

## ■ What Next?

When the surveys come back, you code them, enter the information in a spreadsheet, and analyze the data to understand what people told you. At the very least, you should look at frequencies (How many people said yes; how many said no?) and means (What is the average response on a scale?). You can also correlate one variable to another (As number of visits increase does appreciation for the site also increase?) and look for differences between subsets (Do 4th grade teachers have a more successful field trip experience than 2nd grade teachers?). Your analysis should strive to reduce the data to some meaningful summary statements of what you found and then relate this back to your evaluation questions. See chapter 5 for more on this process.

## ■ Quality Assurance

The information you collect from your survey will be useful in your evaluation if you have

- designed and pilot tested a good survey;
- achieved a high response rate from a randomly selected sample; and
- compiled and reduced the information into meaningful recommendations.

## Tests of Knowledge

**M**any educators want to know whether or not audience members learn from their programs. Of course, you can always ask people if and what they learned, but such self-assessment is not as convincing as an objective measure of audience knowledge before and after a program. A test of knowledge measures knowledge and understanding that could be the result of participation in your program. You are more likely to use a test of knowledge in formative and summative than in front-end evaluation. Prior to development of your program, you may not know enough about what it will cover to design

a good test. In this case, it would be more helpful to learn what the audience knows about your chosen topic through interviews and surveys.

The challenge in using tests of knowledge is designing questions that are fair, reasonable, and appropriate to the program. If there is information unique to your program that participants could not have learned anywhere else, use it. For example, Monterey Bay Aquarium used these questions to assess visitors' knowledge after they watched a program in the auditorium about the aquarium's research activities:

- Who was doing the research?
- Where did the research take place?
- What was the program about?

As it is unlikely that visitors could have known the answers from any other source, these questions would help evaluators know whether people were paying attention to the program and whether they remembered its main ideas.

If your program information could be obtained from many sources—watching TV, reading the newspaper, attending school—then you need to ask learners a few key questions before and after their program participation in order to determine whether it was your program, and not something else, that increased their knowledge. In fact, many tests of knowledge are conducted using this pretest-posttest design to enable evaluators to say that the information gained is due to a particular program. The question in figure 3.3 was used in pretest-posttest fashion to reveal what respondents believe about woody biomass and fossil fuel energy sources before and after a program. A portion of a test of knowledge is provided in figure 3.4.

One challenge with using pretests and posttests to measure knowledge is that program participants will be alerted to important concepts from the pretest and will be more likely to pay attention to that information during the program. For this reason, some educators use assessment tools like this as a regular part of their programs—they help learners learn!

### FIGURE 3.3 Sample Knowledge Question

*Please check how burning wood to generate electricity compares to burning coal for each characteristic.*

	WOOD IS BETTER	THE SAME	WOOD IS WORSE	DON'T KNOW
Air quality	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Global warming	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Local jobs	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Cost	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

## FIGURE 3.4 Sample Test of Knowledge

Circle the letter of the best answer for each of the following questions. Please circle only one answer for each question.

Example: The color of grass is (circle the best answer):

- a. Blue
- b. Green
- c. Red
- d. Orange
- e. Don't know

1. The color of the lake water is determined by (circle the best answer):
  - a. The type of pollution in the water
  - b. The type of algae in the water
  - c. The time of day when viewing the water
  - d. All of the above
  - e. Don't know
2. Wetlands are important because they (circle the best answer):
  - a. Help control climate changes
  - b. Prevent the spread of exotic plant and animal species
  - c. Provide good development sites for businesses and homes
  - d. Filter and store water before it enters lakes, rivers, and streams
  - e. Don't know
3. What % of all the water in the world is suitable for human use (circle the best answer):
  - a. 1%
  - b. 50%
  - c. 10%
  - d. 25%
  - e. Don't know
4. A watershed is (circle the best answer):
  - a. A shelter for treated water
  - b. A large, man-made lake
  - c. A region draining into a body of water such as a river, stream, or lake
  - d. Not located near the Great Lakes
  - e. Don't know
5. Phytoplankton obtain food by (circle the best answer):
  - a. Absorbing sunlight near the surface of the water and producing their own food
  - b. Filter-feeding near the bottom of the water
  - c. Consuming other plankton
  - d. Eating water fleas
  - e. Don't know
6. Some fish, such as Lake Trout and Sturgeon, have a lighter colored belly so they are less visible to predators below them. This is (circle the best answer):
  - a. A mutation
  - b. A handicap
  - c. Unwanted
  - d. An adaptation
  - e. Don't know
7. An exotic species, such as sea lamprey, may be hazardous to an environment because (circle the best answer):
  - a. They are unattractive
  - b. They out-compete native species for resources
  - c. They increase pollution in the ecosystem
  - d. They eat the native species
  - e. Don't know
8. The abiotic factors of an environment are (circle the best answer):
  - a. The nonliving parts of an ecosystem, such as water and minerals
  - b. The living parts of an ecosystem, such as plants and animals
  - c. The interaction of plants and animals
  - d. Both the living and nonliving parts of an ecosystem
  - e. Don't know
9. Though Lake Superior's environment is in good condition, further research is necessary to address which one of the following issues (circle the best answer):
  - a. Reducing toxins in food fish
  - b. Protecting and restoring spawning streams
  - c. Better management of ship ballast water
  - d. All of the above
  - e. Don't know
10. One reason why Lake Victoria, one of the world's largest freshwater lakes, is in danger is because of (circle the best answer):
  - a. Drought
  - b. Fertilizers and pesticides draining into it from nearby farms
  - c. Decreased numbers of Nile perch
  - d. Increased algal blooms depleting the lake of nutrients
  - e. Don't know
11. The primary contaminants of water supplies are (circle the best answer):
  - a. Toxic chemicals and waterborne diseases from runoff
  - b. Exotic aquatic plant species
  - c. Solid waste
  - d. Debris (sediments, etc.)
  - e. Don't know

Source: Kubarek-Sandor (2006).

Answers: 1. d; 2. d; 3. c; 4. d; 5. a; 6. d; 7. b; 8. a; 9. d; 10. b; 11. a

### BOX 3.5

## Developing a Test of Knowledge

The Lagoon Quest Program, a successful educational field trip offered by the Brevard Zoo in Melbourne, Florida, was recently adopted by the county school district as a required portion of the 4th grade curriculum. The school district and the zoo want to know if students are learning about their local Indian River Lagoon as a result of the two-week unit. A test of knowledge was developed by zoo staff and teachers to cover the basic elements of the program: lagoons and estuaries, water quality, and threats to water quality in the watershed. An additional survey was developed for teachers to discover how many program-related activities they implemented in the classroom. By implementing the quiz both before the unit begins and after the lessons are completed and comparing this information to the teacher survey, evaluators can measure student knowledge gain and attribute it to the program (see appendix B).



### 3.11 APPLICATION EXERCISE

Return to your logic model and evaluation plan. Do you have an evaluation question that can be answered through a test of knowledge? If not, imagine that you want to measure an increase in knowledge. Look at the knowledge objectives for your program and the types of indicators that will help you know if learners have gained this understanding. Think about what learners are already likely to know and what questions you could ask to demonstrate that your program is responsible for increasing their knowledge. Write down some of the broad categories of information your program is supposed to teach, and then observe the program presentation to make sure that information is covered adequately. Ask some participants about your categories at the conclusion of the program to make sure you are on the right track.

Another challenge is balancing test length with the time people are able or willing to give. Multiple items that measure the same concept are needed to improve test reliability (one question just won't do it). Yet young children have limitations in reading and test taking, and adults may not want to voluntarily submit to a test. Aim for 15 to 20 questions that cover three to five concepts. You may have to select the most important concepts and discard the rest.

It is tempting to use very specific details, unique to your program, in your test of knowledge so that you can just measure this post-program knowledge (as in the Monterey Bay Aquarium example). If you decide to go this way, make sure the information you test is meaningful as well as memorable (see box 3.5). Focusing on obscure tidbits just because they are unique likely will not produce the positive change in learners that will impress funders. If children can recall the scientific name of the pine cone willow gall, does that mean they are more environmentally literate? If, on the other hand, students can provide correct responses that reflect the state standards because of your program—that could make stakeholders very happy!

### ■ Types of Items on Tests of Knowledge

The types of questions most frequently used to measure knowledge are multiple choice, true/false, and matching. Vocabulary and concept identification questions can also be used to test lower order thinking skills. These questions will have one right answer and be easy to score. In some cases, basic questions may be appropriate, while in others they may be too trivial. Particularly with older learners, aim for questions that provide scenarios and choices as well as opportunities to critique information or provide evidence—that is, include items that will allow an assessment of students' higher order thinking skills and understanding. (Multiple choice questions written in such a way to require students to analyze scenarios test higher order thinking skills and are easier to score than essay questions.) The following information will help you write good test questions.

*Multiple choice* questions are widely used in tests of knowledge. They enable you to include common myths among the wrong responses, which can be helpful in determining if your program is combating these misconceptions.

- Most of the item information should be placed in the stem, leaving the answer choices relatively concise, simple, and clear.
- Response options to multiple choice questions should be as grammatically parallel as possible.
- Four choices per question are typical.
- The incorrect choices are challenging to write. Some test writers aim to have one wildly wrong

answer, two feasible wrong answers (based on common misunderstandings or mistakes), and one unambiguously right answer. The more obviously wrong an answer, however, the easier it is for someone with little knowledge to guess the right answer.

- While negative questions (“Which of the following is not true . . .”) are easy to write (you only have to provide one wrong answer statement, not three), they can either be too simple or they can confuse readers if a negative appears in one of the responses as well. It is best to avoid them, if possible.
- Try not to provide clues about the correct response through grammar (for example, ending the stem with “an” and beginning only one answer choice with a vowel or providing only once choice whose verb agrees with the subject in the stem).
- Vary the correct responses among a, b, c, and d.
- Try to avoid using as responses “all of the above” or “none of the above,” as they are usually the correct responses. As long as the learner knows two responses are likely, the “all of the above” response must be correct.
- Avoid supplying the answer to one question in the stem of another. Make all questions independent.

*Matching* questions are often used to test vocabulary and simple concepts. You can match concepts or terms to definitions, or you can match pictures to terms. It is often difficult to design them to be appropriately challenging.

Here are some tips for writing matching questions.

- Limit the number of items to be matched to less than ten. It is very difficult to keep track of more than ten items.
- Keep the terms and definitions relatively simple. You are not testing reading ability but knowledge.
- Options for answers should be available for more than one match or should exceed the number of available matches. Both strategies make the question more challenging, as process of elimination does not reveal the last answer.

Figure 3.5 shows a good example of a matching question.

**FIGURE 3.5**

### Sample Matching Question

*Match the animal on the left with the category on the right by writing the letter (a, b, c, d, or e) in the space next to correct animal.*

<b>ANIMAL</b>	<b>CATEGORY</b>
___ bat ray	a. mammal
___ sea otter	b. bird
___ brittle star	c. amphibian
___ brown pelican	d. fish
	e. invertebrate



### 3.7 CHECK FOR UNDERSTANDING

What could be changed to improve these questions from a test of knowledge for 4th graders learning about the Indian River Lagoon?

1. The Indian River Lagoon is a (an)
 

a. barrier island	b. lagoon	c. estuary	d. lagoon and estuary
-------------------	-----------	------------	-----------------------
2. Freshwater enters the Indian River Lagoon by way of
 

a. rivers and canals	b. precipitation	c. condensation	d. rivers, canals, and precipitation
----------------------	------------------	-----------------	--------------------------------------
3. Which of the following should NOT be done in the Indian River Lagoon?
 

a. boating	b. fishing	c. bird watching	d. dumping trash
------------	------------	------------------	------------------
4. An increase in turbidity is also described as a decrease in
 

a. salinity	b. pH	c. temperature	d. water clarity
-------------	-------	----------------	------------------
5. The Indian River Lagoon has more species of plants and animals than any other \_\_\_\_\_ in the United States.
 

a. estuary	b. inlet	c. island	d. reef
------------	----------	-----------	---------
6. A lagoon is a shallow body of water protected from the ocean by a (an)
 

a. adaptation	b. inlet	c. decomposer	d. barrier island
---------------	----------	---------------	-------------------

*Answers are found in appendix A.*

True/False questions can also be used to test understanding of definitions and basic concepts, but they can also be written to apply concepts to real-world situations.

- Questions should be unambiguously true or false.
- Avoid using the words “always” and “never” to make an item false. This is usually a strong clue that an item is false, as the world is usually a bit fuzzier than that.

### ■ Pilot Testing and Next Steps

It is essential to pilot your test of knowledge with audiences who have not yet been exposed to your program as well as those who have. Ideally, there will be a difference in their scores. As with a survey, sit down with your pilot-testers and ask them to “think out loud” as they complete the test. Ask them why they think certain

responses are correct. Plan to ask experts to review your test—someone who is an expert in the subject matter, someone who is an expert on the desired audience (a 5th grade teacher, for example, if the test will be given to 5th graders), and someone who is an expert in assessment or test development.

When the completed tests are returned, you will enter the responses into a spreadsheet and analyze the data as described in the quantitative analysis section of chapter 5. If you are conducting both a pretest and posttest, you may be looking for a significant difference between the two. If you are using your test only as a posttest, you might be looking for high scores on the questions about the content covered in your program and for lower scores on questions about material not covered in your program. Or you might be looking for criteria specified in your lesson objectives or evaluation questions. For example, did 80% of program participants correctly answer seven or more of the 10 questions?

You may want to use Cronbach’s alpha—a statistical test that measures reliability—with your pilot test results. It is often used to help eliminate poorly performing test questions and improve tests of knowledge; it can be found in statistical software packages.

### ■ Quality Assurance

Your test of knowledge will be useful in your evaluation if you have

- carefully identified reasonable and appropriate concepts to test;
- designed and pilot tested the test; and
- checked each question for unnecessary clues, obvious answers, and illogical responses.



### 3.8 CHECK FOR UNDERSTANDING

To help determine whether or not a teacher workshop to distribute a new environmental education curriculum was valuable, the following questions were used as a pretest before the workshop and as a posttest after the workshop. Because the questions could be answered correctly without any knowledge of the curriculum, the results didn’t show any improvement, and it was unclear if the teacher workshop had been effective. If you were evaluating the teacher workshop, how could you improve these individual items or the pretest/posttest in general in order to produce a better sense of the effectiveness of the workshop?

- |            |   |
|------------|---|
| True/False | This curriculum can be conducted successfully at various sites within your travel limitations.  |
| True/False | One reason to use the curriculum is to give children hands-on experiences with the environment.   |
| True/False | Environmental education does not need to be a priority in our society today.  |
| True/False | The main objective of this curriculum is to develop individuals whose attitudes and skills foster a reverence and stewardship of the earth. |

*Answers are found in appendix A.*



### 3.12 APPLICATION EXERCISE

Revisit your test of knowledge questions. Make sure they are all well worded, logical, and unambiguous. Ask a colleague if they make sense. Do you have some easy and some difficult questions? Are the questions sequenced appropriately? You may find the Test of Knowledge Checklist in worksheet 3.3 (at the end of this chapter) helpful as you review your test.