

# Building Concepts: Two-Way Tables and Categorical Data

TEACHER NOTES

## Lesson Overview

In this TI-Nspire lesson, students will use two-way tables to investigate associations between two categorical variables. Students will first compare absolute frequency tables and graphs to relative frequency tables and graphs. Then using information gleaned from these sources students will make statements about general association between variables. Finally, students will compare conditional frequencies calculated from these two-way tables.



Two-way tables of *relative frequencies* provide numerical evidence for whether or not there is an association between categorical variables, and the associated relative frequency graphs provide visual evidence.

## Prerequisite Knowledge

*Two-Way Tables and Categorical Data* is part of a series of lessons that explores the concepts of statistics and probability. This lesson builds on the concepts of the previous lessons. Prior to working on this lesson students should have completed *Law of Large Numbers*, *Sample Proportions*, and *Scatter Plots*. Students should understand:

- how to interpret points, clusters, gaps, and outliers on a scatter plot;
- the concept of sampling variability;
- the concept of relative frequency;
- how to read a bar graph.

## Lesson Pacing

This lesson should take 50–90 minutes to complete with students, though you may choose to extend, as needed.

## Learning Goals

1. Recognize whether there is an association between categorical data using two-way tables and stacked bar graphs;
2. Find conditional probabilities given information in a two-way table.

## Vocabulary

- **two-way table**: a chart that organizes information about two categorical variables
- **categorical data**: information grouped together according to the same rule
- **relative frequency**: the number of times an event occurs during experimental trials, divided by the total number of trials conducted
- **conditional probabilities**: probability values that takes into account some additional known information (the “given” or “condition”).

# Building Concepts: Two-Way Tables and Categorical Data

TEACHER NOTES

## Lesson Materials

- Compatible TI Technologies:



TI-Nspire CX Handhelds,



TI-Nspire Apps for iPad®,



TI-Nspire Software

- Two-Way Tables and Categorical Data\_Student.pdf
- Two-Way Tables and Categorical Data \_Student.doc
- Two-Way Tables and Categorical Data.tns
- Two-Way Tables and Categorical Data\_Teacher Notes
- To download the TI-Nspire activity (TNS file) and Student Activity sheet, go to <http://education.ti.com/go/buildingconcepts>.

## Class Instruction Key

The following question types are included throughout the lesson to assist you in guiding students in their exploration of the concept:



**Class Discussion:** Use these questions to help students communicate their understanding of the lesson. Encourage students to refer to the TNS activity as they explain their reasoning. Have students listen to your instructions. Look for student answers to reflect an understanding of the concept. Listen for opportunities to address understanding or misconceptions in student answers.



**Student Activity:** Have students break into small groups and work together to find answers to the student activity questions. Observe students as they work and guide them in addressing the learning goals of each lesson. Have students record their answers on their student activity sheet. Once students have finished, have groups discuss and/or present their findings. The student activity sheet can also be completed as a larger group activity, depending on the technology available in the classroom.



**Deeper Dive:** These questions are provided for additional student practice and to facilitate a deeper understanding and exploration of the content. Encourage students to explain what they are doing and to share their reasoning.

# Building Concepts: Two-Way Tables and Categorical Data

TEACHER NOTES

## Mathematical Background

Scatterplots are useful tools for investigating associations between paired *measurement* variables. In this lesson, students take a careful look at relationships between paired *categorical* variables. Categorical variables represent data that can be divided into distinct groups. For example, eye color, gender, or types of music. Two-way tables provide a tool for investigating whether or not there is an association between being a member of a category of one type with being a member of a category of another type. For example, students may be interested in relationships between gender and sports participation or age and parental curfews.

Two-way tables report a concise summary of either the membership counts (frequencies) or membership proportions (relative frequencies) in each category. The simplest kinds of two-way table have two row categories and two column categories. Each row corresponds to the members of categories of one type (example: row 1 represents Boys and row 2 represents Girls) and each column corresponds to members of the other categories of interest (example: Curfew and No Curfew). The tables and associated stacked bar graphs can be used to address questions of association between paired categorical data.

Relative frequencies have been addressed in an earlier lesson (Lesson 16, *Law of Large Numbers*), and they are especially important in the use of two-way tables.

Students have already been introduced to probability ideas in previous lessons. Two-way tables can serve to further introduce students to the idea and language of a *conditional probability*. A conditional probability value takes into account some additional known information (the “given” or “condition”), such as “what is the probability that a student has a curfew given that the student is a boy?”



# Building Concepts: Two-Way Tables and Categorical Data

TEACHER NOTES

## Part 1, Page 1.3

Focus: Relative frequency tables and graphs are important in identifying associations between categories. Relative frequencies make it easier to judge whether an association exists or not, especially when the counts (absolute frequencies) between two categories are very different.

On Page 1.3, students work with an interactive version of a two-way table with exactly two rows for the first category and two columns for the second category. Students can edit the row labels and column labels as well as counts in the cells. The totals for each row update automatically.

**Table** allows a choice of counts or relative frequencies (decimal or percent) to be displayed in the table.

**Graph** show/hides stacked bar graphs of the counts (absolute frequency) and relative frequency (decimal or percent) of the data in the table.

**Clear** clears the data from the table including the row and column headings.

**Reset** hides the graphs and resets the data to the original data set.

	Indiv	Team	Total
Boys	15	35	50
Girls	20	50	70

### TI-Nspire Technology Tips

**menu** accesses page options.

**tab** toggles through the editable cells in the table and the dec/% for the relative frequency graph

**enter** turns on/off the stacked bar graphs of counts and relative frequency.

**ctrl del** resets the page to the original screen.



### Class Discussion

The following questions focus on the importance of using both graphs and relative frequencies to identify evidence of association between categories. The examples are chosen so that the existence of any association is fairly obvious. There is no attempt at this level to quantify any measures of association.

Have students...

- *Who do you think likes team sports more than individual sports, boys or girls?*

Look for/Listen for...

Answers may vary. Boys may like team sports more than girls.



# Building Concepts: Two-Way Tables and Categorical Data

TEACHER NOTES



## Class Discussion (continued)

*The editors of the school newspaper were curious about this very question and other questions about the differences between boys' and girls' preferences. They gave a survey to 120 randomly chosen students (70 girls and 50 boys). The first question on the survey was "which kind of sport do you prefer watching: individual sports or team sports?" The responses to this question were as follows: Of the 50 boys, 15 said "individual sports" and 35 said "team sports." Of the 70 girls, 20 said "individual sports" and 50 said "team sports."*

*The survey data have been entered into page 1.3 of the TNS activity.*

- What do the numbers in the Totals Column represent?**

Answer: These are the totals of the numbers in the cells for each row, so the first row total is the total number of boys surveyed and the second row total is the total number of girls surveyed.
- Samee says that more girls than boys prefer watching individual sports. From the data, do you think she is correct?**

Answers will vary. The raw counts support the statement, but there were more girls surveyed than boys overall.
- Select Show Rel. Freq. What does this graph show?**

Answer: The left bar in the graph shows the percent of the boys who answered individual (green) and team sports (purple). The right bar shows the percent of girls who answered individual (green) and team sports (purple).
- Does this graph support Samee's statement?**

Answer: The percentages shown are almost identical, so this does not support Samee's statement if Samee meant that girls are more likely to prefer watching individual sports.

*Remember the idea of relative frequency from Activity 16 Law of Large Numbers. Select menu> Table> Rel. Freq. (decimal) to show the relative frequencies of the responses in the table.*

- What does the value 0.30 in the cell for Boys Indiv mean and how was it calculated?**

Answer: 0.3 represents the proportion of boys who liked individual sports. There were 15 out of 50 boys who responded "individual sports," so the relative frequency of that response is  $\frac{15}{50} = 0.30$ .
- How does 0.30 in this cell relate to the Rel. Freq. graph shown?**

Answer: It corresponds to the green part of the Boys bar showing 30% responded "individual sports."
- What do the values 1.00 in the totals column of the table represent?**

Answer: The sum of the relative frequencies adds up to 1.

# Building Concepts: Two-Way Tables and Categorical Data

TEACHER NOTES

## Class Discussion (continued)

- ***What do the values 1.00 in the totals column of the table represent?***
- ***Sometimes relative frequencies are represented as decimals and sometimes as percents. Select menu> Table> Rel. Freq. (percent). What do these values represent?***

***Select Show Freq. button to see a bar graph with Boys and Girls categories on the horizontal axis.***

- ***What do the stacked colored bars represent?***
- ***Describe how the Relative Frequency graph differs from the Counts graph?***
- ***Why are the combined stacked bars in the relative frequency graph exactly the same height, even though more girls than boys were surveyed?***

Answer: The sum of the relative frequencies adds up to 1.

Answer: These are the relative frequencies expressed as percentages. Now the total of each row is 100%.

Answer: The heights of the green bars show the counts (frequencies) of the individual sports responses for boys and girls, and the heights of the purple bars show the counts (frequencies) of the team sports responses for boys and girls. The total heights of the stacked bars show the total counts of boys and girls surveyed.

Answer: The Counts graph shows the actual number of responses (called the absolute frequencies) from the survey. In the Relative Frequency graph, the heights of the green bars show the relative frequencies of the individual sports responses for boys and girls, and the heights of the purple bars show the frequencies of the team sports responses for boys and girls.

Answer: The first combined bar shows the total of the relative frequencies of all the boys surveyed, and the second combined bar shows the total of the relative frequencies of all the girls surveyed. That will be 100% for each group, no matter what the actual numbers of girls and boys surveyed.



# Building Concepts: Two-Way Tables and Categorical Data

TEACHER NOTES



## Student Activity Questions—Activity 1

1. A possible association between two categorical variables can be revealed by looking at relative frequencies. For example, if girls have a much greater relative frequency of liking team sports than boys, then this suggests that gender and sports preferences are associated: girls are more likely to like team sports than boys.

a. Why might counts (absolute frequencies) not be as helpful in seeing an association between categorical variables?

Answer: If the count totals of the two row categories are very different, then the values in the cells could be very different even if there is no association.

b. If the relative frequencies are not very different, then what does that indicate about the association between gender and preference for team sports?

Answer: If the relative frequencies are not very different, then there is little evidence of association.

c. Reset page 1.3. Given the data from the school newspaper survey, do you think there is an association between gender and sports preference based on this school's survey data?

Answer: Based on the relative frequencies, there does not appear to be much of an association.

2. Select *Reset* and *Show Freq.*

a. Another school gave the same survey question to some of their students. The responses from the 50 boys surveyed were exactly the same (15 preferred watching individual sports and 35 preferred watching team sports). However, the responses from the girls were 20 preferring individual sports and 20 preferring team sports. Make the change in the table by editing the Girls Team cell from 50 to 20. What do you notice in the Counts graph?

Answer: More of the girls' bar is green than the boys' bar, but the counts are different since the total bars have different heights.

b. Now select *Show Rel. Freq.* What does this graph tell you?

Answer: The relative frequency graph gives more clear evidence of an association, with the girls preferring individual sports more than boys and the boys preferring team sports more than girls. More of the girls bar is green than the boys bar, but now the total bars are the same height, so I can compare more easily.

c. A third school also gave the same survey question to some of their students. In this school exactly the same number of boys and girls were surveyed, with 10 boys preferring individual sports and 50 boys preferring team sports, and 35 girls preferring individual sports and 25 girls preferring team sports.

Edit the table cells to show these counts. What do you notice about the two graphs now?

Answer: Now both the Counts graph and Relative Frequency graph have two stacked bars of exactly the same height, so it is easy to compare in either graph. Both graphs show some evidence of association between gender and sports preferences.



# Building Concepts: Two-Way Tables and Categorical Data

TEACHER NOTES

**Teacher Tip:** Emphasize that, when the total counts for two categories are the same, then the relative frequency bars will be scaled proportionately to the absolute frequency bars. The relative frequency graphs are particularly important when comparing categories where the total counts are different.

## Part 2, Page 1.3

Focus: A *conditional probability* is a probability value that takes into account some additional known information (the “given” or “condition”).

	Indiv	Team	Total
Boys	15	35	50
Girls	20	50	70



## Class Discussion

The following questions introduce the idea of conditional probability.

Have students...

Reset page 1.3 to show the original school’s survey results.

- ***What is the probability that a randomly chosen surveyed student answered “team sports” given that the student is a boy?***
  
- ***Given that a randomly chosen surveyed student answered “individual sports”, what is the probability that student was a girl?***

Look for/Listen for...

Answer: Of the 50 boys who answered the survey, 35 of them preferred team sports, so the probability that a randomly chosen boy answered “team sports” is  $\frac{35}{50} = 0.70$ , or 70%. Note: In this case, the relative frequency table provides the conditional probability because the condition corresponded to a row category.

Answer: There were a total of 35 students who answered “individual sports” and 20 of these were girls. So the probability is given by  $\frac{20}{35} \approx 0.57$ , or 57%.

Note: Often column totals are also provided in displays of two-way tables. In this case, students will need to calculate the first column total themselves in order to find the conditional probability asked for.





# Building Concepts: Two-Way Tables and Categorical Data

TEACHER NOTES



## Class Discussion (continued)

*The original school's survey also asked the students whether or not their parents set a curfew for them (a curfew is a fixed time to be back home each evening). The responses for this question were as follows: 30 of the boys indicated they had a curfew and 20 indicated they did not, while 60 of the girls indicated they had a curfew and 10 indicated they did not.*

*Edit the heading for the first column to “curfew” and the second column heading to “no cur” and then fill in the table on page 1.3 with these response counts. Use your new table to answer the following questions.*

- *Is there an association between student gender and curfew? Explain your thinking.*
- *Amanda says that based on the survey, twice as many girls indicated they had a curfew as boys, so girls are twice as likely to have a curfew as boys. Do you agree with Amanda? Why or why not?*
- *What is the probability that a surveyed student will have a curfew given that the student is a boy? How did you determine your answer?*
- *What is the probability that a surveyed student will have a curfew given that the student is a girl?*
- *Based on your answers to the two previous questions, what would you tell Amanda about her claim?*
- *Given that a randomly chosen surveyed student answered “no curfew,” what is the probability that this student is a boy?*

Answer: There appears to be evidence for an association. Some students may use the table, and others may use the graphs (be sure they refer to the relative frequency graph).

Answer: It is true that twice as many girls answered curfew than boys, but there were more girls surveyed than boys, but to say “twice as likely” would mean that the relative frequency of girls with curfews is twice the relative frequency of boys with curfew.

Answer: The Relative Frequency table shows that the percentage of boys with a curfew is 60%.

Answer: The Relative Frequency table shows that the percentage of girls with a curfew is 85%.

Answer: Girls are more likely than boys to have a curfew, but the percentage (85%) of girls with curfews is not nearly twice that of boys (60%).

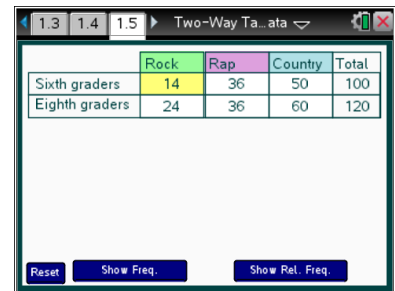
Answer: 20 boys and 10 girls answered “no curfew,” so the total number of “no curfew” responses is 30 and the conditional probability that a randomly chosen student from this group is a boy is  $\frac{20}{30} = \frac{2}{3} \approx 0.67$ , or 67%.

# Building Concepts: Two-Way Tables and Categorical Data

## Part 3, Page 1.5

Focus: Students further investigate conditional probability.

Page 1.5 provides an interactive two-way table with category labels appropriate for addressing a question posed in the following story setting: “Is there a difference between sixth graders and eighth graders with regard to their preference for rock, rap, or country music?” Otherwise, this interactive table and the graphs function in the same way as those on Page 1.3.



### Class Discussion

The following questions introduce the idea of conditional probability.

**A random sample of 100 6<sup>th</sup> graders and 120 8<sup>th</sup> graders were surveyed with regard to their preference for Rock, Rap, or Country music. The table on page 1.5 reports the results of the survey (shown below), with the frequencies (counts) of responses for each category of music preference for 6<sup>th</sup> graders and 8<sup>th</sup> graders.**

	Rock	Rap	Country	Total
6 <sup>th</sup> graders	14	36	50	100
8 <sup>th</sup> graders	24	36	60	120

- **Now select Show Freq. This shows a bar graph with the three colored bars stacked vertically. Explain how this bar graph reflects the frequencies in the table.**

Answer: The total height of the stacked bars represents the total counts (frequencies) of 6<sup>th</sup> graders and 8<sup>th</sup> graders surveyed. Within the stacked bars, the height of each colored bar represents the count or frequency of the students indicating a preference for a particular category of music, with green representing Rock, purple Rap, and blue Country.

# Building Concepts: Two-Way Tables and Categorical Data

TEACHER NOTES

## Class Discussion (continued)

Have students...

- ***Samuel wonders about the association between grade level and music preference. He said that the table showed that the 8<sup>th</sup> and 6<sup>th</sup> graders equally preferred Rap. Samuel also said that 8<sup>th</sup> graders liked Country and Rock more than 6<sup>th</sup> graders. Do you agree with Samuel's statements?***

Look for/Listen for...

Answer: The counts of students are equal for Rap for the two grades and the counts of 8<sup>th</sup> graders is larger than the counts of 6<sup>th</sup> graders for Country and Rock, but the total count of 8<sup>th</sup> graders (120) is larger than the total count of 6<sup>th</sup> graders (100), so it is not clear that this table supports Samuel's statements. Students should consider relative frequencies. From the relative frequency table or graph, there seems to be a slight association based on the survey: 8<sup>th</sup> graders appear slightly more likely than 6<sup>th</sup> graders to prefer rock music (0.2 compared to 0.14 relative frequency), and 6<sup>th</sup> graders appear slightly more likely than 8<sup>th</sup> graders to prefer rap music (0.36 compared 0.3 relative frequency).



## Student Activity Questions—Activity 2

1. Using the frequency table (counts) or the relative frequency table, answer these questions.

- a. Is a 6<sup>th</sup> grader or an 8<sup>th</sup> grader more likely to like rap?

Answer: If chosen randomly from the group of students surveyed, 6<sup>th</sup> graders are slightly more likely to like rap (0.36 compared to 0.30). (These values are close enough that it would be hard to say that a 6<sup>th</sup> grader or 8<sup>th</sup> grader chosen randomly from the entire school population is more likely to like rap.)

- b. If you randomly selected an 8<sup>th</sup> grader who had taken the survey, what is the probability that the 8th grader preferred rap music?

Answer: The probability is 0.30, represented by the relative frequency shown in the table (30% of the 8<sup>th</sup> graders surveyed preferred rap music).

- c. If a student who took the survey prefers rap music, what is the probability that this student is an eighth grader?

Answer: Of those students taking the survey, there were exactly the same number of 6<sup>th</sup> graders as 8<sup>th</sup> graders who preferred rap music (36). If one of these 72 students were chosen at random, the probability of that student being an 8<sup>th</sup> grader would be  $\frac{36}{72} = \frac{1}{2} = 0.5$

- d. Is there an association between grade level and music preference?

Answer: There is a slight preference of 6<sup>th</sup> graders for rap and 8<sup>th</sup> graders for rock.



# Building Concepts: Two-Way Tables and Categorical Data

TEACHER NOTES



## Student Activity Questions—Activity 2 (continued)

- e. What music do most of the 8<sup>th</sup> and 6<sup>th</sup> graders prefer?

Answer: Country music, with half (0.5) of both 6<sup>th</sup> graders and 8<sup>th</sup> graders survey indicating this preference.



## Deeper Dive — Page 1.3

- **Create a table on page 1.3 for Boys and Girls where the frequency of responses for Boys is twice that for Girls, but the relative frequency of Type 1 responses for Boys is half that for Girls.**

Answer: Tables will vary.



## Deeper Dive — Page 1.5

**Suppose you were given this relative frequency table for the music preference responses from another school.**

	Rock	Rap	Country	Total
6 <sup>th</sup> graders	0.20	0.40	0.40	100
8 <sup>th</sup> graders	0.40	0.20	0.40	120

- **A 6<sup>th</sup> grader responding to the survey is chosen at random. What is the probability that this student prefers Country music?**

Answer: 0.40 or 40%

- **A student is chosen at random from those who responded that they preferred Country music. What is the probability that this student is a 6<sup>th</sup> grader?**

Answer: It depends on the number of 6<sup>th</sup> graders and 8<sup>th</sup> graders. For example, suppose there were 50 6<sup>th</sup> graders and 100 8<sup>th</sup> graders surveyed. Then  $(0.40)(50) = 20$  6<sup>th</sup> graders responded Country and  $(0.40)(100) = 40$  8<sup>th</sup> graders responded Country. Of the 60 total students who responded

Country, the probability that a student is a 6<sup>th</sup> grader is  $\frac{20}{60} \approx 0.333$ . If there had been only 50 8<sup>th</sup>

graders surveyed so that the total counts were the same, then the probability that a student who

responded Country is a 6<sup>th</sup> graders would be  $\frac{20}{40} = 0.50$ .



# Building Concepts: Two-Way Tables and Categorical Data

TEACHER NOTES

## Sample Assessment Items

After completing the lesson, students should be able to answer the following types of questions. If students understand the concepts involved in the lesson, they should be able to answer the following questions without using the TNS activity.

- Mr. Wallace surveyed 75 students at Poole Middle School to find out the students' favorite place to eat lunch. The results are shown below.

**Favorite Place to Eat Lunch**

	Cafeteria	Outside	Total
Boys	16	21	37
Girls	24	14	38

Which table shows the approximate relative frequencies of Mr. Wallace's data?

a. **Favorite Place to Eat Lunch**

	Cafeteria	Outside	Total
Boys	16%	21%	37%
Girls	24%	14%	38%

b. **Favorite Place to Eat Lunch**

	Cafeteria	Outside	Total
Boys	21%	28%	49%
Girls	32%	19%	51%

c. **Favorite Place to Eat Lunch**

	Cafeteria	Outside	Total
Boys	40%	60%	49%
Girls	60%	40%	51%

d. **Favorite Place to Eat Lunch**

	Cafeteria	Outside	Total
Boys	43%	57%	100%
Girls	63%	37%	100%

**Answer: d)**

(Adapted from the ENGAGE NY draft released items for New York State Testing Program Grade 8 Common Core Mathematics Test).



# Building Concepts: Two-Way Tables and Categorical Data

2. This table appeared in a report from the Center for Disease Control, and it describes risk factors for obesity associated with age.

**Obesity**

	Not at Risk	At Risk	Row Total
Ages 18 to 24	57.3	42.7	100
Ages 25 to 44	38.6	61.4	100

- a. What do these numbers represent? Are they frequencies or relative frequencies or neither?

**Answer:** *These numbers represent relative frequencies, but they are being reported as percentages between 0 and 100 rather than as decimal numbers between 0 and 1.*

- b. Does there appear to be an association between age group and obesity risk?

**Answer:** *There does appear to be an association with the older group (Age 25 to 44) over 60% at risk compared to approximately 40% at risk for the younger group (Age 18 to 24).*

3. Two classes of students at a school are surveyed as to their favorite colors. The results of the survey are shown in the table below.

**Favorite Color**

	Red	Blue	Green	Yellow	Purple	Totals
Boys	2	10	1	6	1	20
Girls	6	2	4	4	x	25

- a. How many girls chose purple?

**Answer:** 9

- b. A boy is chosen at random. What is the probability that his favorite color is yellow.

**Answer:**  $\frac{6}{20} = 0.30$

- c. One of the students who likes red is chosen at random. What is the probability that this student is a girl?

**Answer:**  $\frac{6}{8} = 0.75$

- d. One of the students is chosen at random. What is probability that the student's favorite color is purple?

**Answer:**  $\frac{10}{45} \approx 0.222$

# Building Concepts: Two-Way Tables and Categorical Data

TEACHER NOTES

## Student Activity Solutions

In these activities you will find conditional probabilities given information in a two-way table. After completing the activities, discuss and/or present your findings to the rest of the class.



### Activity 1 [Page 1.3]

1. A possible association between two categorical variables can be revealed by looking at relative frequencies. For example, if girls have a much greater relative frequency of liking team sports than boys, then this suggests that gender and sports preferences are associated: girls are more likely to like team sports than boys.
  - a. Why might counts (absolute frequencies) not be as helpful in seeing an association between categorical variables?  
*Answer: If the count totals of the two row categories are very different, then the values in the cells could be very different even if there is no association.*
  - b. If the relative frequencies are not very different, then what does that indicate about the association between gender and preference for team sports?  
*Answer: If the relative frequencies are not very different, then there is little evidence of association.*
  - c. Reset page 1.3. Given the data from the school newspaper survey, do you think there is an association between gender and sports preference based on this school's survey data?  
*Answer: Based on the relative frequencies, there does not appear to be much of an association.*
2. Select **Reset** and **Show Freq.**
  - a. Another school gave the same survey question to some of their students. The responses from the 50 boys surveyed were exactly the same (15 preferred watching individual sports and 35 preferred watching team sports). However, the responses from the girls were 20 preferring individual sports and 20 preferring team sports. Make the change in the table by editing the Girls Team cell from 50 to 20. What do you notice in the Counts graph?  
*Answer: More of the girls' bar is green than the boys' bar, but the counts are different since the total bars have different heights.*
  - b. Now select **Show Rel. Freq.** What does this graph tell you?  
*Answer: The relative frequency graph gives more clear evidence of an association, with the girls preferring individual sports more than boys and the boys preferring team sports more than girls. More of the girls bar is green than the boys bar, but now the total bars are the same height, so I can compare more easily.*



## Building Concepts: Two-Way Tables and Categorical Data

### TEACHER NOTES

- c. A third school also gave the same survey question to some of their students. In this school exactly the same number of boys and girls were surveyed, with 10 boys preferring individual sports and 50 boys preferring team sports, and 35 girls preferring individual sports and 25 girls preferring team sports.

Edit the table cells to show these counts. What do you notice about the two graphs now?

*Answer: Now both the Counts graph and Relative Frequency graph have two stacked bars of exactly the same height, so it is easy to compare in either graph. Both graphs show some evidence of association between gender and sports preferences.*



### Activity 2 [Page 1.5]

1. Using the frequency table (counts) or the relative frequency table, answer these questions.

- a. Is a 6<sup>th</sup> grader or an 8<sup>th</sup> grader more likely to like rap?

*Answer: If chosen randomly from the group of students surveyed, 6<sup>th</sup> graders are slightly more likely to like rap (0.36 compared to 0.30). (These values are close enough that it would be hard to say that a 6<sup>th</sup> grader or 8<sup>th</sup> grader chosen randomly from the entire school population is more likely to like rap.)*

- b. If you randomly selected an 8<sup>th</sup> grader who had taken the survey, what is the probability that the 8th grader preferred rap music?

*Answer: The probability is 0.30, represented by the relative frequency shown in the table (30% of the 8<sup>th</sup> graders surveyed preferred rap music).*

- c. If a student who took the survey prefers rap music, what is the probability that this student is an eighth grader?

*Answer: Of those students taking the survey, there were exactly the same number of 6<sup>th</sup> graders as 8<sup>th</sup> graders who preferred rap music (36). If one of these 72 students were chosen at random, the probability of that student being an 8<sup>th</sup> grader would be  $\frac{36}{72} = \frac{1}{2} = 0.5$*

- d. Is there an association between grade level and music preference?

*Answer: There is a slight preference of 6<sup>th</sup> graders for rap and 8<sup>th</sup> graders for rock.*

- e. What music do most of the 8<sup>th</sup> and 6<sup>th</sup> graders prefer?

*Answer: Country music, with half (0.5) of both 6<sup>th</sup> graders and 8<sup>th</sup> graders survey indicating this preference.*