportunity.

1.c. Observers should consider whether students made a permanent record of their prior knowledge that they could return to at a later time. If most students had opportunities to record aspects of their prior knowledge, the rating should be a "4." If groups of students recorded their collective prior knowledge, this rating may be lowered if there was not an expectation that all students' ideas would be written down.

Students writing in their notebooks would factor into the rating for 1.c. as notebooks would provide a permanent record. Students recording their ideas on whiteboards, which may be erased (and not provide a permanent record) should factor into 1.d.

1.d. Observers should consider whether students had opportunities to share their thinking. It's not necessary that each student in the class share, as long as their ideas were brought up by others. If the observer thinks that most students' ideas were made public, this indicator should be rated a "4." If only some students had this opportunity, the rating should be lowered.

1.e. Observers should consider whether or not the teacher commented on the correctness of the students' initial ideas/prior knowledge. If so, this rating should be lowered.

This indicator is concerned only with whether the teacher, who students may consider an expert, judged initial ideas, not the other students. The extent to which students judge one another's ideas (whether positively or negatively) should be addressed in the Classroom Culture section.

	Not at all			To a great extent
2. To what extent were the opportunities provided to surface students' prior knowledge *aligned* with the targeted idea?	1	2	3	4

B. Synthesis Rating: Surfacing Prior Knowledge

The synthesis rating describes the sufficiency of the opportunities to surface prior knowledge for understanding the targeted idea. The synthesis rating is not intended to be an arithmetic average of the individual indicator ratings. It is the observer's responsibility, with knowledge of the purposes of the science instruction and the experience of being there, to come up with a holistic rating. However, item 2 is a critical indicator; the synthesis rating cannot be higher than the rating of this indicator.

To what extent were opportunities for students to surface their prior knowledge likely *sufficient* for their **learning of the targeted idea**?

1	2	3	4
Not at all likely	Minimally likely	Moderately likely	Very likely

2. Observers should consider the extent to which opportunities to surface students' prior knowledge were aligned with the targeted idea. The quality of instruction and amount of time devoted to teaching this idea should not factor into this rating.

B. Synthesis Rating: Surfacing Prior Knowledge

Observers should consider the extent to which opportunities to surface students' prior knowledge were sufficient for their learning of the targeted idea. The synthesis rating cannot be higher than the rating for alignment in 2.

Learning theory indicates that students need to commit to their ideas, either through writing or by making them public. However, it is not necessary for them to do both. If students did only one (either 1.c. or 1.d.) it should not impact the synthesis rating.

C. Supporting Evidence for Synthesis Rating

For each key indicator listed above, and the synthesis rating, provide 1) a description of the relevant instruction, 2) a statement of quality of the instruction for the indicator, and 3) the rationale and evidence for your statement of quality. Additionally, if anything else factored into your synthesis rating, please describe it below.

Description of Opportunities to Surface Prior Knowledge of the Targeted Idea

1.a. Deliberate opportunities provided to surface students' prior knowledge were structured/implemented so that students would be aware of their own prior knowledge.

1.b. Deliberate opportunities provided to surface students' prior knowledge surfaced students' reasons for how they were thinking.

1.c. Deliberate opportunities provided to surface students' prior knowledge had students record aspects of their prior knowledge.

1.d. Deliberate opportunities provided to surface students' prior knowledge had students make public aspects of their prior knowledge.

1.e. Deliberate opportunities provided to surface students' prior knowledge allowed students' ideas to be surfaced without judgment.

C. Supporting Evidence for Synthesis Rating

Observers should thoroughly describe what happened in the instruction related to this subsection, as well as when (what day, which lesson) it occurred. In addition, for each key indicator listed in this subsection, observers should provide a detailed description of the evidence used to determine the ratings. For example:

Description of Opportunities to Surface Prior Knowledge of the Targeted Idea: *There were two opportunities for surfacing students' prior knowledge related to the following targeted idea:*

> *A force is a push or pull interaction between two objects. Forces can vary in strength and direction.*
>
> *On 11/15, students generated their own definitions for force. On 11/17, they voted for when they thought the force of a kick stopped acting on a ball.*

1.a. *Opportunities for surfacing prior knowledge were structured/implemented such that students would be aware of their prior knowledge to a great extent. The teacher asked students to come up with their own definitions of "force" and then called on each student to share his/her definition aloud. The teacher also asked students to write down predictions about when the force of a kick stopped action on a ball before voting by show of hands for when they thought the force of a kick stopped acting on a ball.*

1.b. *Students' reasons for how they were thinking were surfaced to a moderate extent. The teacher had students write down predictions about when the force of a kick stopped acting on a ball and reasons for their thinking. He then had all students vote by show of hands and called on five students to explain their thinking aloud. However, during the force definition activity, the teacher did not ask for students' reasons.*

1.c. *The opportunities to surface student ideas had them record aspects of their prior knowledge to a moderate extent. All students wrote down their predictions and reasons during the ball kicking activity. However, students did not write down their definitions of "force."*

1.d. *The opportunities to surface student ideas had them make public aspects of their prior knowledge to a great extent. All the students voted for when they thought the force stopped during the ball kicking activity and five students shared their thinking aloud. Additionally, all students shared their own definitions of "force" aloud.*

1.e. *Students' ideas were allowed to be surfaced without judgment.*

2. To what extent were the opportunities provided to surface students' prior knowledge aligned with the targeted idea?

Synthesis

2. *The opportunities to surface prior knowledge were well aligned to the targeted idea. The instruction included an opportunity for students to consider their initial ideas about what a force is. Students were also prompted to think about when a force stops acting on an object.*

Synthesis

There were two opportunities for students to surface their ideas related to the targeted idea. In both instances, students made aspects of their prior knowledge public, either verbally or in writing. In addition, all student ideas were surfaced without judgment. However, students were not asked to provide reasons for how they came up with definitions of "force." Further, students were not given opportunities to surface their ideas about all parts of the idea, specifically that a force is an interaction between two objects. Therefore, opportunities to surface students' prior knowledge were minimally likely to be sufficient for students' learning of the targeted idea.

III. Engaging with Examples/Phenomena

Regardless of the mode of instruction (e.g., hands-on laboratory activity, lecture), learning theory suggests that effective science instruction includes experiences that engage students with examples and phenomena that are designed to build student understanding of targeted science ideas. Although a hands-on experience may be interesting to and fun for students, it is unlikely to further most students' conceptual understanding if it is not explicitly focused on a meaningful question and provides evidence of the targeted idea. In addition, learning theory indicates that students should focus on aspects of the experience that are key to developing an understanding of the targeted idea. For example, a teacher may develop students' understanding of average speed (i.e., a measure of distance divided by time) by pushing a ball. However, if students are focused on the strength of the push rather than the time it takes the ball to travel a certain distance, they are not focused on the relevant aspects of the activity.

Learning theory also indicates that classroom activities must be accessible to students. For example, if students are asked to add 15 milliliters of water to a solution but are unable to accurately use graduated cylinders to measure liquids, the activity is likely not accessible to them. Similarly, if students are using probes to take dissolved oxygen measurements at a local stream but do not know what the probe readings represent/how to interpret them, the activity is likely not accessible to them.

Three key features of engaging with examples/phenomena are rated on the protocol:

- The examples/phenomena are accessible to students,
- Students are focused on the relevant aspects of the examples/phenomena, and
- Students describe and/or keep a record of the processes they use or the data they generate.

The following two scenarios illustrate classroom experiences intended to address the targeted idea that ***in a contact push/pull interaction, the force ceases to exist as soon as contact between the interacting objects is lost.***

Scenario A: Engaging with Examples/Phenomena

Pairs of students each receive a toy car with a piece of soft foam attached to the top.

The students push on the back edge of the foam to get the car moving. As the car is moving, the students give the back edge of the foam several more quick pushes, observing the shape of the foam before, during, and in between pushes. The pairs of students discuss their observations and record them in their science notebooks. They then repeat the activity, this time making quick pulls on the front edge of the foam.

Scenario B: Engaging with Examples/Phenomena

Pairs of students each receive a toy car with a piece of soft foam attached to the top.

The students measure how long it takes for the car to move a specified distance when they push on the foam softly. The pairs of students record their measurements in their science notebooks. They then repeat the activity, this time pushing harder on the foam and measuring how long it takes to cover the same distance.

The activities in both scenarios are accessible to students and require them to keep a record of the data they generate. Yet, only Scenario A focuses students on evidence about when forces start/stop acting on an object. Although Scenario B may be a fun and engaging activity, it focuses on measuring speed, which is not part of the targeted idea. Students would be unlikely to learn the targeted idea as a result of this activity because they are not focused on the relevant aspects. Therefore, Scenario A would receive a synthesis rating of a "4" on the protocol and Scenario B would be rated lower.

An annotated version of the Engaging with Examples/Phenomena component is provided below. Each comment is keyed to the area or item to which it refers.

III. Engaging with Examples/Phenomena

A. Ratings of Key Indicators

☐ Check here if the observed instruction addressing the targeted idea did not include opportunities for students to engage with examples/phenomena for developing or reinforcing science ideas and skip to Using Evidence to Draw Conclusions and Make Claims about the Examples/Phenomena.

Of the observed lessons that addressed the targeted idea, how many included opportunities to engage with examples/phenomena? _____

The first part of this section asks you to rate the extent to which several key features of engaging students with examples/phenomena were present, regardless of how well aligned the examples/phenomena were to the targeted idea. The second part asks for your judgment of that alignment.

	Not at all			To a great extent
1. Deliberate opportunities to engage with examples/phenomena:				
a. included examples/phenomena *that were accessible to students*.	1	2	3	4
b. focused students on *relevant aspects* of the examples/phenomena.	1	2	3	4
c. had students describe and/or keep a record of the processes they used/data they generated.	1	2	3	4
	Not at all			To a great extent
2. To what extent did the examples/phenomena *provide evidence for* the targeted idea?	1	2	3	4

III. Engaging with Examples/Phenomena

A. Ratings of Key Indicators

NOTE: The ratings in the first part of this subsection (1.a.–c.) do not take into consideration how well aligned opportunities to engage with examples/phenomena were to the targeted idea.

1.a. To rate indicator 1.a., observers should first identify the examples/phenomena that the students engaged with. Observers should then consider whether students could: (1) understand and relate to the examples/phenomena based on their previous experiences (in or out of school), (2) successfully carry out the tasks and procedures so they could collect meaningful data, and (3) understand the meaning of any representations or examples that were used. If some of the examples/phenomena were not accessible to students, the rating should be lowered. If many of the examples/phenomena were not accessible to students, the rating should be lowered even more.

1.b. Observers should consider whether the students were attending to the important features of the examples/phenomena that will later be used as evidence. If students were focused on aspects of the examples/phenomena that were not aligned with the targeted idea or not important for developing an understanding of the idea, the rating should be lowered.

1.c. Observers should consider whether students described and/or kept a record of the process they used (e.g., steps for carrying out a procedure) and/or data they generated (e.g., the time it takes a moving car to go a certain distance). The rating should not be lowered if students only did one of these two things.

2. Observers should consider the extent to which the examples/phenomena provide evidence for the targeted idea. The quality of instruction and amount of time devoted to teaching this idea should not factor into this rating.

   If the examples/phenomena provide evidence for only part of the targeted idea, this indicator should still be rated high. However, the synthesis should be rated lower.

B. Synthesis Rating: Engaging with Examples/Phenomena

The synthesis rating describes the sufficiency of the students' engagement with examples/phenomena for understanding the targeted idea. The synthesis rating is not intended to be an arithmetic average of the individual indicator ratings. It is the observer's responsibility, with knowledge of the purposes of the science instruction and the experience of being there, to come up with a holistic rating. However, item 2 is a critical indicator; the synthesis rating cannot be higher than the rating of this indicator.

To what extent were opportunities for students to engage with examples/phenomena likely *sufficient* for their **learning of the targeted idea**?

1	2	3	4
Not at all likely	Minimally likely	Moderately likely	Very likely

B. Synthesis Ratings: Engaging with Examples/Phenomena

Observers should consider the extent to which opportunities for students to engage with examples/phenomena were sufficient for their learning of the targeted idea. Included in this rating is consideration of whether or not the phenomena allowed for confrontation of common misconceptions/the misconceptions the students in the class had. The synthesis rating cannot be higher than the rating for A.2.

C. Supporting Evidence for Synthesis Rating

For each key indicator listed above, and the synthesis rating, provide 1) a description of the relevant instruction, 2) a statement of quality of the instruction for the indicator, and 3) the rationale and evidence for your statement of quality. Additionally, if anything else factored into your synthesis rating, please describe it below.

Description of Opportunities to Engage with Examples/Phenomena Aligned to the Targeted Idea

1.a. Deliberate opportunities to engage with examples/phenomena included examples/phenomena that were accessible to students.

1.b. Deliberate opportunities to engage with examples/phenomena focused students on relevant aspects of the examples/phenomena.

C. Supporting Evidence for Synthesis Rating

Observers should thoroughly describe what happened in the instruction, as well as when (what day, which lesson) it occurred. In addition, for each key indicator listed in this subsection, observers should provide a detailed description of the evidence used to determine the ratings. For example:

Description of Opportunities to Engage with Examples/Phenomena Aligned to the Targeted Idea: *There were several opportunities for students to engage with phenomena aligned with the following targeted idea:*

A force is a push or pull interaction between two objects. Forces can vary in strength and direction.

On 11/17, students examined the effects of pushing and pulling on a piece of foam and established that the foam could serve as a force detector. On 11/18, students were instructed to use the foam force detector to examine when forces existed in conjunction with several rolling cart scenarios. On 11/19 the class examined several examples of forces acting without contact between objects, including magnets applying force on paper clips and a balloon applying force on tiny pieces of paper. On 11/20, students again used the foam force detectors to investigate the duration of forces when varied pushes and pulls were applied to a wooden block.

1.a. *The phenomena the students engaged with were readily accessible to them. The students have experienced object movements like those encountered in the tasks of this lesson. Relevant changes in the foam shape were visible when attended to. Students at this age generally have some familiarity with identifying changes in movement with respect to objects such as magnets, paper clips, pieces of paper, blocks, or toy vehicles. In addition, the magnets were of sufficient strength that their attraction/repulsion could be felt.*

1.b. *Overall, the instruction focused students on relevant aspects of the phenomena to a moderate extent. The directions presented generally had students focus on appropriate pushes and pulls applied and their effects on the foam's shape and the motion of the objects. However, students inconsistently attended to the relevant aspects as directed. In a few instances, many students were not engaged (e.g., simply played with the materials supplied rather than focusing on applying specified pushes/pulls or observing relevant changes).*

1.c. Deliberate opportunities to engage with examples/phenomena had students describe and/or keep a record of the processes they used/data they generated.

2. To what extent did the examples/phenomena provide evidence for the targeted idea?

Synthesis:

1.c. *The engagement with examples/phenomena had students describe/record data to a limited extent. In many cases, students were instructed to perform the activities in small groups, but were not given explicit directions for recording or describing the results of their actions. For example, no mechanism was used to have students identify or collect data for cart motion, foam shape change, or motion resulting from "non-touching" interactions.*

2. *The phenomena provided evidence for the targeted idea to a great extent. Pushing and pulling on the foam makes it clearly visible when a force was being applied to the cart/block. The use of strong magnets and the charged objects also provided tangible evidence of forces acting at a distance.*

Synthesis

It is moderately likely that the students' engagement with the phenomena was sufficient for their learning of the targeted idea. The examples were accessible to the students and provided evidence for the targeted idea. Much of the engagement with phenomena had students attend to application of contact forces and changes in object motion (e.g., during and after pushes/kicks). The instruction also focused students on non-contact interactions and resulting motion changes (e.g., falling ball, magnetic attraction). However, the inconsistent attention to the relevant aspects of the phenomena likely decreased the probability of the instruction being sufficient for most students to develop an understanding of the target idea.

IV. Using Evidence to Draw Conclusions and Make Claims about the Examples/Phenomena

Science is not just the memorization of facts, but a process by which understanding of the natural world is developed and revised using evidence. Similarly, learning theory suggests that opportunities to use evidence to and make and critique claims about examples/phenomena help deepen students' conceptual understanding. Evidence refers to analyzed or processed data that are used to support a scientific claim or conclusion. The data used as evidence to make claims can either be directly gathered by students (e.g., observations, experiments) or those collected by others and given to students. In either case, it is important that students understand what the data represent. For example, young students may have difficulty understanding that a table containing data on the velocity of an object represents the speed and direction of the moving object.

It is also important that students use data as evidence to make claims about what is being studied rather than just restating observations. For example, "I saw that the car was red," or "We measured the car to be 15 centimeters," is simply restating data and are not claims based on data. Rather, making claims requires going beyond observations and data to draw meaning from them (e.g., "The red car crossed the finish line in 10.5 seconds and the blue car crossed the finish line in 15.2 seconds, so the red car must have been moving at a higher speed."). Student understanding is further solidified if they have opportunity to examine others' claims, critiquing the substance of those claims using existing evidence.

Four key features of deliberate opportunities to use evidence to make claims about examples/phenomena are rated on the protocol:

- The opportunities help students understand what the data represent,
- The opportunities facilitate students' interpretation/analysis of the data,
- Students use evidence to support their claims, and
- Students use evidence to critique claims.

The following two scenarios demonstrate the use of evidence to make claims about the targeted idea that ***in a contact push/pull interaction, the force ceases to exist as soon as contact between the interacting objects is lost.***

Scenario A: Using Evidence to Draw Conclusions and Make Claims

After completing an activity with a car and foam (pushing and pulling), the teacher asks students,

"When was a force acting on the car and how do you know?"

Several students share aloud:

Student 1: "A force was acting on the car when we were pushing on the foam, because that's when it was dented in."

Student 2: "A force was also acting on the car when we were pulling on the foam, because we could see the foam stretch out."

Student 3: "I agree with both of them. A force was acting on the car when we pushed and pulled on the foam. When we were actually touching the form in some way we could see it change shape."

Scenario B: Using Evidence to Draw Conclusions and Make Claims

After completing an activity with a car and foam (pushing and pulling), the teacher leads a large-group discussion and each student is asked to share an observation:

Student 1: "I saw the car speed up."

Student 2: "Pushing on the foam made the car go forward."

Student 3: "The car moved when we pulled on it."

Only the prompt in Scenario A is aligned to the targeted idea and requires students to consider their evidence in order to answer. Although students in Scenario B are asked to share an observation, they are not asked to analyze/interpret data or use evidence to support a claim. Therefore, Scenario A would receive a synthesis rating of a "4" on the protocol and Scenario B would be rated lower.

An annotated version of the Using Evidence to Draw Conclusions and Make Claims component is provided below. Each comment is keyed to the area or item to which it refers.

IV. Using Evidence to Draw Conclusions and Make Claims about the Examples/ Phenomena

A. Ratings of Key Indicators

☐ Check here if the observed instruction addressing the targeted idea did not include using of evidence to draw conclusions and make claims for the purpose of developing or reinforcing science ideas and skip to V. Sense Making of Targeted Idea.

Of the observed lessons that addressed the targeted idea, how many included opportunities to use evidence to draw conclusions and make claims about the examples/phenomena? _____

The first part of this section asks you to rate the extent to which several key features of using evidence to draw conclusions about the examples/phenomena were present, regardless of how well aligned these opportunities were to the targeted idea. The second part asks for your judgment of that alignment.

	Not at all			To a great extent
1. Deliberate opportunities provided to students to draw conclusions/make claims:				
a. helped students understand what the data represent.	1	2	3	4
b. facilitated students' interpretation/analysis of the data.	1	2	3	4
c. had students use evidence to support their claims about the examples/phenomena they were investigating.	1	2	3	4
d. had students use evidence to critique claims (made by them and/or others) about the examples/phenomena they were investigating.	1	2	3	4

IV. Using Evidence to Draw Conclusions and Make Claims about the Examples/Phenomena

A. Ratings of Key Indicators

NOTE: The ratings in the first part of this subsection (1.a.–d.) do not take into consideration how well aligned opportunities to surface students' prior knowledge were to the targeted idea.

The students must be the ones making the claims. If the teacher makes all the claims, observers should not complete this subsection. However, if the teacher helps students articulate their claims, this subsection should be rated.

1.a. Observers should consider whether or not the students understand representations being used to depict data (e.g., numbers from a measuring device, graphs, pictures). For example, if students are using a motion sensor to get a graph of speed vs. time, do they understand what the graph represents in relation to the movement of the actual object? The opportunities must also actually help students understand what the data represent, not just be intended to help. If students do not seem to understand what the data represent, this rating should be lowered.

1.b. Observers should consider whether students had opportunities to: (1) organize data and look for relationships and/or patterns, and (2) analyze/interpret data in order to make claims. Students' drawing conclusions is a key component of this indicator. If students have few or no have opportunities to draw conclusions from their data (e.g., the teacher tells students what the data mean), this rating should be lowered.

1.c. Observers should consider whether students had opportunities to use data as evidence to support their claims. This rating should not be lowered if students use evidence to support claims for part of the targeted idea (rather than the entire idea). If there were no instances where students used evidence to support claims, or only brief opportunities for a few students to do so, this rating should be lowered.[3]

1.d. Observers should consider whether students had opportunities to use data as evidence to critique claims. This indicator includes agreeing with a claim or disagreeing with a claim using evidence. This rating should not be lowered if students use evidence to critique claims about part of the targeted idea (rather than the entire idea). The key point here is students using evidence to examine others' claims; if students did not use evidence at all, or if there were only brief opportunities for a few students to do so, this rating should be lowered. Note, if all students arrive at the same conclusion based on the same evidence, it may not be

[3] The teacher may help the students articulate a claim as long as the claim comes from students' documented ideas (either verbally or in writing).

	Not at all			To a great extent
2. To what extent were the opportunities provided to students to draw conclusions/make claims about the examples/phenomena *aligned* with the targeted idea?	1	2	3	4

B. Synthesis Rating: Using Evidence to Draw Conclusions and Make Claims

The synthesis rating describes the sufficiency of using the evidence for understanding the targeted idea. The synthesis rating is not intended to be an arithmetic average of the individual indicator ratings. It is the observer's responsibility, with knowledge of the purposes of the science instruction and the experience of being there, to come up with a holistic rating. However, item 2 is a critical indicator; the synthesis rating cannot be higher than the rating of this indicator.

To what extent were opportunities for students to use evidence to draw conclusions and make claims likely *sufficient* for their **learning of the targeted idea**?

1	2	3	4
Not at all likely	Minimally likely	Moderately likely	Very likely

possible to critique claims. In this case, this indicator would be rated low, but would not affect the synthesis rating for this subsection.

2. Observers should consider the extent to which opportunities provided to students to draw conclusions/make claims about the examples/phenomena aligned with the targeted idea. The quality and amount of time devoted to drawing conclusions/making claims should not factor into this rating.

   If the opportunities for students to draw conclusions/make claims are aligned only with part of the targeted idea this indicator should still be rated high. However, the synthesis should be rated lower.

B. Synthesis Rating: Using Evidence to Draw Conclusions and Make Claims

Observers should indicate the extent to which opportunities for students to use evidence to draw conclusions and make claims were sufficient for their learning of the targeted idea. The synthesis rating cannot be higher than the rating for key indicator 2.

C. Supporting Evidence for Synthesis Rating

For each key indicator listed above, and the synthesis rating, provide 1) a description of the relevant instruction, 2) a statement of quality of the instruction for the indicator, and 3) the rationale and evidence for your statement of quality. Additionally, if anything else factored into your synthesis rating, please describe it below.

Description of Opportunities to Use Evidence to Draw Conclusions and Make Claims about the Targeted Idea

1.a. Deliberate opportunities provided to students to draw conclusions/make claims helped students understand what the data represent.

1.b. Deliberate opportunities provided to students to draw conclusions/make claims facilitated students' interpretation/analysis of the data.

C. Supporting Evidence for Synthesis Rating

Observers should thoroughly describe what happened in the instruction, as well as when (what day, which lesson) it occurred. In addition, for each key indicator listed in this subsection, observers should provide a detailed description of the evidence used to determine the ratings. For example:

Description of Opportunities to Use Evidence to Draw Conclusions and Make Claims about the Targeted Idea: *In three of the observed lessons, students had opportunities to use evidence to draw conclusions and make claims about the following targeted idea:*

A force is a push or pull interaction between two objects. Forces can vary in strength and direction.

On 11/17, students made claims about when forces were present, referring to pushes and pulls applied to a piece of foam. On 11/18, students again made claims about when forces were and were not present based on their experiences with a foam/car assembly. On 11/19, students made claims about the presence of non-contact forces based on engaging with magnets and paperclips, charged balloons and hair/bits of paper, and dropping a tennis ball.

1.a *There were no deliberate opportunities to help students understand what the data represented. However, the data that students generated were simply observations of straightforward phenomena, so this indicator did not factor into the synthesis rating.*

1.b. *To a minimal extent, the opportunities provided to students to draw conclusions/make claims facilitated their interpretation/analysis of the data. After students engaged with each phenomenon, the teacher led discussions of what the students had observed and what relationships they could infer. However, these discussions were often too brief or unsuccessful in helping students analyze their observational data. For example, after students experimented with pushing the foam/car configuration on 12/2, the teacher led a discussion about their observations related to when the force stopped acting on the car. Student responses were mainly focused on how fast the car was moving (rather than changes in speed or the changing shape of the foam) resulting in very little productive analysis of the data.*

1.c. ***Deliberate opportunities provided to students to draw conclusions/make claims had students use evidence to support their claims about the examples/phenomena they were investigating.***

1.d. ***Deliberate opportunities provided to students to draw conclusions/make claims had students use evidence to critique claims (made by them and/or others) about the examples/phenomena they were investigating.***

2. ***To what extent were the opportunities provided to students to draw conclusions/make claims about the examples/phenomena aligned with the targeted idea?***

Synthesis:

1.c. *To a minimal extent, students used evidence to support claims. Although on a few of occasions students did make claims based on evidence, on most occasions students either made only observations or made inaccurate claims not supported with appropriate evidence.*

1.d. *Students did not use evidence to critique claims (made by them and/or others) about the examples/phenomena they were investigating.*

2. *To a great extent, the opportunities provided to students to make claims about the phenomena were aligned with the targeted idea. Students had opportunities to make claims about what constitutes a force, when a force stops acting, and whether forces can act without contact.*

Synthesis

It is minimally likely that the opportunities for students to use evidence to draw conclusions and make claims were sufficient for their learning of the targeted idea. Although there were a few occasions when students made evidence-based claims, on most occasions students did not. For the most part, during the opportunities for students to make claims, they offered just their observations or made inaccurate claims without appropriate evidence. In addition, students' analysis of the data was inadequately facilitated by the teacher and they did not have an opportunity to critique each other's claims.

V. Sense Making of the Targeted Idea

Learning theory suggests that effective science instruction requires opportunities for students to make sense of ideas they have engaged with. Sense making may include helping students make connections between what they did in a lesson and what they were intended to learn. Sense making may also involve asking students to reflect on their prior ideas and consider how their thinking has changed over time, helping students connect targeted science ideas to what they have leaned previously, and applying new concepts to new contexts. Because students are usually unable to make such connections on their own, the teacher is largely responsible for helping students make sense of their science learning experiences through activities such as purposeful questioning and guided class discussions.

Four key features of sense making are rated on the protocol, deliberate opportunities for students to:

- Connect what they did in the instruction to the targeted idea,
- Use evidence from multiple phenomena to support/critique claims about the targeted idea,
- Compare their emerging understanding of the targeted idea to their prior ideas, and
- Compare their emerging understanding of the targeted idea to other science ideas they already know.

Two scenarios are shown below depicting sense making opportunities around the targeted idea that ***in a contact push/pull interaction, the force ceases to exist as soon as contact between the interacting objects is lost***.

Scenario A: Sense Making

To conclude the lesson, the teacher leads a class discussion about the students' observations and their implications for when a force was acting on a toy car. He then asks the students to answer the following questions in their notebooks:

> *Based on your experience pushing the toy car, was a force pushing the car forward while your finger was **not** touching the car, in between your pushes? How do you know?*

Scenario B: Sense Making

To conclude the lesson, the teacher asks the students:

> *"What is one thing you learned today?"*

The students individually answer the question in their notebooks, and then discuss their answers as a whole group.

Although students in both scenarios are prompted to consider what they learned during the lesson, only the prompt in Scenario A requires students to use evidence from the activities they engaged with to explain why they thought a force was or was not acting on the toy car in between pushes (one of the three types of sense making). In contrast, the prompt in Scenario B is very open-ended and does not require students elaborate on their responses. Therefore Scenario A would receive a synthesis rating of a "4" on the protocol while Scenario B would be rated lower.

An annotated version of the Sense Making component is provided below. Each comment is keyed to the area or item to which it refers.

V. Sense Making of the Targeted Idea

A. Ratings of Key Indicators

☐ Check here if the observed instruction addressing the targeted idea did not include opportunities for students to connect what they did in the instruction to science ideas for the purpose of developing or reinforcing those ideas.

Of the observed lessons that addressed the targeted idea, how many included opportunities for sense making? _____

The first part of this section asks you to rate the extent to which several key features of sense making were present, regardless of how well aligned these opportunities were to the targeted idea. The second part asks for your judgment of that alignment.

	Not at all			To a great extent
1. Deliberate opportunities for sense making:				
a. helped students connect what they did in the instruction to the targeted idea.	1	2	3	4
b. had students use evidence from multiple phenomena to support/critique claims about the targeted idea.	1	2	3	4
c. helped students compare their emerging understanding of the targeted idea to their prior ideas.	1	2	3	4
d. helped students compare their emerging understanding of the targeted idea to other science ideas that they already know.	1	2	3	4

	Not at all			To a great extent
2. To what extent did the sense making opportunities *address* the targeted idea?	1	2	3	4

B. Synthesis Rating: Sense Making of the Targeted Idea

The synthesis rating describes the sufficiency of the opportunities to make sense of the targeted idea. The synthesis rating is not intended to be an arithmetic average of the individual indicator ratings. It is the observer's responsibility, with knowledge of the purposes of the science instruction and the experience of being there, to come up with a holistic rating. However, item 2 is a critical indicator; the synthesis rating cannot be higher than the rating of this indicator.

To what extent were opportunities for students to make sense likely *sufficient* for their **learning of the targeted idea**?

1	2	3	4
Not at all likely	Minimally likely	Moderately likely	Very likely

V. Sense Making of the Targeted Idea

A. Ratings of Key Indicators

NOTE: Unlike in previous sections, all of the indicators in the Sense Making section are related to the targeted idea. Thus, if for any indicator the sense-making opportunities were not aligned to the targeted idea, the rating should be lowered.

1.a. Observers should consider whether students used evidence to connect what they did in the instruction to the targeted idea. If few or no students make connections, this rating should be lowered.

1.b. Observers should consider whether students are supporting and critiquing claims by using evidence from more than one phenomenon. If students do not use evidence to support/critique claims, or use evidence from only a single phenomenon, this rating should be lowered.

1.c. Observers should consider whether students revisited their prior ideas (e.g., ideas written in their notebooks) to see how they changed and/or stayed the same. If few or no students revisit their prior thinking, this rating should be lowered.

1.d. Observers should consider whether students are taking what they know about the targeted idea, and applying it to science ideas they already know. If few or no students apply their emerging understanding to previously learned ideas, this rating should be lowered.

2. Observers should consider the extent to which sense making opportunities addressed the targeted idea. The quality and amount of time devoted to sense making should not factor into this rating.

If the opportunities for sense making address only part of the targeted idea, this indicator should be rated high. However, the synthesis should be rated lower.

B. Synthesis Rating: Sense Making of the Targeted Idea

Observers should consider the extent to which opportunities for students to make sense were sufficient for their learning of the targeted idea.

Because of the nature of sense-making experiences, it is not expected that all features of sense making necessarily happen for every idea. Therefore, if one or two ratings in 1.a.–d. are lowered, it does not necessarily mean the synthesis should be lowered. However, key indicator 2 is critical for deciding upon a synthesis rating in this subsection; the synthesis cannot be higher than the rating for key indicator 2. In addition, the synthesis rating for sense making cannot be higher than the synthesis

C. Supporting Evidence for Synthesis Rating

For each key indicator listed above, and the synthesis rating, provide 1) a description of the relevant instruction, 2) a statement of quality of the instruction for the indicator, and 3) the rationale and evidence for your statement of quality. Additionally, if anything else factored into your synthesis rating, please describe it below.

Description of Sense-making Opportunities Aligned to the Targeted Idea

1.a Deliberate opportunities for sense making helped students connect what they did in the instruction to the targeted idea.

1.b. Deliberate opportunities for sense making had students use evidence from multiple phenomena to support/critique claims about the targeted idea.

rating was for using evidence to draw conclusions and make claims about the examples/phenomena.

C. Supporting Evidence for Synthesis Rating

Observers should thoroughly describe what happened in the instruction, as well as when (what day, which lesson) it occurred. In addition, for each key indicator listed in this subsection, observers should provide a detailed description of the evidence used to determine the ratings. For example:

Description of Sense-making Opportunities Aligned to the Targeted Idea: *Students had opportunities in two of the observed lessons to make sense about the following targeted idea:*

A force is a push or pull interaction between two objects. Forces can vary in strength and direction.

On 11/17, the teacher led a class discussion about whether there was a force on the toy car when it started to move and whether pulls were forces. On 11/18, students considered when a force stops acting on an object as they wrote down their responses to the making sense questions, and a few students shared their responses with the class.

1.a. *To a moderate extent, opportunities for sense making helped students connect what they did in the instruction to the targeted idea. The making sense questions that the students engaged with helped them connect what they did in the instruction to the targeted idea. All students had the opportunity to consider when a force stops acting on an object as they individually wrote responses to the making sense questions. However, because only a few students participated in the class discussions, the extent to which most students made sense of the activities was not clear.*

1.b. *Opportunities for sense-making had students use evidence from multiple phenomena to support/critique claims about the targeted idea to a minimal extent. In answering the making sense questions related to forces on a toy car, students had experience with multiple phenomena in the lesson to draw upon. However, only one student used evidence to support his claim, and only when answering the first question. Students were not expected to draw from multiple phenomena to answer the making sense questions about when a force stops acting.*

1.c. ***Deliberate opportunities for sense making helped students compare their emerging understanding of the targeted idea to their prior ideas.***

1.d. ***Deliberate opportunities for sense making helped students compare their emerging understanding of the targeted idea to other science ideas that they already know.***

2. ***To what extent did the sense-making opportunities address the targeted idea?***

Synthesis:

1.c. *Sense-making opportunities did not help students compare their emerging understanding of the targeted idea to their prior ideas. Students were not asked to refer back to their initial ideas and compare those ideas to their emerging understanding.*

1.d. *There were not any deliberate opportunities to help students compare their emerging understanding of the targeted idea to other science ideas that they already knew.*

2. *The sense-making opportunities addressed the target idea to a great extent. The sense-making opportunities on 11/16 addressed the targeted idea that pushes and pulls are forces by prompting students to consider the motion of a toy car when they applied a force to it. On 11/17 the making sense questions required students to use what they had learned in the toy car activity to determine when a force stops acting on an object.*

Synthesis

It is minimally likely that the opportunities for students to make sense were sufficient for their learning of the targeted idea. The sense-making opportunities were highly aligned with the targeted idea and helped students connect what they did in the instruction to the targeted idea. However, students generally did not use evidence from multiple phenomena to support/critique claims about the targeted idea. Further, students were not required to compare their emerging understanding of the targeted idea to their prior ideas. Additionally, because only a few students participated in the class discussions, the extent to which most students made sense of the activities was not clear.

Section Three: Ratings for Classroom Culture

Classroom Culture is the final section of the AIM COP. This section provides an opportunity to rate the extent to which key features of culture were present in the observed instruction (e.g., high expectations for student learning, collegial working relationships among students). In addition, a synthesis rating captures the extent to which the classroom culture interfered with or facilitated student learning of the targeted idea. As in Section Two, culture ratings are supported by detailed rationales with illustrative examples.

This section also includes a checkbox for the observer to indicate that the classroom culture varied greatly across the observed lessons. However, if this box is checked, the remainder of the section should still be completed, averaging across all lessons. Observers should use the available evidence to provide an average rating for each key indicator.

An annotated version of Section Three (Ratings for Classroom Culture) is provided below. The annotations include sample questions that observers should ask themselves in order to assign ratings. Each comment is keyed to the area or item to which it refers.

SECTION THREE: RATINGS FOR CLASSROOM CULTURE

A. Ratings of Key Indicators

		Not at all			To a great extent
1.	There was a climate of respect for students' ideas, contributions, and emerging knowledge.	1	2	3	4
2.	Students appeared motivated to learn.	1	2	3	4
3.	There appeared to be high expectations for student learning.	1	2	3	4
4.	Interactions reflected collegial working relationships among students.	1	2	3	4
5.	The instruction built on the collective knowledge of the students.	1	2	3	4
6.	Students demonstrated a willingness to share ideas and take intellectual risks.	1	2	3	4

☐ The culture in this classroom varied greatly across the lessons.

B. Synthesis Rating: Classroom Culture

The synthesis rating is not intended to be an arithmetic average of the individual indicator ratings. It is the observer's responsibility, with knowledge of the purposes of the science instruction and the experience of being there, to come up with a holistic rating.

How did the classroom culture affect student **learning of the targeted idea**?

1	2	3	4
Greatly interfered	Somewhat interfered	Somewhat facilitated	Greatly facilitated

A. Ratings of Key Indicators

For each key indicator, observers should consider the following as evidence for their ratings:

Examples to consider in rating this indicator:

1. Were students encouraged to contribute to class discussions and activities? Was the teacher respectful of students' ideas, etc.? Did the teacher facilitate constructive and respectful exchange of ideas among students?

2. Did students willingly participate in class discussions and activities throughout the instruction? Was student behavior generally on-task?

3. Did the teacher take steps to ensure that all (or almost all) students participated in class discussions and activities? Did the teacher take steps to help all (or almost all) students learn the targeted science ideas (e.g., re-teaching ideas, providing additional examples/practice)? Did the teacher prompt students to explain/provide reasons for how they were thinking?

4. Did the students work well together (as appropriate)? Did students listen to each other's ideas, both in large group and small group discussions?

5. Were student questions/ideas addressed during the instruction? Were student questions/ideas taken into account during instruction? Did the teacher summarize students' ideas and/or make common patterns of thinking explicit (e.g., "As I walked around the room, these are the types of things I heard you saying…")?

6. Did students contribute to class discussions and activities? Did students disagree with one another's ideas?

B. Synthesis Rating: Classroom Culture

Observers should indicate the extent to which the classroom culture affected student learning of the targeted idea.

C. Supporting Evidence for Synthesis Rating

For each key indicator listed above, please provide a *detailed description of the evidence* you used to determine your rating.

1. ***There was a climate of respect for students' ideas, contributions, and emerging knowledge.***
2. ***Students appeared motivated to learn.***
3. ***There appeared to be high expectations for student learning.***
4. ***Interactions reflected collegial working relationships among students.***
5. ***The instruction built on the collective knowledge of the students.***

C. Supporting Evidence for Synthesis Rating

For each key indicator, observers should provide a summary of the evidence used to determine the rating as well as 1 or 2 illustrative examples (e.g., conversations, scenarios).

1. *There was a climate of respect for students' ideas, contributions, and emerging knowledge to a moderate extent. Students were generally encouraged to participate in partner/small group discussions. The teacher also regularly called on select students to participate in large group discussions and activities.*

2. *Students appeared motivated to learn to a great extent. Students generally attempted to participate in pair and whole class discussions. Students often raised their hands to vote for outcomes of scenarios or to indicate agreement/disagreement with ideas. Additionally, student behavior was on-task and they actively participated in the hands-on activities.*

3. *There appeared to be high expectations for student learning to a minor extent. When the teacher tried to involve multiple students in the class discussion, she usually asked questions that required one word answers and did not prompt students to explain reasons for their answers. Additionally, on several occasions the teacher carried on a conversation with the one or two "smartest" kids in the room while the other students listened. Even when other students raised their hands to volunteer responses, the teacher called on the select students.*

4. *Interactions reflected collegial working relationships among students to a great extent. For example, on 11/19, students worked well together to perform hands-on activities with magnets, balls, and balloons. On 11/25, students collaborated in small groups to experiment with CD pucks. On 12/5, students listened to each other's ideas and respectfully agreed/disagreed about the order in which three cars would finish a race.*

5. *The instruction built on the collective knowledge of the students to a minor extent. For example, on 11/4/11 the teacher related the calculation for speed to a school race that most students were familiar with. Additionally, the teacher occasionally tried to build on or summarize patterns in student ideas. For example, on 11/16/11 the teacher summarized student observations regarding the motion of CD pucks. However, attempts to build on the collective knowledge of the students occurred infrequently and inconsistently.*

6. Students demonstrated a willingness to share ideas and take intellectual risks.

Synthesis:

6. *The students demonstrated a willingness to share ideas and take intellectual risks to a great extent. Most students regularly discussed scenarios/questions with partners and many students raised their hands to share answers aloud. Additionally, on 11/25 the students challenged the teacher's assertion that the CD puck was moving toward the heavier cup during the class demonstration (they were correct).*

Synthesis

The classroom culture moderately facilitated students' opportunity to learn the targeted idea. Students appeared motivated to learn and worked well with one another. However, the teacher frequently carried out conversations with the one or two "smart" kids even though other students raised their hands. In addition, the teacher did not regularly ask students to provide reasons for how they were thinking and instead accepted brief one or two word answers before moving on.

AIM
Classroom Observation Protocol

SECTION ONE: DESCRIPTION OF THE INSTRUCTION

I. Overview of Instruction

Teacher ID: _____
Researcher completing this COP: _____
Targeted Idea: _____
Number of lessons observed that addressed the targeted idea: _____

Date of Observed Lesson about Idea						
Name of Observer						
1. For each observed lesson that addressed the targeted idea, please give the length of the lesson, teacher's stated purpose for the lesson,[1] approximate length of time spent on the targeted idea, the type of instructional materials used for teaching the targeted idea (textbook, science kit, teacher created, materials from the PD, etc.), and the title (if commercially published).						
Length of lesson						
Teacher's stated purpose for the lesson						
Approximate length of time spent on targeted idea						
Instructional materials used for teaching the targeted idea						
2. Of the time spent on the targeted idea in each observed lesson approximately how long did students spend working in the following arrangements (in minutes)?						
Individually						
Small Groups						
Whole Class						
3. For the time spent on the targeted idea in each observed lesson, please note if each of the following occurred. (Write "yes" or "no" in each box.)						
Teacher lecture[2]						
Teacher demonstration						
Class discussion[3]						
Writing in journal, recording answers to open-ended questions, etc.						
Completing worksheets, reading textbooks, doing homework, etc.						
Hands-on activities						
Other						

[1] Observers may ask the teacher for the purpose of the lesson, find it posted in the room, or the teacher may state to the class the purpose for the lesson.

[2] A lecture is when the instructor is providing information (facts, definitions, etc.) to the learners as opposed to this information coming from hands-on activities or a written source. Delivering instructions for an activity should NOT be considered lecture.

[3] Discussion is when learners are talking with each other or with the teacher. Discussion can occur as part of a lecture, demonstration or hands-on activity.

II. Comments

Your ratings in Section Two will focus on the instruction about the targeted idea. Use this space for any information on the lesson and/or classroom that you consider so salient that it needs to get "on the table" right away to help explain your ratings.

III. Description of Instruction about Targeted Idea

Describe what happened in the observed instruction. The description should include enough rich detail that readers have a sense of what happened during the instruction, but should not include a judgment of quality. Include:

- A description of the structure and flow of the instruction related to the targeted idea (i.e., the nature and progression of the activities).
- A description of the level of intellectual engagement (i.e., who is doing the intellectual work—the teacher, some students, most students).

SECTION TWO: RATINGS FOR STUDENT OPPORTUNITY TO LEARN

I. Science Content[4]

A. Ratings of Key Indicators

The first part of this section asks you to rate the extent to which several key features of the science content were present, regardless of how well aligned the content was to the targeted idea. The second part asks for your judgment of that alignment.

	Not at all			To a great extent
1. The science content addressed in this instruction was:				
a. accurate.	1	2	3	4
b. developmentally appropriate.	1	2	3	4

	Not at all			To a great extent
2. To what extent was the science content addressed in this instruction *aligned* with the targeted idea?	1	2	3	4

B. Synthesis Rating: Science Content

The synthesis rating is not intended to be an arithmetic average of the individual indicator ratings. It is the observer's responsibility, with knowledge of the purposes of the science instruction and the experience of being there, to come up with a holistic rating. For the science content synthesis rating, if any indicator is rated a "1," the synthesis rating must a "1"; a synthesis rating of "4" can be given only if each key indicator is a "4."

For the purposes of this lesson and considering this group of students, the science content **related to the targeted idea** was:

1	2	3	4
Inappropriate (i.e., unaligned, inaccurate, and/or developmentally inappropriate)	Minimally appropriate	Moderately appropriate	Very appropriate (i.e., aligned, accurate, and developmentally appropriate, and *covers the entire targeted idea*).

[4] Science Content can include disciplinary science concepts and/or ways of knowing in science.

C. Supporting Evidence for Synthesis Rating

Description of Science Content Related to the Targeted Idea

1.a. The science content addressed in this instruction was accurate.

1.b. The science content addressed in this instruction was developmentally appropriate.

2. To what extent was the science content addressed in this instruction aligned with the targeted idea?

Synthesis:

II. Opportunities to Surface Prior Knowledge

A. Ratings of Key Indicators

☐ Check here if the observed instruction addressing the targeted idea did not include opportunities to surface students' prior knowledge and skip to III. Engaging with Examples/Phenomena.

Of the observed lessons that addressed the targeted idea, how many included opportunities to surface prior knowledge? _____

The first part of this section asks you to rate the extent to which several key features of surfacing students' prior knowledge engaged students, regardless of how well aligned the surfacing was to the targeted idea. The second part asks for your judgment of that alignment.

	Not at all			To a great extent
1. Deliberate opportunities provided to surface students' prior knowledge:				
a. were structured/implemented so that students would be aware of their own prior knowledge.	1	2	3	4
b. surfaced students' reasons for how they were thinking.	1	2	3	4
c. had students record aspects of their prior knowledge.	1	2	3	4
d. had students make public aspects of their prior knowledge.	1	2	3	4
e. allowed students' ideas to be surfaced without judgment.	1	2	3	4

	Not at all			To a great extent
2. To what extent were the opportunities provided to surface students' prior knowledge *aligned* with the targeted idea?	1	2	3	4

B. Synthesis Rating: Surfacing Prior Knowledge

The synthesis rating describes the sufficiency of the opportunities to surface prior knowledge for understanding the targeted idea. The synthesis rating is not intended to be an arithmetic average of the individual indicator ratings. It is the observer's responsibility, with knowledge of the purposes of the science instruction and the experience of being there, to come up with a holistic rating. However, item 2 is a critical indicator; the synthesis rating cannot be higher than the rating of this indicator.

To what extent were opportunities for students to surface their prior knowledge likely *sufficient* for their **learning of the targeted idea**?

1	2	3	4
Not at all likely	Minimally likely	Moderately likely	Very likely

C. Supporting Evidence for Synthesis Rating

For each key indicator listed above, and the synthesis rating, provide 1) a description of the relevant instruction, 2) a statement of quality of the instruction for the indicator, and 3) the rationale and evidence for your statement of quality. Additionally, if anything else factored into your synthesis rating, please describe it below.

Description of Opportunities to Surface Prior Knowledge of the Targeted Idea

1.a. Deliberate opportunities provided to surface students' prior knowledge were structured/implemented so that students would be aware of their own prior knowledge.

1.b. Deliberate opportunities provided to surface students' prior knowledge surfaced students' reasons for how they were thinking.

1.c. Deliberate opportunities provided to surface students' prior knowledge had students record aspects of their prior knowledge.

1.d. Deliberate opportunities provided to surface students' prior knowledge had students make public aspects of their prior knowledge.

1.e. Deliberate opportunities provided to surface students' prior knowledge allowed students' ideas to be surfaced without judgment.

2. To what extent were the opportunities provided to surface students' prior knowledge aligned with the targeted idea?

Synthesis:

III. Engaging with Examples/Phenomena

A. Ratings of Key Indicators

☐ Check here if the observed instruction addressing the targeted idea did not include opportunities for students to engage with examples/phenomena for developing or reinforcing science ideas and skip to Using Evidence to Draw Conclusions and Make Claims about the Examples/Phenomena.

Of the observed lessons that addressed the targeted idea, how many included opportunities to engage with examples/ phenomena? _____

The first part of this section asks you to rate the extent to which several key features of engaging students with examples/phenomena were present, regardless of how well aligned the examples/phenomena were to the targeted idea. The second part asks for your judgment of that alignment.

	Not at all			To a great extent
1. Deliberate opportunities to engage with examples/phenomena:				
a. included examples/phenomena *that were accessible to students*.	1	2	3	4
b. focused students on *relevant aspects* of the examples/ phenomena.	1	2	3	4
c. had students describe and/or keep a record of the processes they used/data they generated.	1	2	3	4

	Not at all			To a great extent
2. To what extent did the examples/phenomena *provide evidence for* the targeted idea?	1	2	3	4

B. Synthesis Rating: Engaging with Examples/Phenomena

The synthesis rating describes the sufficiency of the students' engagement with examples/phenomena for understanding the targeted idea. The synthesis rating is not intended to be an arithmetic average of the individual indicator ratings. It is the observer's responsibility, with knowledge of the purposes of the science instruction and the experience of being there, to come up with a holistic rating. However, item 2 is a critical indicator; the synthesis rating cannot be higher than the rating of this indicator.

To what extent were opportunities for students to engage with examples/phenomena likely *sufficient* for their **learning of the targeted idea**?

1	2	3	4
Not at all likely	Minimally likely	Moderately likely	Very likely

C. Supporting Evidence for Synthesis Rating

For each key indicator listed above, and the synthesis rating, provide 1) a description of the relevant instruction, 2) a statement of quality of the instruction for the indicator, and 3) the rationale and evidence for your statement of quality. Additionally, if anything else factored into your synthesis rating, please describe it below.

Description of Opportunities to Engage with Examples/Phenomena Aligned to the Targeted Idea

1.a. Deliberate opportunities to engage with examples/phenomena included examples/phenomena that were accessible to students.

1.b. Deliberate opportunities to engage with examples/phenomena focused students on relevant aspects of the examples/ phenomena.

1.c. Deliberate opportunities to engage with examples/phenomena had students describe and/or keep a record of the processes they used/data they generated.

2. To what extent did the examples/phenomena provide evidence for the targeted idea?

Synthesis:

IV. Using Evidence to Draw Conclusions and Make Claims about the Examples/Phenomena

A. Ratings of Key Indicators

☐ Check here if the observed instruction addressing the targeted idea did not include using of evidence to draw conclusions and make claims for the purpose of developing or reinforcing science ideas and skip to V. Sense Making of Targeted Idea.

Of the observed lessons that addressed the targeted idea, how many included opportunities to use evidence to draw conclusions and make claims about the examples/phenomena? _____

The first part of this section asks you to rate the extent to which several key features of using evidence to draw conclusions about the examples/phenomena were present, regardless of how well aligned these opportunities were to the targeted idea. The second part asks for your judgment of that alignment.

	Not at all			To a great extent
1. Deliberate opportunities provided to students to draw conclusions/make claims:				
a. helped students understand what the data represent.	1	2	3	4
b. facilitated students' interpretation/analysis of the data.	1	2	3	4
c. had students use evidence to support their claims about the examples/phenomena they were investigating.	1	2	3	4
d. had students use evidence to critique claims (made by them and/or others) about the examples/phenomena they were investigating.	1	2	3	4

	Not at all			To a great extent
2. To what extent were the opportunities provided to students to draw conclusions/make claims about the examples/phenomena *aligned* with the targeted idea?	1	2	3	4

B. Synthesis Rating: Using Evidence to Draw Conclusions and Make Claims

The synthesis rating describes the sufficiency of using the evidence for understanding the targeted idea. The synthesis rating is not intended to be an arithmetic average of the individual indicator ratings. It is the observer's responsibility, with knowledge of the purposes of the science instruction and the experience of being there, to come up with a holistic rating. However, item 2 is a critical indicator; the synthesis rating cannot be higher than the rating of this indicator.

To what extent were opportunities for students to use evidence to draw conclusions and make claims likely *sufficient* for their **learning of the targeted idea**?

1	2	3	4
Not at all likely	Minimally likely	Moderately likely	Very likely

C. Supporting Evidence for Synthesis Rating

For each key indicator listed above, and the synthesis rating, provide 1) a description of the relevant instruction, 2) a statement of quality of the instruction for the indicator, and 3) the rationale and evidence for your statement of quality. Additionally, if anything else factored into your synthesis rating, please describe it below.

Description of Opportunities to Use Evidence to Draw Conclusions and Make Claims

1.a. Deliberate opportunities provided to students to draw conclusions/make claims had students reflect, individually or in groups, on the meaning of the data (i.e., what the data represent).

1.b. Deliberate opportunities provided to students to draw conclusions/make claims facilitated students' interpretation/analysis of the data.

1.c. Deliberate opportunities provided to students to draw conclusions/make claims had students use evidence to support their claims about the examples/phenomena they were investigating.

1.d. Deliberate opportunities provided to students to draw conclusions/make claims had students use evidence to critique claims (made by them and/or others) about the examples/phenomena they were investigating.

2. To what extent were the opportunities provided to students to draw conclusions/make claims about the examples/phenomena aligned with the targeted idea?

Synthesis:

V. Sense Making of the Targeted Idea

A. Ratings of Key Indicators

☐ Check here if the observed instruction addressing the targeted idea did not include opportunities for students to connect what they did in the instruction to science ideas for the purpose of developing or reinforcing those ideas.

Of the observed lessons that addressed the targeted idea, how many included opportunities for sense making? _____

The first part of this section asks you to rate the extent to which several key features of sense making were present, regardless of how well aligned these opportunities were to the targeted idea. The second part asks for your judgment of that alignment.

	Not at all			To a great extent
1. Deliberate opportunities for sense making:				
a. helped students connect what they did in the instruction to the targeted idea.	1	2	3	4
b. had students use evidence from multiple phenomena to support/critique claims about the targeted idea.	1	2	3	4
c. helped students compare their emerging understanding of the targeted idea to their prior ideas.	1	2	3	4
d. helped students compare their emerging understanding of the targeted idea to other science ideas that they already know.	1	2	3	4

	Not at all			To a great extent
2. To what extent did the sense making opportunities *address* the targeted idea?	1	2	3	4

B. Synthesis Rating: Sense Making of the Targeted Idea

The synthesis rating describes the sufficiency of the opportunities to make sense of the targeted idea. The synthesis rating is not intended to be an arithmetic average of the individual indicator ratings. It is the observer's responsibility, with knowledge of the purposes of the science instruction and the experience of being there, to come up with a holistic rating. However, item 2 is a critical indicator; the synthesis rating cannot be higher than the rating of this indicator.

To what extent were opportunities for students to make sense likely *sufficient* for their **learning of the targeted idea**?

1	2	3	4
Not at all likely	Minimally likely	Moderately likely	Very likely

C. Supporting Evidence for Synthesis Rating

For each key indicator listed above, and the synthesis rating, provide 1) a description of the relevant instruction, 2) a statement of quality of the instruction for the indicator, and 3) the rationale and evidence for your statement of quality. Additionally, if anything else factored into your synthesis rating, please describe it below.

Description of Sense-Making Opportunities Aligned to the Targeted Idea

1.a. Deliberate opportunities for sense-making helped students connect what they did in the instruction to the targeted idea.

1.b. Deliberate opportunities for sense-making had students use evidence from multiple phenomena to support/critique claims about the targeted idea.

1.c. Deliberate opportunities for sense-making helped students compare their emerging understanding of the targeted idea to their prior ideas.

1.d. Deliberate opportunities for sense-making helped students compare their emerging understanding of the targeted idea to other science ideas that they already know.

2. To what extent did the sense-making opportunities address the targeted idea?

Synthesis:

SECTION THREE: RATINGS FOR CLASSROOM CULTURE

A. Ratings of Key Indicators

		Not at all			To a great extent
1.	There was a climate of respect for students' ideas, contributions, and emerging knowledge.	1	2	3	4
2.	Students appeared motivated to learn.	1	2	3	4
3.	There appeared to be high expectations for student learning.	1	2	3	4
4.	Interactions reflected collegial working relationships among students.	1	2	3	4
5.	The instruction built on the collective knowledge of the students.	1	2	3	4
6.	Students demonstrated a willingness to share ideas and take intellectual risks.	1	2	3	4

☐ The culture in this classroom varied greatly across the lessons.

B. Synthesis Rating: Classroom Culture

The synthesis rating is not intended to be an arithmetic average of the individual indicator ratings. It is the observer's responsibility, with knowledge of the purposes of the science instruction and the experience of being there, to come up with a holistic rating.

How did the classroom culture affect student **learning of the targeted idea**?

1	2	3	4
Greatly interfered	Somewhat interfered	Somewhat facilitated	Greatly facilitated

C. Supporting Evidence for Synthesis Rating

For each key indicator listed above, please provide a *detailed description of the evidence* you used to determine your rating.

Description of the Classroom Culture Across All Observed Lessons

1. ***There was a climate of respect for students' ideas, contributions, and emerging knowledge.***
2. ***Students appeared motivated to learn.***
3. ***There appeared to be high expectations for student learning.***
4. ***Interactions reflected collegial working relationships among students.***
5. ***The instruction built on the collective knowledge of the students.***
6. ***Students demonstrated a willingness to share ideas and take intellectual risks.***

Synthesis: