Some Ideas on Teaching Data Management

A. Craig Loewen

The two most significant recent changes to Alberta's elementary mathematics program include the dramatic change in philosophy and the revisions which resulted in the new data management strand.[1]

Why Teach Data Management?

The data management strand (Alberta Education 1994) involves all the concepts formerly contained within the graphing strand of the 1982 Alberta curriculum but goes much further to entail concepts in collecting, recording, organizing, displaying and interpreting data, as well as concepts in probability, including theoretical and experimental probabilities.[2] These additions to the math program are extremely positive for several reasons.

Life Skills

The new data management topics represent a positive addition to the new program in that the skills addressed are in fact life skills. Our students need to know how to manipulate data, and, in an increasingly technological society, the ability to code, sort and interpret data has become paramount. It is perhaps arguable that data management skills are almost as important as number facts and number sense. The value of these skills justifies the increased amount of time spent in the data management strand.

Relevance

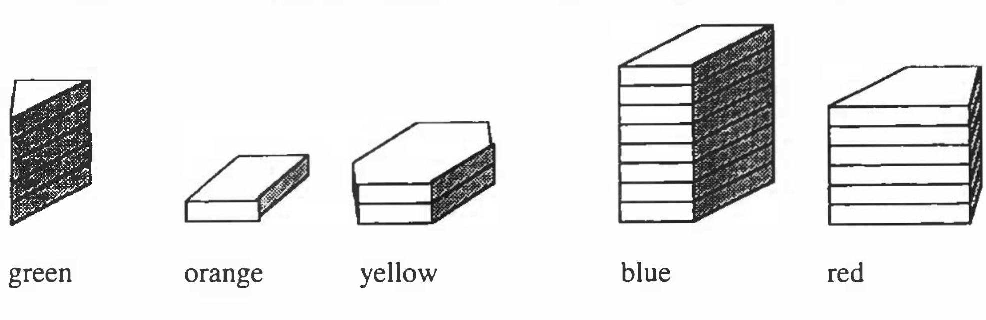
Mathematics has traditionally been criticized for its perceived lack of relevance at all grade levels. Even the simplest concepts have often been taught without context and without reference to any physical object. For example, it is possible to teach a simple addition fact, such as 2 + 1 = 3, without referring to concrete objects, but, when we contextualize it by talking about the joining of sets of fruit or children or bicycles, the concept takes on much more interest and meaning. In short, we forget at times that mathematics has a purpose: mathematics summarizes and represents patterns, events and regularities within our world. Mathematics describes the world around us. The data management strand has its strength in that it emphasizes that numbers and graphs represent information and as such are interpretable. This emphasis enhances the general sense of the relevance of the study of math.

Integration

The data management strand is easily integrated with most other strands of the elementary curriculum. For example, one can easily visualize how collecting temperatures on given school days for a school year and charting them would address objectives within the data management and measurement strands. Whenever we strive to integrate strands, concepts or subjects, students are typically left with an enhanced sense of understanding and purpose of the study.

Discussion

The topics in the data management strand lend themselves nicely to discussion. A familiar example involves having students select a pattern block of their favorite color. Once the selections are made, students are asked to come forward by color and stack their pattern blocks in towers of the same color. A concrete graph is created as follows:



Viewing the concrete graph, questions such as the following emerge: Do more people prefer red or blue? How many people are in the classroom? What is the least favorite color? The last question typically surprises students as it is not answerable from the graph alone. All possible colors are not represented in a set of pattern blocks—pattern blocks are painted red, green, yellow, blue, white or orange. Is it possible that brown or purple is the least favorite color? It is not difficult to see how an interesting discussion quickly emerges. Of course, interesting discussions are possible within any strand of the curriculum, but interpreting displayed data is a particularly rich source of such dialogue.

Managing Data: An Example

The data management strand separates into three main topics: collecting and recording data, organizing and displaying data, and interpreting displayed data (Alberta Education 1994). The following example illustrates these components:

> My friend is opening a men's clothing store. He asked me to help him decide how many of each size belt he should stock.

Collecting and Recording Data

To assist my friend, I agreed to stand on a busy street corner and ask the first 50 men who passed to tell me their belt size (a risky proposition, but I thought it might be interesting). I recorded the following numbers:

30 44 32 36 32 38 32 44 28 42 30 56 34 36 34 34 26
28 42 36 32 26 30 38 34 40 32 46 30 28 24 36 40 34
30 38 32 32 36 36 30 42 38 50 40 40 36 28 28 38

It is easy to see that the collection of data above has little obvious order. We might notice that the numbers are all even (belts are usually sold only in even sizes) and that several values are in the "30 range." Beyond these simple generalizations (which are not much help to my friend), the only order we know exists is that the numbers are listed according to the sequence in which they were collected. The lack of order forces us to consider how we might better manage or organize the data.

Organizing and Displaying Data

Organizing and displaying the data involves many levels of sophistication. For example, a simple reorganization involves merely listing the numbers from lowest to highest:

24 26 26 28 28 28 28 28 30 30 30 30 30 30 32 32 32 32 32
32 32 34 34 34 34 34 36 36 36 36 36 36 36 38 38 38 38 38
40 40 40 40 42 42 42 44 44 46 50 56

Once the data is sorted and listed, information is much easier to find, such as, What is the largest belt size? What is the smallest belt size? What is the range of belt sizes? We may even (with a little more effort) be able to determine the most and least common belt sizes. With such a list, we can also determine the median belt size (by finding the middle value in the list) or the most common belt size (the mode).

However, our list is still somewhat visually intimidating. To find the most common belt sizes requires us to count items in our list, and it is hard to get a quick sense of how the values cluster. So, we may consider more sophisticated ways to display the data, such as tally charts and stem-and-leaf plots. The tally chart requires us to make some decisions, such as, do we wish to group together any ranges (or sizes) to create a more succinct presentation? For example, in the tally chart below, the data is presented in ranges of 24–28, 30–34, 36–40, 42–48 and 50–56. These ranges may approximate the categories small, medium, large, extra large and extra-extra large.

24–28	S	𝍸 |||
30–34	M	𝍸 𝍸 𝍸 |||
36–40	L	𝍸 𝍸 ||
42–48	XL	𝍸 |
50–56	XXL	||

Notice however that our grouping of data has caused us to lose some information: it is easier to see how many people fall in particular ranges, but we no longer know exactly how many people claimed to have a 36-inch waist. In comparison, the stem-and-leaf plot helps us maintain this information. In the stem-and-leaf plot, we define the stem as the digit in the tens place and the leaf as the digit in the ones or units place. We list all the leaves after the appropriate stem. Thus, our method of coding the data enables us to regenerate the entire list of raw data. A stem-and-leaf plot for our data is shown below.

2	4 6 6 8 8 8 8 8
3	0 0 0 0 0 0 2 2 2 2 2 2 2 4 4 4 4 4 6 6 6 6 6 6 6 8 8 8 8 8
4	0 0 0 0 2 2 2 4 4 6
5	0 6

In this plot, we can quickly see that more people are using a belt in the 30–38-inch range than in any other range. In comparison to the tally chart, in the stem-and-leaf plot, the means, median and mode are no less difficult to find; however, our ranges for grouping data are determined by place value, and in this case do not easily translate to the familiar method of sizing (S, M, L, XL, XXL). The primary advantages of the stem-and-leaf plot are that it is relatively easy to construct and that it can be easily created from unsorted raw data.

Though the stem-and-leaf plot is more visually interesting and appealing than the tally chart, other forms of graphs such as bar graphs, line graphs, histograms and pictographs are even more visually captivating. Figure 1 shows our data represented as a bar graph. This graph is easy to read but is significantly more difficult to construct than either the tally chart or stem-and-leaf plot. Creating a bar graph also requires making many decisions. We must decide what to place along our vertical and horizontal axes, not to mention vertical and horizontal spacing and what the spacing may represent. But, in a quick glance, one can easily determine the range, identify the most and least common measurements, compare frequencies of given measurements and get a strong sense of the distribution of the data.

Interpreting Displayed Data

Whenever we try to get information from a graph or chart, we are interpreting displayed data. Many examples have already been given in the discussion above, but more difficult questions can also be posed. For example, Are any belt sizes sufficiently uncommon that you would not want to order for your store? Do any values in our graph seem surprising (for example, why is the 34-inch belt size less frequent than the 32-and 36-inch belt sizes when we would expect it to be more common)? How should we account for this anomaly in placing our order?

We may wish to experiment with our axes. For example, let's assume that my friend wishes to order approximately 500 belts for his store. By changing the values along the vertical axes (that is, multiplying each value by 10), we essentially scale the data and get a reasonable estimate of how many of each size belt should be ordered (Figure 2).

The example developed above helps us to understand the general flow involved in managing data. We begin by collecting the data (either first- or secondhand) and then proceed to the next step of sorting, organizing and displaying the data in different ways. Once displayed, our challenge is to interpret the data reasonably to solve the given problem or task. At each step along the way, the data becomes increasingly meaningful and useful.

Sample Activities

Defining an imaginary problem or posing a question like the one above is an interesting and enjoyable class activity. A collection of other activities and problems emphasizing particular stages of the data-management process are presented below. The activities represent a broad range of difficulty from relatively easy (for example, Race to the Top) to very difficult (for example, Bar Graph Clues).

Activity One: Find Me Five (Game)

Topic

Collecting Data

Objective

Formulate the questions and categories for data collection, and actively collect firsthand information (Alberta Education 1995, 54).

Materials

Paper, pencils

Rules

- Each player begins by selecting a question to which a person would answer *yes* or *no*. For example, the player may ask, "Did you have toast for breakfast this morning?"
- Once each player has selected a question, the players circulate around the room asking the question of 10 people.
- Players should keep lists of the names of people who answered *yes* and who answered *no*.
- The player who receives exactly five affirmative responses to his or her question wins.

Discussion

The difficulty of this game rests in trying to determine a question that would elicit exactly five affirmative responses. Asking students how they might adapt or change their question to achieve the goal may result in some interesting dialogue. The game also lends itself to discussions of how you may be able to bias your sample. For example, you could ask, "Are you wearing white socks?" and proceed to deliberately ask five people who you know are wearing white socks and five people you know are not. Answers to some questions are easier to predict than others.

Activity Two: Tally Woe (Problem)

Topic

Collecting Data

Objective

Organize data, using such graphic organizers as diagrams, charts and lists (Alberta Education 1995, 54).

Materials

Paper, pencil

Problem

Anita has noticed that half the students in her class have blue eyes. In your class, do more than half of the students have blue eyes?

Discussion

The solver will need to set up some method for surveying his or her classmates to determine their eye color and develop some method for recording the results. Depending on the grade level at which this problem is introduced, the teacher may need to discuss the notion of one-half, specifically the number of students that would constitute one-half in that class. The teacher may also wish to discuss how tally charts are constructed and used. For example, does it make sense to collect information in a list or other form and later sort it into a tally chart, or are there benefits (and difficulties) associated with entering the data directly into a tally chart as it is collected?

Activity Three: Race to the Top (Game)

Topic

Organizing and Displaying Data

Objective

Organize data, using such graphic organizers as diagrams, charts and lists. Describe the likelihood of an outcome using such terms as more likely, less likely, chance (Alberta Education 1995, 54 and 58).

Materials

Bar Graph Gameboard (Figure 3), pencil crayons (red, green, yellow, blue), Race to the Top Spinner Mats (Figure 4), overhead spinner

Rules

- This is a game for two players playing on the same gameboard (bar graph).
- The first player places the overhead spinner on any one of the spinner mats and twirls the spinner. The player colors a space at the bottom of the bar graph according to the results of the spin. For example, if the spinner points to red, the player would color the first block in the bar labeled red.
- The spinner in the bottom right-hand corner of the spinner mat has two possible outcomes: none and choice. If the player chooses that spinner, and the spinner points to none, then the player passes the turn to his or her opponent. If the spinner points to choice, the player must shade the next block in any one bar she or he chooses.
- The second player now chooses a spinner mat and twirls the spinner, recording the results of his or her spin.
- The players continue taking turns building up the bars of the bar graph until one bar reaches a height of 10.
- The player who completes any bar graph (that is, shades the tenth block in any bar) wins.

Discussion

The teacher may find it useful to discuss strategy with the students. Students should realize that turning over the bar graph to their opponents is not desirable while the height of any bar rests at nine. This strategy requires careful choice of the spinner mat on each turn. If working with younger students, the teacher may wish to let them each have their own bar graph and race each other to complete any bar or the entire graph. To add further variety to the game, the teacher can encourage students to create their own spinner mats. This game emphasizes organizing and displaying data in that students are essentially creating bar graphs to display the results of their experiment with the spinner.

Activity Four: Bar Graph Clues (Problem)

Topic

Organizing and Displaying Data

Objective

Organize data, using such graphic organizers as diagrams, charts and lists (Alberta Education 1995, 54).

Materials

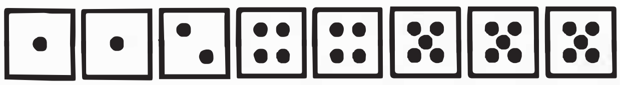
Paper, pencils, beans or small markers

Problem

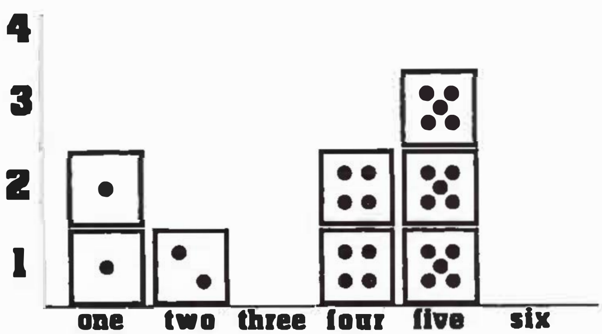
Gwenyth asked eight people to roll a die while she recorded the value they rolled (one through six). Gwenyth noticed more people rolled a 4, 5 or 6 than a 1, 2 or 3. She also noticed that more odd than even numbers were rolled and that the sum of the values rolled was 27. Construct a bar graph to show how many people rolled each of the values one through six.

Discussion

This somewhat difficult problem has four solutions. The teacher may wish to introduce the problem by distributing dice to the students and having them conduct the same experiment as Gwenyth. The students could talk about their firsthand data before being introduced to Gwenyth's secondhand data. There are many approaches to solving this problem, including making a list or chart and adjusting that list or chart until all conditions of the problem are met. The teacher may also consider having students model the problem using eight dice and a guess-and-check strategy. In this way, students can easily manipulate the dice until all conditions are met. See the following example:



The students can then rearrange the dice to create a concrete graph like the one below before constructing their bar graph to complete the problem.



Activity Five: Bar Graph Battleship (Game)

Topic

Interpreting Displayed Data

Objective

Construct and label concrete/object graphs, pictographs and bar graphs. Generate new questions from displayed data (Alberta Education 1995, 54).

Materials

10 each of red, green, blue and yellow counters; pencil; Bar Graph Gameboard (Figure 3) for each player; paper bag

Rules

- This is a game for two players.
- Each player randomly draws a number of red, green, blue and yellow counters from a paper bag and creates a bar graph on the left-hand side of his or her gameboard to display the counters drawn. This bar graph is kept hidden from his or her opponent.
- Once both players have created their bar graphs, they take turns asking each other questions which can be answered with *yes* or *no*, attempting to determine how many of each color the other player drew. For example, the first player may ask, "Did you draw more red than blue?"
- The second player truthfully answers the question, and the first player notes the answer on the right side of his or her gameboard. The second player now asks a question.
- Players continue asking each other questions until one player announces, "I know how many of each color you have drawn." This player must construct/complete the bar graph she or he believes his or her opponent created at the beginning of the game.
- Players now compare the appropriate bar graphs. If the player has determined the correct number (and has constructed the correct bar graph), that player wins. If a mistake has been made, the other player wins.

Discussion

This game is not difficult for most students who are familiar with similar battleship games. Nonetheless, the teacher may wish to introduce the game using only one color, and over time introduce the remaining three colors. Some discussion time may be required to help students develop a system for recording opponents' answers. Teachers may also wish to discuss the elimination problem-solving strategy before starting the game. Younger students tend to ask simple direct questions such as, "Do you have five blue counters?" The teacher may wish to encourage other types of questions as well.

Activity Six: Train of Thought (Problem)

Topic

Interpreting Displayed Data

Objective

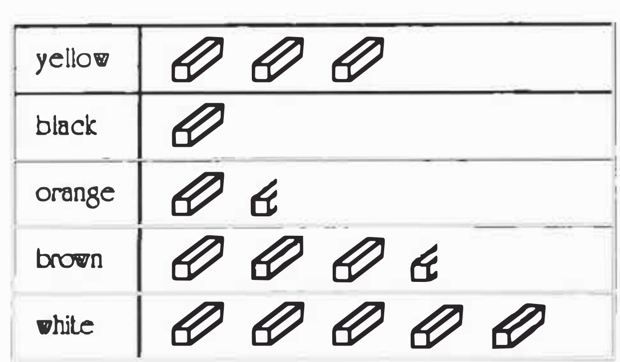
Discuss data, and draw and communicate appropriate conclusions (Alberta Education 1995, 54).

Materials

Paper, pencil, colored counting rods, metre stick

Problem

Tannis selected several colored counting rods and created a pictograph to show how many she picked of each color. If her rods were placed end to end, how long would her train be?



Scale: = 2 blocks

Discussion

A reasonable approach to this problem would be to find a set of colored counting rods and generate the set of rods represented by the given pictograph. These rods could then be placed end to end along the edge of a metre stick to create a train. Another solution would require finding one rod of each color mentioned, measuring it, and then multiplying and adding the lengths to determine the total length. Other interesting problems could be posed, such as How many of each color rod listed would you need to create a train exactly 1 m in length? What train lengths could you create if you must have the same number of each color of rod? Students may enjoy creating their own problems and graphs to challenge their friends and classmates.

Notes

1. The concepts found in the data management strand of the October 1994 Alberta Program of Studies are contained under the statistics and probability strand in the Western Canada Protocol.

2. These topics are presented in two different strands in the Western Canada Protocol: statistics and probability (data analysis) and statistics and probability (chance and uncertainty).

References

Alberta Education. *Program of Studies Elementary Schools: Mathematics (Interim Document)*. Edmonton: Author, 1994.

——. *The Common Curriculum Framework for K–12 Mathematics*. Edmonton: Author, 1995.

Figure 1: Belt Sizes Data Presented as a Bar Graph

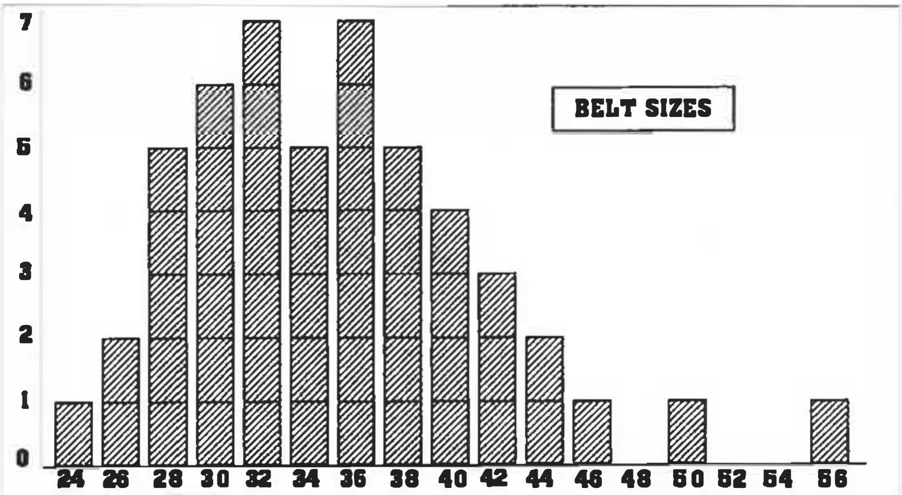


Figure 2: Bar Graph with Vertical Axis Scale Adjusted for 500 Belts

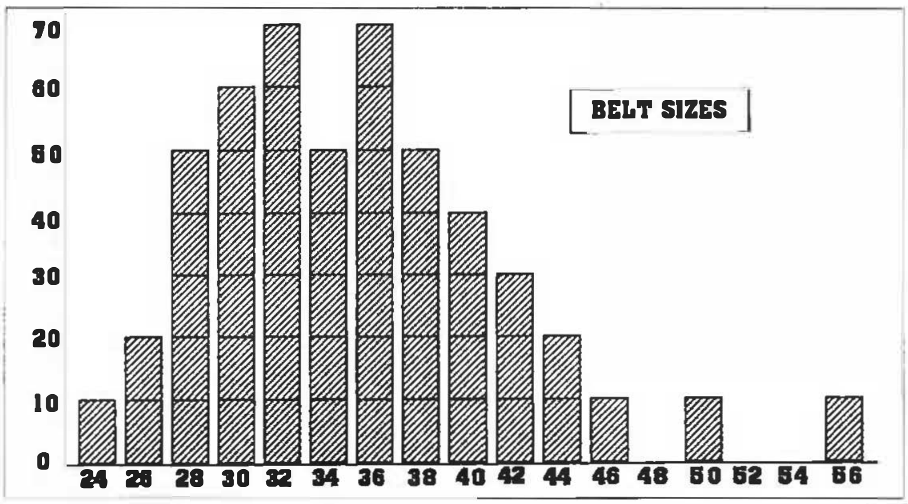


Figure 3: Bar Graph Gameboard

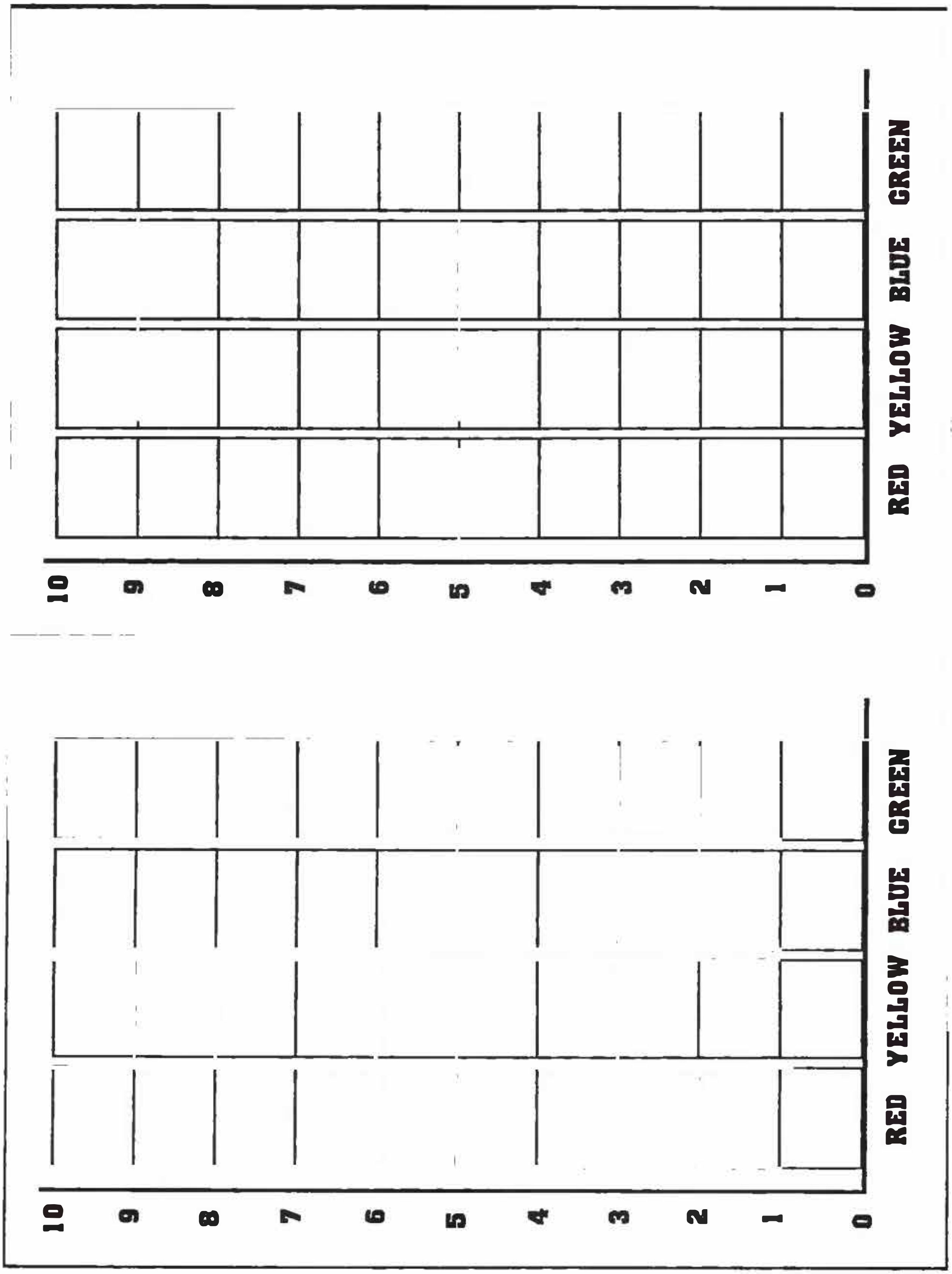


Figure 4: Race to the Top Spinner Mats