Assessment Overview







twig SCIENCE

NGSS/3-D Science Assessment System

Twig Science supports educators to deliver the three-dimensional learning required by the NGSS and other 3-D science standards, and it provides a three-dimensional assessment system that allows educators to evaluate student attainment of Performance Expectations.

The assessment strategies measure students' knowledge and ability, favor performance tasks over rote memorization, and include multiple measures of performance, such as written assignments, collaborative engineering design challenges, and oral presentations.

There are frequent opportunities to evaluate student progress against the standards within each module. The Module Assessment Overviews show where and how each standard is assessed.

The Twig Science assessment system includes diagnostic Pre-Assessments (known as Pre-Explorations), formative Reflect Activities, rich Performance Tasks, summative Multiple Choice Assessments, and summative, performance-based Benchmark Assessments, which were developed in partnership with the SCALE Science at WestEd SCALE team.

Pre-Explorations (All Grades)

The Pre-Explorations assess students' prior knowledge and identify the misconceptions they have about the module content. They also assess students' ability to understand and apply the Science and Engineering Practices (SEPs) and Crosscutting Concepts (CCCs). Each module contains a number of Pre-Explorations, embedded at strategic points.

Students complete the Pre-Explorations in their Twig Books and Twig Journals, which employ grade-appropriate strategies to elicit understanding. Pre-Explorations are quick to administer, and educators are supported to track progression of student understanding, with suggestions for how to tailor instruction if misconceptions persist. Pre-Explorations have been included in Twig Science because assessing prior knowledge ranks highly in international meta-analysis of the impacts on student attainment.

Examples of Pre-Explorations

Grade 1



Grade 8



Reflect Activities (All Grades)

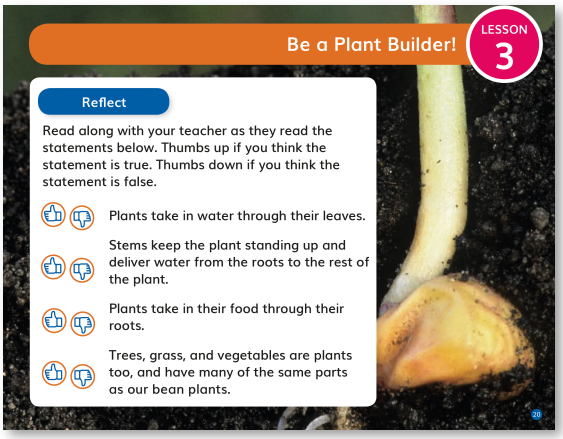
If students understood everything they were taught, then assessment would not be necessary. Unfortunately, this is not the case, which is why opportunities for formative assessment are embedded throughout the program, and suggested questions are frequently included in the Teacher Edition to gauge student understanding.

It is vital to include regular activities that give students the opportunity to demonstrate what they do and do not know, as this helps to guide future lesson planning. In many classrooms, formative assessment is improvised on the spot—in Twig Science, the activities are thought through to ensure they link well to the lesson objectives and the standards.

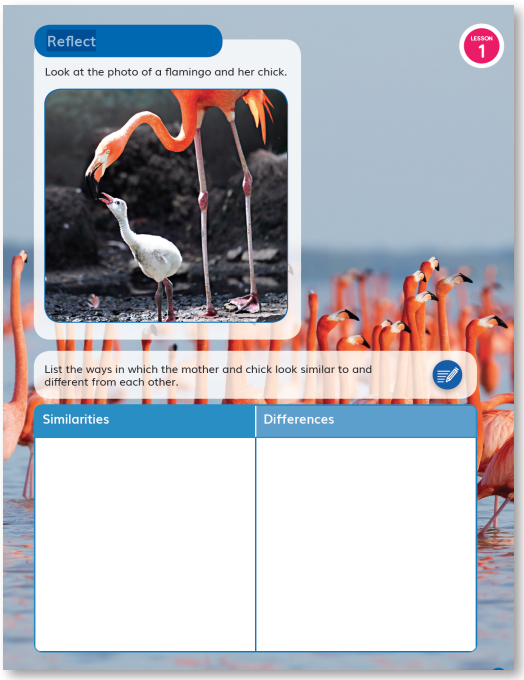
The Reflect Activities are grade-appropriate, but vary in approach. They include KLEW charts, short written reflections, annotated sketches, peer-to-peer discussions, and graphic checklists. They offer students the chance to self-reflect on their own growing understanding of the Driving Question and Module Phenomenon.

Most Reflect Activities are short and easy to implement, which cuts down on the grading burden. Since educators plan and deliver at least 5 hours of content a day, this approach gives them a quick win at the end of a lesson, which can also be a talking point at the start of the next. There is evidence to show that this kind of approach can have a meaningful impact on student attainment, especially when it is linked to quality feedback that moves learning on.

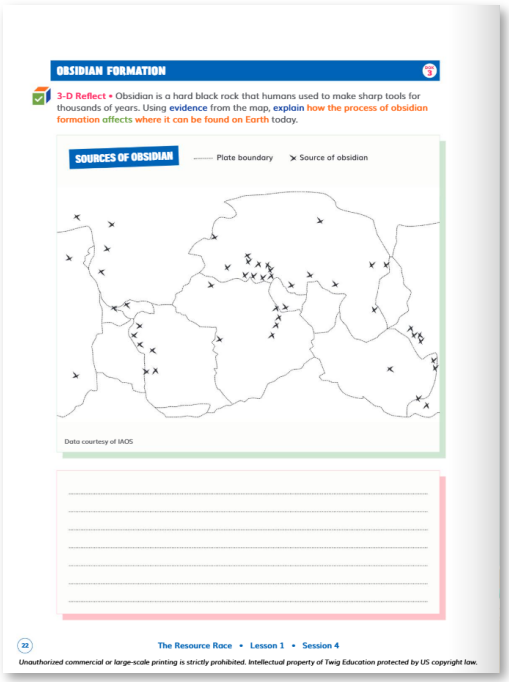
Examples of Reflect Activities
Grade 1



Grade 3



Grade 7



Performance Tasks (All Grades)

The Performance Tasks have been carefully designed to captivate students' imaginations, connect their classroom experiences to real-world problems, and reflect how engineers and scientists behave in the field. They allow students to apply the CCCs, SEPs, and their knowledge of the DCIs to demonstrate their growing mastery of the Performance Expectations, as well as their understanding of the Module Phenomenon.

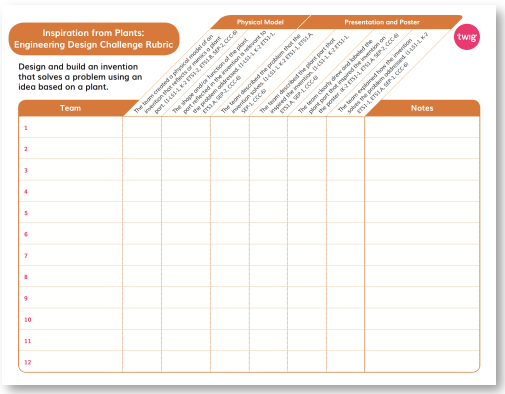
These three-dimensional Performance Tasks target discrete sets of Performance Expectations, and deliver data that educators can use to tailor their instruction to support students to meet and exceed the standards. They typically unfold over a number of lessons, score a portfolio of work, and include several components, such as research and written tasks, engineering design challenges, and oral presentations.

Rubrics support educators to assess the level of student attainment in relation to the Performance Expectations. The rubrics grow in complexity as the grades progress, from simple checklists to multi-leveled rubrics: Emerging, Developing, Proficient, and Advanced. From the end of Grade 2, these rubrics are shared with students, so they have a clear expectation of what success looks like. Students are supported to self-assess their own work and that of their peers. There is strong evidence that this strategy has a meaningful impact on student attainment. Therefore, in many instances in Twig Science, students are asked to compare their work with

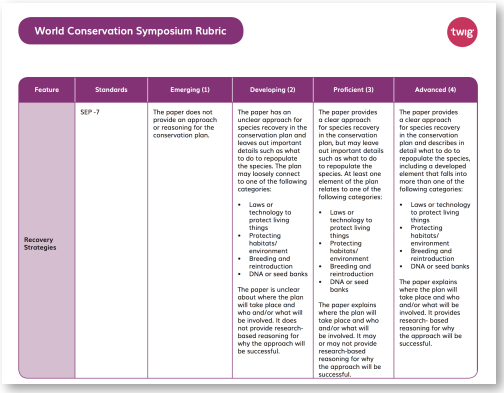
that of peers, and discuss where they think they are on the rubric collaboratively. This is a very good way for students to develop an understanding of what quality work looks like, and will be a vital skill when students reach the workplace.

Examples of Performance Task Rubrics

Grade 1



Grade 6



Benchmark Assessments (Grades 3–8)

The Twig Science Benchmark Assessments were developed in partnership with the SCALE Science at WestEd SCALE team. They are underpinned by research carried out by the Stanford NGSS Assessment Project (SNAP), which sought to design a model system of assessment that was truly three-dimensional in nature. Over three years, the SNAP team developed and tested assessment items in 1,400 classrooms.

The SCALE team has applied everything they learned from SNAP, and their familiarity with state science assessment, to the Twig Science Benchmark Assessments.

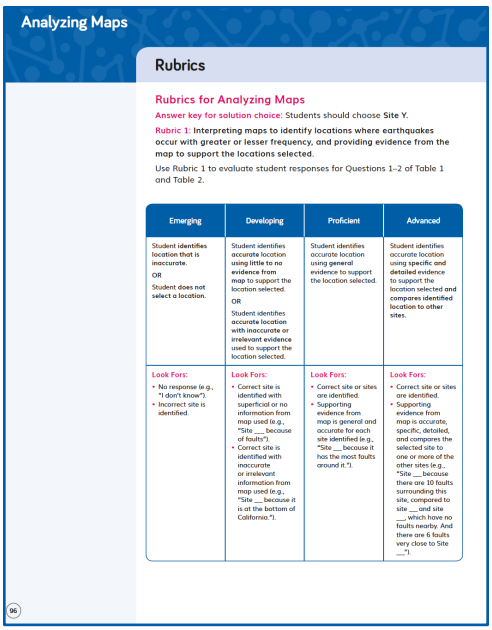
These summative assessments are designed to be taken individually at the end of a Driving Question. They provide a reliable and fair measure of students' progress with integrating the three dimensions of the PEs.

The assessments include open-ended problems that challenge students to apply what they have learned so far in the module to new contexts. This requires analysis and reasoning, rather than merely rote memorization. The Benchmark Assessments vary in length, from 30 minutes to 3 or 4 lessons, depending on the complexity of the task. Longer assessments include formatively assessed group work before students embark on an individual assignment.

Just as with the Performance Tasks, the leveled rubrics used to assess the Benchmark Assessments give a fair measure of student attainment against four levels: Emerging,

Developing, Proficient, and Advanced. The rubrics support educators to evaluate the level of conceptual understanding and include examples of student responses ("Look Fors"). Each Benchmark Assessment is correlated to a DoK level.

Example of a Benchmark Assessment Rubric



Analyzing Maps

Rubrics

Rubrics for Analyzing Maps

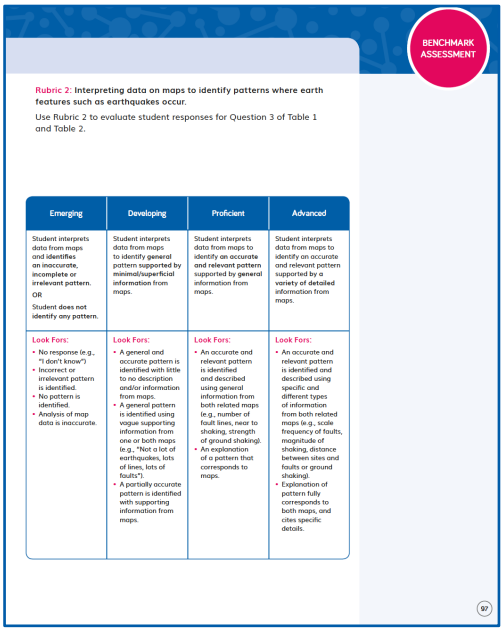
Answer key for solution choice: Students should choose Site Y.

Rubric 1: Interpreting maps to identify locations where earthquakes occur with greater or lesser frequency, and providing evidence from the map to support the locations selected.

Use Rubric 1 to evaluate student responses for Questions 1–2 of Table 1 and Table 2.

Emerging	Developing	Proficient	Advanced
Student identifies location that is inaccurate. OR Student does not select a location.	Student identifies accurate location using little to no evidence from map to support the location selected. OR Student identifies accurate location with inaccurate or irrelevant evidence used to support the location selected.	Student identifies accurate location using general evidence to support the location selected.	Student identifies accurate location using specific and detailed evidence to support the location selected and compares identified location to other sites.
Look Fors: • No response (e.g., "I don't know"). • Incorrect site is identified.	Look Fors: • Correct site is identified with superficial or no information from map used (e.g., "Site ___ because of faults"). • Correct site is identified with inaccurate or irrelevant information from map used (e.g., "Site ___ because it is at the bottom of California.").	Look Fors: • Correct site or sites are identified. • Supporting evidence from map is general and accurate for each site identified (e.g., "Site ___ because it has the most faults around it.").	Look Fors: • Correct site or sites are identified. • Supporting evidence from map is accurate, specific, detailed, and compares the selected site to one or more of the other sites (e.g., "Site ___ because there are 10 faults surrounding this site, compared to site ___ and site ___ which have no faults nearby. And there are 6 faults very close to Site ___").

96



BENCHMARK ASSESSMENT

Rubric 2: Interpreting data on maps to identify patterns where earth features such as earthquakes occur.

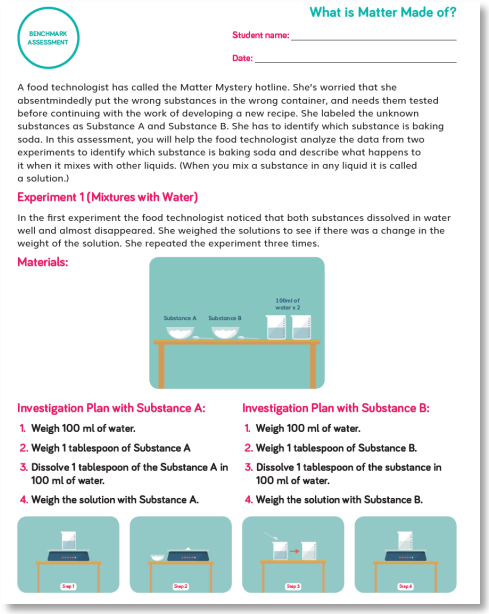
Use Rubric 2 to evaluate student responses for Question 3 of Table 1 and Table 2.

Emerging	Developing	Proficient	Advanced
Student interprets data from maps and identifies an inaccurate, incomplete or irrelevant pattern. OR Student does not identify any pattern.	Student interprets data from maps to identify general pattern supported by minimal/superficial information from maps.	Student interprets data from maps to identify an accurate and relevant pattern supported by general information from maps.	Student interprets data from maps to identify an accurate and relevant pattern supported by a variety of detailed information from maps.
Look Fors: • No response (e.g., "I don't know") • Incorrect or irrelevant pattern is identified. • No pattern is identified. • Analysis of map data is inaccurate.	Look Fors: • A general and accurate pattern is identified with little to no description and/or information from maps. • A general pattern is identified using vague supporting information from one or both maps (e.g., "Not a lot of earthquakes, lots of lines, lots of faults"). • A partially accurate pattern is identified with supporting information from maps.	Look Fors: • An accurate and relevant pattern is identified and described using general information from both related maps (e.g., number of fault lines, near to shaking, strength of ground shaking). • An explanation of a pattern that corresponds to maps.	Look Fors: • An accurate and relevant pattern is identified and described using specific and different types of information from both related maps (e.g., scale frequency of faults, magnitude of shaking, distance between sites and faults or ground shaking) • Explanation of pattern fully corresponds to both maps, and cites specific details.

97

Examples of Benchmark Assessments

Grade 5



BENCHMARK ASSESSMENT

What is Matter Made of?

Student name: _____

Date: _____

A food technologist has called the Matter Mystery hotline. She's worried that she absentmindedly put the wrong substances in the wrong container, and needs them tested before continuing with the work of developing a new recipe. She labeled the unknown substances as Substance A and Substance B. She has to identify which substance is baking soda. In this assessment, you will help the food technologist analyze the data from two experiments to identify which substance is baking soda and describe what happens to it when it mixes with other liquids. (When you mix a substance in any liquid it is called a solution.)

Experiment 1 (Mixtures with Water)

In the first experiment the food technologist noticed that both substances dissolved in water well and almost disappeared. She weighed the solutions to see if there was a change in the weight of the solution. She repeated the experiment three times.

Materials:

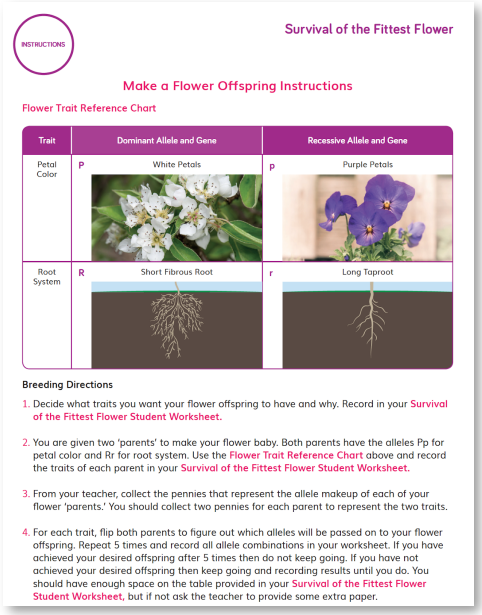
Investigation Plan with Substance A:

1. Weigh 100 ml of water.
2. Weigh 1 tablespoon of Substance A
3. Dissolve 1 tablespoon of the Substance A in 100 ml of water.
4. Weigh the solution with Substance A.

Investigation Plan with Substance B:

1. Weigh 100 ml of water.
2. Weigh 1 tablespoon of Substance B.
3. Dissolve 1 tablespoon of the substance in 100 ml of water.
4. Weigh the solution with Substance B.

Grade 6



INSTRUCTIONS

Survival of the Fittest Flower

Make a Flower Offspring Instructions

Flower Trait Reference Chart

Trait	Dominant Allele and Gene	Recessive Allele and Gene
Petal Color	P White Petals	p Purple Petals
Root System	R Short Fibrous Root	r Long Taproot

Breeding Directions

1. Decide what traits you want your flower offspring to have and why. Record in your Survival of the Fittest Flower Student Worksheet.
2. You are given two 'parents' to make your flower baby. Both parents have the alleles Pp for petal color and Rr for root system. Use the Flower Trait Reference Chart above and record the traits of each parent in your Survival of the Fittest Flower Student Worksheet.
3. From your teacher, collect the pennies that represent the allele makeup of each of your flower 'parents.' You should collect two pennies for each parent to represent the two traits.
4. For each trait, flip both parents to figure out which alleles will be passed on to your flower offspring. Repeat 5 times and record all allele combinations in your worksheet. If you have achieved your desired offspring after 5 times then do not keep going. If you have not achieved your desired offspring then keep going and recording results until you do. You should have enough space on the table provided in your Survival of the Fittest Flower Student Worksheet, but if not ask the teacher to provide some extra paper.

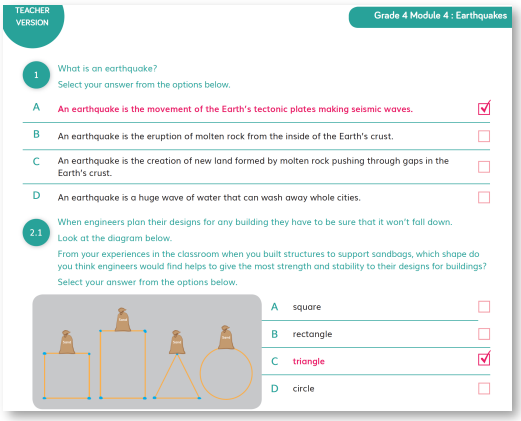
Multiple Choice Assessments (All Grades)

Twig Science Multiple Choice Assessments are a summative tool intended to be assigned at the end of a module. They can be used to assess student understanding of the concepts covered in the module, and identify areas of weakness that may need more attention.

Each question is correlated to an NGSS/3-D science standard and a Depth of Knowledge level. Assessments are multi-dimensional in nature, and reflect multiple choice items in state science tests.

The suggested pacing for Sections A and B is around 20 minutes. For some modules, there is a third section (Section C), which comprises an extended 10-minute question that focuses on a single PE. Sometimes Section C requires students to interact with an interactive or a data set they have not seen before, and apply their learning. Educators are advised to assign Section C to their advanced learners.

Examples of Multiple Choice Assessments (Teacher versions)


TEACHER VERSION

Grade 4 Module 4 : Earthquakes

1 What is an earthquake?

Select your answer from the options below.

A An earthquake is the movement of the Earth's tectonic plates making seismic waves. ☑

B An earthquake is the eruption of molten rock from the inside of the Earth's crust. ☐

C An earthquake is the creation of new land formed by molten rock pushing through gaps in the Earth's crust. ☐

D An earthquake is a huge wave of water that can wash away whole cities. ☐

2.1 When engineers plan their designs for any building they have to be sure that it won't fall down.

Look at the diagram below.

From your experiences in the classroom when you built structures to support sandbags, which shape do you think engineers would find helps to give the most strength and stability to their designs for buildings?

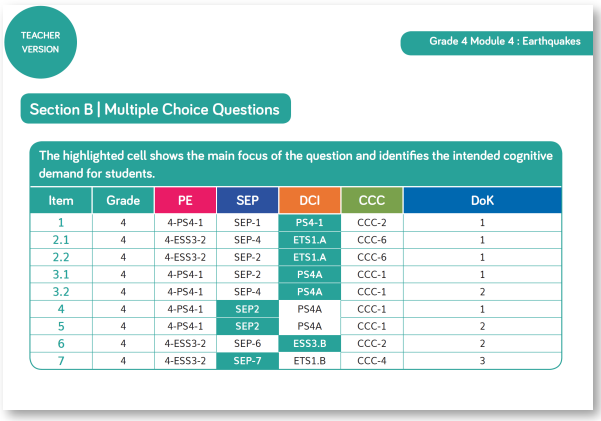
Select your answer from the options below.

A square ☐

B rectangle ☐

C triangle ☑

D circle ☐

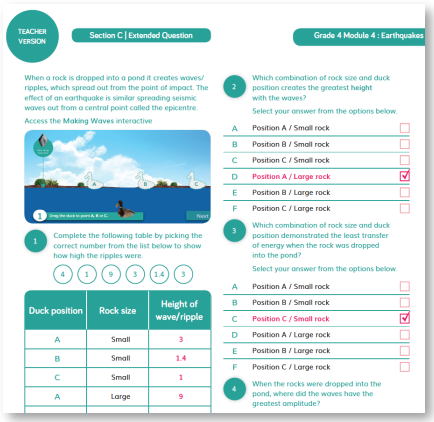

TEACHER VERSION

Grade 4 Module 4 : Earthquakes

Section B | Multiple Choice Questions

The highlighted cell shows the main focus of the question and identifies the intended cognitive demand for students.

Item	Grade	PE	SEP	DCI	CCC	DoK
1	4	4-PS4-1	SEP-1	PS4-1	CCC-2	1
2.1	4	4-ESS3-2	SEP-4	ETS1.A	CCC-6	1
2.2	4	4-ESS3-2	SEP-2	ETS1.A	CCC-6	1
3.1	4	4-PS4-1	SEP-2	PS4A	CCC-1	1
3.2	4	4-PS4-1	SEP-4	PS4A	CCC-1	2
4	4	4-PS4-1	SEP2	PS4A	CCC-1	1
5	4	4-PS4-1	SEP2	PS4A	CCC-1	2
6	4	4-ESS3-2	SEP-6	ESS3.B	CCC-2	2
7	4	4-ESS3-2	SEP-7	ETS1.B	CCC-4	3


TEACHER VERSION

Section C | Extended Question

Grade 4 Module 4 : Earthquakes

When a rock is dropped into a pond it creates waves/ripples, which spread out from the point of impact. The effect of an earthquake is similar spreading seismic waves out from a central point called the epicentre.

Access the Making Waves interactive

1 Complete the following table by picking the correct number from the list below to show how high the ripples were.

4 1 9 3 1.4 3

Duck position	Rock size	Height of wave/ripple
A	Small	3
B	Small	1.4
C	Small	1
A	Large	9

2 Which combination of rock size and duck position creates the greatest height with the waves?

Select your answer from the options below.

A Position A / Small rock ☐

B Position B / Small rock ☐

C Position C / Small rock ☐

D Position A / Large rock ☑

E Position B / Large rock ☐

F Position C / Large rock ☐

3 Which combination of rock size and duck position demonstrated the least transfer of energy when the rock was dropped into the pond?

Select your answer from the options below.

A Position A / Small rock ☐

B Position B / Small rock ☐

C Position C / Small rock ☑

D Position A / Large rock ☐

E Position B / Large rock ☐

F Position C / Large rock ☐

4 When the rocks were dropped into the pond, where did the waves have the greatest amplitude?