CLEMSON MIDDLE SCHOOL MATHEMATICS PROJECT
UNIT 2: STATISTICAL STRATEGIES

PROBLEM 1: CURVING YOUR GRADES

Suppose that you and 9 friends take a quiz on Patterns and Functions. The scores are shown in the frequency table below.

SCORE	FREQUENCY
68	2
72	1
76	3
80	2
86	1
88	1

Because of an error on the quiz, your teacher decides to "curve" the scores and is considering two plans. In the first plan, your teacher adds 10 points to each student's score. In the second plan, your teacher adds 10% to each student's score (i.e., multiplies each score by 1.10).

A. For each plan, determine what happens to the mean score, the median score, the mode, the range of scores, and the Interquartile Range (IQR).

B. Construct a boxplot for the three sets of scores. Comment on the changes that the transformations cause to the graphs.

C. What will happen to the mean, median, mode, range, and IQR of any group of scores if x points are added to each score or if each score is multiplied by y?

MATERIALS

Casio *Algebra FX 2.0* Graphing Calculator

EXTENSION

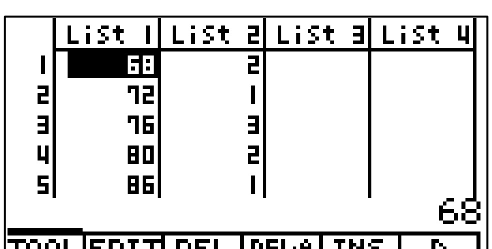
Determine what happens to the mean, median, mode, range, and IQR of a group of scores if 10 points are added to each score and then 5% is subtracted. Also determine what happens to the measures if 5% is first subtracted and then 10 points are added.

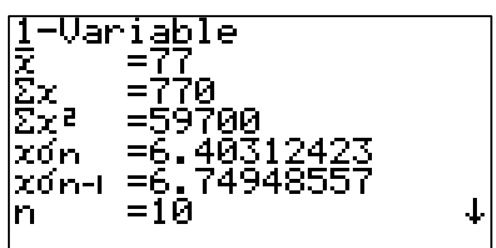
ONE SOLUTION TO PROBLEM 1: CURVING YOUR GRADES

A. For each plan, determine what happens to the mean score, the median score, the mode, the range of scores, and the Interquartile Range (IQR).

Let's first determine the statistics for the original data. From the MAIN MENU, choose STAT. Then,

- Clear Lists 1 through 4 if needed. To do so, have the cursor in the list you wish to delete, then press F6 (until you see DEL-A above F4) , then F4 for DEL-A (delete all) , then EXE for "yes" and the list should clear. Use the left or right arrow on the disc to move the cursor into any list that needs to be deleted and repeat the same process.
- Put the cursor in List 1. Type in each score (68, 72, 76, 80, 86, and 88) into List 1, pressing EXE after each entry.
- Move the cursor into List 2 and type in the frequencies (2, 1, 3, 2, 1, 1), again pressing EXE after each entry. See below left for the beginning of the lists.
- Press F6 until CALC (calculate) is available on the screen. Press F2 for calculate.
- Press 4 to check the setup for the calculations. 1Var XList should be List1. If not, when 1Var XList is highlighted, press F1 and type 1 then EXE . To set the frequencies, move the cursor down so that 1Var Freq is highlighted. If it says List2 then no change is necessary. If it does not say List2 then press F2 then 2 then EXE . Press ESC to return to the primary statistics screen.
- Press F2 for calculate and 1 for one variable. (One variable is the score on the quiz.) See below right for the beginning of the screen.





You can use the down arrow to move through the statistics. The statistics we are looking for from our original data are as follows:

?? The mean, denoted by $\bar{x}$, is 77.

?? The median, denoted Med, is 76.

?? The mode, denoted Mod, is 76.

?? The range, determined by subtracting the minimum (minX) from the maximum (maxX), is 88 ? 68 ? 20.

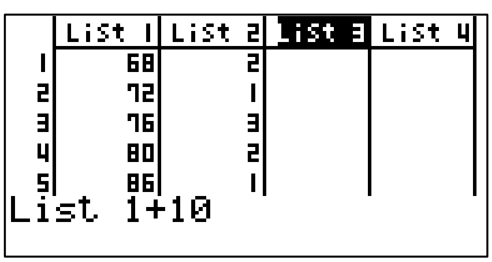
?? The Interquartile range, determined by subtracting the first quartile (Q1) from the third quartile (Q3) is 80 ? 72 ? 8.

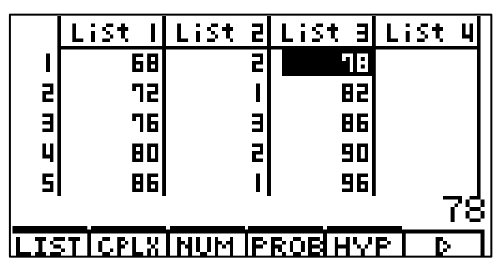
We now wish to explore what happens under the two plans. We will use List 3 and List 4 for our transformed scores. In List 3, we want to store the scores when 10 points are added to each score. We could do this quite easily by entering the transformed data one at a time, but the calculator can also do this with a technique that can be helpful when the transformations are more complex.

x If you are viewing the statistics, press [ESC] to return to the primary statistics screen.

x Move the cursor over to List 3 and then up so that List 3 is highlighted.

x Press [OPTN], then [F1] for the LIST menu, [1] to access individual Lists, and [1] to represent List 1, the addition sign, and 10. See below left. Press [EXE] and the list will be generated for you. See below right.

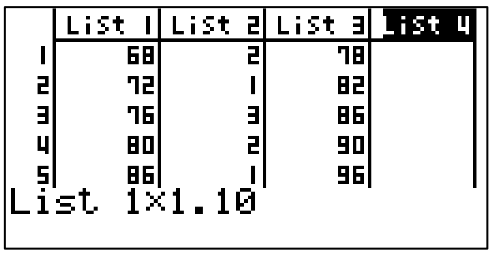


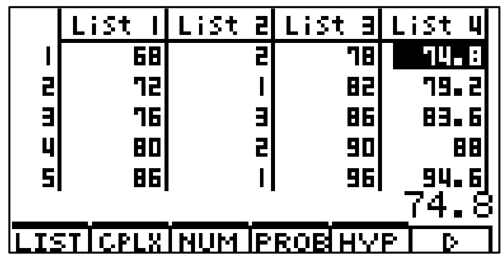


STATISTICS

Before we analyze the statistics, we will create the scores under the teacher's second plan in List 4. To obtain this list, we need to multiply each score by 1.10.

- Move the cursor over to List 4 and then up so that List 4 is highlighted.
- If List is not a menu option on the screen, press OPTN . Otherwise, press F1 for List menu options, 1 to choose List, then 1 again to choose List1, the multiplication sign, and 1.10. See below left. Press EXE and the new data should automatically appear. See below right.





We now wish to explore the statistics.

- Press ESC to return to the primary statistics screen.
- Press F2 for calculate.
- Press 4 for Set . For the 10-point addition plan, from List3, 1Var XList should be List 3. For the 10% plan, from List4, 1Var XList should be List 4. For both plans, the 1VarFreq should remain List 2. When you have the desired setup, press ESC to return.
- Press F2 for calculate and 1 for one variable.

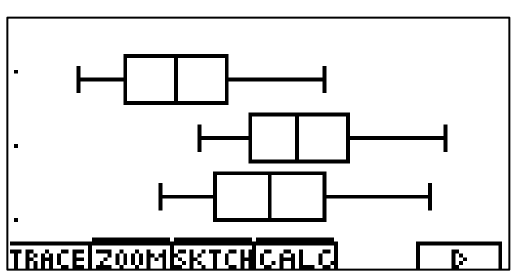
The statistics of interest are shown in the table below.

STAT/PLAN	**ORIGINAL**	**ADD 10 POINTS**	**ADD 10%**
MEAN	77	87	84.7
MEDIAN	76	86	83.6
MODE	76	86	83.6
RANGE	20	20	22
IQR	8	8	8.8

Under the 10-point addition plan, the mean, median, and mode increased by 10, while the range and IQR remained the same. When 10% was added to each score, all of the measures, including the range and IQR, increased by 10%.

B. Construct a boxplot for the three sets of scores. Comment on the changes that the transformations cause to the graphs.

- From the primary statistics screen, press CTRL F3 for the SET UP. Make sure the StatWind is set to Auto (press F1 if necessary when StatWind is highlighted). Press ESC to return.
- Press F1 for graph, 5 for Set (set-up). For StatGraph1, use the down arrow to GraphType. Press F6 for more options, and F2 for Box. Move down and highlight Xlist; select F1 if needed to change to List 1. Move down to Frequency and press F2 if needed to change to List 2.
- Use the up arrow key. Highlight StatGraph1, and press F2 to set StatGraph2. Again choose box plot by repeating the above steps, but change Xlist to List 3. Frequency should remain List 2.
- Use the up arrow key again to highlight StatGraph2 and press F3 to set up StatGraph3 as a box plot with List 4 as the Xlist and List 2 as the Frequency.
- When all three graphs have been set, press ESC . Then press F1 for graph. Press 4 for Select. Use F1 and the down arrow to turn on all three graphs, then press F6 to draw the graphs.



It should be readily apparent that both plans shift the original data to the right. Also, the first plan shifts the minimum, the quartiles, the median, and the maximum all farther to the right than does the second plan. Students should analyze why. (Because even the highest score is less than 100, adding 10 points increases a score more than adding 10%. For example, for a person who scores 80, adding 10% only increases the score by 8 points.)

What may not be so obvious is that the bottom graph, the one in which 10% has been added, is slightly wider than the others. The numbers are not dramatically different, but note that the range on the first two graphs (easily obtained either from tracing the graphs or looking back at the statistics) is 20, but the range of the last graph is 22. This is because multiplying the lower scores does not add as much to the grades as does multiplying the higher scores.

C. What will happen to the mean, median, mode, range, and IQR of any group of scores if *x* points are added to each score or if each score is multiplied by *y*?

It may take students several examples to discover the following rules:

1) If *x* points are added to each score,
   - ?? *x* points are added to the mean
   - ?? *x* points are added to the median
   - ?? *x* points are added to the mode
   - ?? the range remains unchanged
   - ?? the IQR remains unchanged.

2) If each score is multiplied by *y*,
   - ?? the mean is multiplied by *y*
   - ?? the median is multiplied by *y*
   - ?? the mode is multiplied by *y*
   - ?? the range is multiplied by *y*
   - ?? the IQR is multiplied by *y*

PROBLEM 2: WHERE SHOULD WE MEET?

Five people who live along the same highway want to get together for a meeting. They live at mile markers 2, 4, 16, 28, and 50.

A. Where should they meet so that the total number of miles driven is the least possible?

B. Now compute the mean and median for the values. How does your result compare with these measures? Do you think this will always be the case, no matter what five values are used? Support your answer.

C. Suppose a sixth person, who lives at mile marker 80, joins the group. Where should the group meet so that the total number of miles traveled is the least possible? Generalize your results for any six numbers.

D. Finally, generalize your results to any set of numbers.

MATERIALS

Casio *Algebra FX 2.0* Graphing Calculator

ONE SOLUTION TO PROBLEM 2: WHERE SHOULD WE MEET?

A. Where should they meet so that the total number of miles driven is the least possible?

To begin, let's explore the problem without the use of technology. Suppose that a student suggests that we should use mile marker 30 to hold the meeting. We might use a table to determine how many miles must be traveled.

PERSON	DISTANCE FROM 30
2	28
4	26
16	14
28	2
50	20

Adding the numbers in the second column, we find that, if the meeting is held at mile marker 30, the total distance traveled is 90 miles.

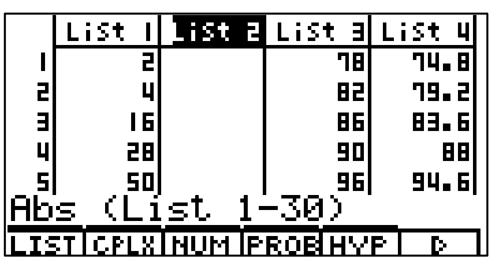
We now wish to see if we can do better. After students have tried a few values on their own in their search for the minimum distance, the teacher can demonstrate how technology can make their exploration easier. We can create a list in the calculator and then have the calculator compute the sum for us.

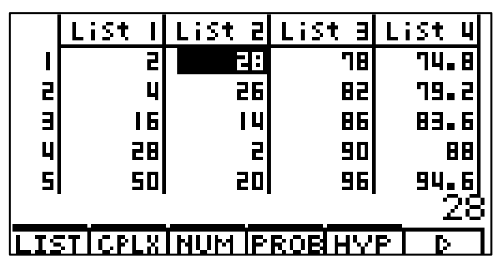
Before doing this, though, let's first make sure we know exactly what we want the calculator to do. For our guess of mile marker 30, we want to find the difference between 30 and the mile marker at which each person lives. However, we need to be careful; the distance must be positive. Consequently, for the person at mile marker 2, we want to subtract 2 from 30; however, for the person at mile marker 50, we want to subtract 20 from 50. In other words, sometimes we want to subtract the mile marker from 30, but at others we want to subtract 30 from the mile marker.

Fortunately, in mathematics, we have an easy way to ensure that the difference is positive – absolute value. With absolute value, no matter which value we put in first or second in our subtraction problem, the result is positive. We simply need to tell the calculator to compute the absolute value of the difference between 30 (or whatever our guess is) and the mile markers.

From the MAIN MENU, choose STAT. If necessary, press [ESC] until the lists appear on the screen. Then,

- Delete the values in Lists 1 and 2 by highlighting a number in the list, pressing [F6] until DEL-A appears at the bottom of screen, then press [F4] , then [EXE] to clear the list. Make sure both lists are clear.
- Move the cursor back into List 1. Type in the numbers 2, 4, 16, 28, and 50, pressing [EXE] after each entry.
- Let's again suppose our first guess was to meet at mileage marker 30. To determine how far each person would drive, we want to take the absolute value of the difference between the numbers in List 1 and 30. Move the cursor into List 2 and then up so that List 2 is highlighted. Press [OPTN] , [F3] for number, [1] for absolute value, the left parenthesis, [OPTN] , [F1] for the LIST menu, [1] for specific lists, [1] for List 1, the subtraction sign, 30, and the right parenthesis. See below left for this command. Press [EXE] . See below right for List 2.



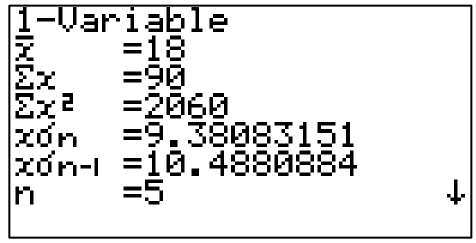


We will now check the statistics for our guess.

- Press [ESC] to return to the primary statistics screen and [F6] until CALC is at the bottom of the screen. Press [F2] for calculate and [4] for SET.

x Press [F1] to select a list number. Enter 2 to choose List 2, then [EXE] . Move down to 1VarFreq and press [F1] so that each value in List 2 is used exactly once. See below left. Then press [ESC] to return.

x Press [F2] to calculate and [1] for one variable. See below right for statistics.



We are looking for the sum of the numbers in the list, which is denoted by ? x. If the group meets at mile marker 30, the total distance traveled will be 90 miles.

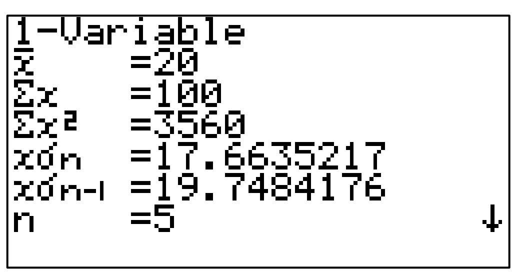
After perhaps many guesses, students may discover that the meeting should be held at mile marker 16. This would result in a total of 72 miles that must be driven.

B. Now compute the mean and median for the values. How does your result compare with these measures? Do you think this will always be the case, no matter what five values are used? Support your answer.

To obtain the statistics for our original list, press [ESC] so that CALC shows at the bottom of the screen.

x Press [F2] for calculate and [4] for SET.

x Press [F1] for LIST when the 1Var XList is highlighted. Then press 1 to choose list 1 and [EXE] . The frequency should still be one. Press [ESC] to return, [F2] to calculate, and [1] for one variable. See statistics below.



Many may have guessed that the mean, denoted by $\bar{x}$ and shown to be 20 miles, might be the solution to our problem. This, however, is not the case. Also, note that the median is 16 miles, which is, in fact, the answer to our problem. Students should explore many data sets to determine whether or not the median will always represent the value that results in the least total distance traveled. To begin the exploration, have them consider only data sets with an odd number of points. Once students have found that the median seems to work for all of these sets, we will return to the original set of values, {2, 4, 16, 28, and 50}, to try to see why this is true.

Let's start with the result, 15. If we move to the right of 15 a distance of one mile, we, of course, reach mile marker 16. This adds one mile to the people living at mile markers 2, 4, and 16, while reducing by one mile the people living at mile markers 28 and 50. Combining these, we see that moving to the right from 16 toward 28 adds one mile to the total distance traveled. If we move past 28 again toward the right, we add one mile to the people at markers 2, 4, 16, and 28, while reducing by one mile the person at marker 50. This, of course, adds 3 miles to the distance traveled.

Similarly, if we start at mile marker 16 and move left, we reduce by one mile the distance that the people at markers 2 and 4 must travel, but add one mile to the distance people at markers 16, 28, and 50 must travel. Combined then, we are increasing the total distance traveled.

To generalize, if we have an odd number of data points, the median will provide the smallest total distance to travel. As we move away from the median one mile at a time, we reduce by one mile the distance that those on that side of the median must travel. This, however, is one less than half the total number of people. We also add one mile to the distance that those on the other side of the median must travel, but we also add one mile to the distance the person at the median must travel. Consequently, the total distance traveled increases as we move away from the median. Thus, the median provides the minimum distance we can obtain.

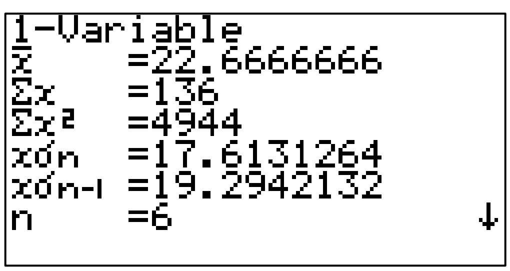
C. Suppose a sixth person, who lives at mile marker 80, joins the group. Where should the group meet so that the total number of miles traveled is the least possible? Generalize your results for any six numbers.

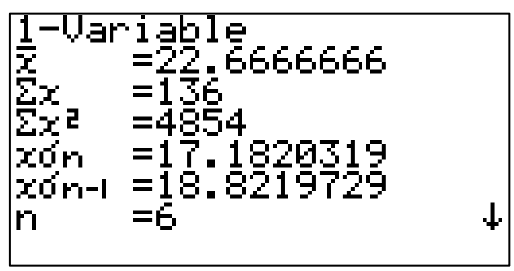
Our list now consists of {2, 4, 16, 28, 50, and 80}. As before, we can allow the calculator to determine the median for us and check the sum of the distances for us. First, we will go to List 1 at the main statistics screen and add the 80 at the bottom of the list. Simply go down to the bottom of the list, type in 80, and press [EXE] .

Next, we'll compute the statistics for List 1 to determine the median. Using the same steps as before (pressing [F2] followed by [1] to calculate the one variable statistics), we find that the median is 22. (Scroll down until Med appears on the screen.)

To see how many miles must be driven for the meeting, we will find the absolute value of the difference between our guess, 22, and each of the numbers in List 1. Using the techniques described earlier and then finding the statistics on List 2, see below left, we find that the sum of these distances is 136. See below left.

We now wish to determine if this is the best we can do. Let's move to right, one mile at a time. In other words, let's see the total distance that will have to be driven if the meeting is held at mile marker 23. Using the same techniques as before, we find that it is the same! See below right.





If we continue to move to the right or left of the median, we continue to obtain the same total distance, at least until we move to the left of 16 or to the right of 28. In other words, the meeting can be held at any mile marker between 16 and 28, inclusive, to minimize the distance. The question we should ask ourselves is why.

As before, let's consider what happens as we move to the right of the median, 22. When we move from 22 to 23, we add one mile to the people who live at markers 2, 5, and 16, but subtract one mile for the people who live at mile markers 28, 50, and 80. The result of adding three and then subtracting three is, of course, no net change. Until we move to the right of 28 (or left of 16), we will continue to add one mile to three people but subtract one mile from the other three. Once we move to the right of 28 (or left of 16), we add a mile to four people while subtracting one from only two.

D. Finally, generalize your results to any set of numbers.

Before working this problem, many people may have thought that the mean minimizes the total distance from each point. This, however, is not true. The median, which is not affected by outliers (extreme values) accomplishes the task.

With an odd number of data points, there is a single middle number in the set. For example, if a set has 9 values, the 5^{th} value in order is the median – four values are below this and four are above. In general, if there are n values and n is odd, the median is the $\frac{n ? 1}{2}^{st}$ value, assuming the numbers are in either ascending or descending order. The sum of the distances from this value is the minimum possible.

With an even number of data points, the minimum total distance is obtained by using any value between the two middle-most points, including these two points. Consider the possible implication. Often students are taught to find the mean of the two middle-most points when calculating the median for a data set with an even number of data points. This problem suggests, however, that any value between the two middle-most data points, including the points, can be used. What do you think?

PROBLEM 3: WHAT SCORE DO I NEED?

Suppose you need an average of 90 to earn an A in a course, a grade you desperately want. Your scores so far are 96, 99, 84, 85, and 91, and you have one last test to take.

A. What score do you need on the sixth test to obtain the 90 average, assuming the teacher uses the mean to compute the average? What score do you need on the sixth test if the teacher used the median to determine the grade? Verify your results by computing the mean on your calculator.

B. Change the third score from 84 to 0. Now what score do you need on the sixth test to obtain a mean of 90? What score do you need to obtain a median of 90? Again verify your results.

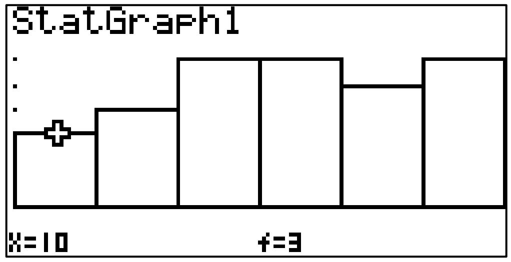
C. Should the teacher consider the mean or the median when determining grades? Is there another measure that might be better to use? Support your answer. Be sure to consider the issue of fairness in your response.

EXTENSIONS

Create a procedure that will determine the final value needed for a set of numbers to obtain a preset mean. (For example, create a method that will determine the needed sixth score for the scores above to obtain a mean of 90 or any other number you might choose.) Compare your procedure to the others in your class and decide who has created the most efficient method.

PROBLEM 4: CLASS GRADES

Consider the histogram depicting the scores of 30 students on a 20-point quiz. The scores were all even numbers, ranging from 10 to 20. Each bar represents a different score, and the vertical scale has been set at 1.



A. Construct a frequency table that depicts the scores. Use your calculator to construct a histogram graph to confirm your results. (From the Statistics menu, check the set up, CTRL F3 , and make sure the StatWind is manual.)
B. Determine the class median, mean, range, and Inter-Quartile-Range.
C. Construct a box and whisker plot of the data.

TEXT SECTION CORRESPONDENCES

The materials in this module are compatible with the following sections in the listed texts.

TEXT	SECTION
AW – Foundations of Algebra and Geometry (1998)	1.2-3; 2.3; 3.3; 8.1
Glencoe – Mathematics Applications and Connections C1 (1995)	2.6-7; 3.10
Glencoe – Mathematics Applications and Connections C2 (1995)	3.2-3; 3.5; 3.7
Glencoe – Mathematics Applications and Connections C3 (1995)	2.6; 4.2; 4.5-7
Houghton Mifflin – The Mathematics Experience I (1992)	12.4; 15.4
Houghton Mifflin – The Mathematics Experience II (1992)	7.5; 14.3
McDougal Littell – Gateways to Algebra and Geometry (1994)	5.2; 5.4-5; 6.7; 11.7
Prentice Hall – Middle Grades Mathematics C1 (1995)	1.3; 1.7; 10.8
Prentice Hall – Middle Grades Mathematics C2 (1995)	1.1; 1.4; 6.7
Prentice Hall – Middle Grades Mathematics C3 (1995)	1.4; 1.8
SFAW – Middle School Math C1, V1 (1999)	1.7-9
SFAW – Middle School Math C1, V2 (1999)	
SFAW – Middle School Math C2, V1 (1999)	1.4-5
SFAW – Middle School Math C2, V2 (1999)	
SFAW – Middle School Math C3, V1 (1999)	1.2; 1.3
SFAW – Middle School Math C3, V2 (1999)	11.1
SFAW: UCSMP – Transition Mathematics, Part 1 (1998)	4.1
SFAW: UCSMP – Transition Mathematics, Part 2 (1998)	11.3

