Exploring the Confusions: Bar Graphs

Susan Hillman

Saginaw Valley State University, Bay City, USA

Professional development with practicing teachers of mathematics in the middle grades from the U.S. and India revealed confusions about the construction and interpretation of data with bar graphs when working with certain kinds of data. Textbooks often influence teachers' decisions about what and how mathematics is taught. Textbooks used in both countries were examined for definitions, descriptions, and examples that highlighted (or obscured) features of bar graphs. Potential sources of confusion for teachers (and ultimately their students) regarding this type of graph were identified. Implications and recommendations for teaching data analysis with bar graphs are provided.

Introduction

Data analysis is increasingly important in school curricula due to our sophisticated technological world that relies on interpreting many forms of information for a multitude of purposes. Representations of data frequently used and seen in various media include bar graphs. Bar graphs highlight comparisons of certain aspects of data, making them more or less appropriate for certain types of data, and for certain purposes. Representing numerical data with bar graphs may be challenging, with potential for confusion when teaching bar graph construction and interpretation of data displayed in bar graphs.

Bar graphs use bars to indicate something about the data. What do these bars represent? What does the length of each bar represent? What features of the data does a bar graph highlight? When representing numerical data, when is it appropriate to use a bar graph? These questions surfaced while working with maths teachers in South India. Three teachers from different schools taught the same lesson on bar graphs using the same curriculum. During each lesson, the maths teacher constructed a bar graph to represent numerical data. I anticipated seeing a bar graph where the length of each bar represented the frequency (e.g., number of students) of each data element (e.g., height). However, each teacher (without consulting one another) constructed a bar graph with one bar for each frequency (number of students) along a carefully drawn scale for the frequencies and the lengths of each bar represented the data elements (height) – the opposite of what was expected. Discussions with these teachers were based around the questions above and prompted further collaborative study.

Teachers of maths in the U.S. revealed similar confusion when representing data with bar graphs. After all, one can "read" the same information (e.g., five people are 164 centimetres tall) from a bar for the frequency (5 people) with a bar length showing a data element (164 centimetres tall) as from bars representing data elements with bar lengths showing frequencies of the data elements. The latter is a conventional bar graph, where it is easy to compare frequencies (or magnitudes) of data elements since the bar lengths highlight this aspect of a set of data. Additionally, switching the meaning of the bars and bar lengths is problematic when there is more than one data element with the same frequency or magnitude. Understanding how bar graphs show these aspects of reading and interpreting data is essential to consider, when constructing them to display data meaningfully.

Theoretical framework

Three areas of research relate to this study: mathematical

knowledge for teaching mathematics (Ball, et al., 2005; Ma, 1999); graph comprehension (Friel et al., 2001; Russell et al., 2002; Shah & Hoeffner, 2002); and the nature and role of curriculum for teaching mathematics (Hiebert, 2008; Remillard, 2005; Stein et al., 2000; Tarr et al., 2006). Teachers' knowledge of mathematical content (e.g., constructing and interpreting data from bar graphs) interacts with the intended curriculum as presented in textbooks (e.g., definitions and examples of bar graphs) resulting in the enacted curriculum (i.e., learning opportunities for students).

Mathematics Knowledge for Teaching Mathematics

Ma (1999) studied the mathematical knowledge of teachers and described teachers' "profound understanding of fundamental mathematics" as one factor that significantly influenced learning opportunities for students. Others describe mathematical knowledge for teaching as "a kind of professional knowledge of mathematics different from that demanded by other mathematically intensive occupations" (Ball et al., 2005). Aspects of this knowledge that relate to this paper include teachers' ability to do the maths expected of students, explain the meanings of concepts and procedures in appropriate ways for students' understanding, and ability to use appropriate and meaningful representations (Ball et al., 2005).

How teachers understand the features of bar graphs related to highlighting aspects of categorical and numerical data may influence learning opportunities. Questions regarding how the bars represent data, and sorting out important characteristics of data such as whether the data is numerical or categorical and whether numbers indicate magnitude or frequencies, are important to consider when using bar graphs. Specifically, when using numerical data, one must be clear about which set of numbers represents the cases or categories and which set of numbers represents the case values (magnitudes) or frequencies for appropriate graph construction and comprehension.

Graph comprehension

Research on graph comprehension provided some guidance for critical factors to consider when a search for research on bar graph comprehension brought no results. Interpretation of graphs involves characteristics of visual display, knowledge of conventions in graph construction, and knowledge of the content represented in the graph (Shah & Hoeffner, 2002). Bars are the main visual characteristic of bar graphs, so conventions for using bars to represent data involves knowledge of types of bar graphs and data, and conventions in construction such as using the length of a bar to represent frequencies or magnitudes of data elements, equal-width bars, and equally spaced bars.

Different types of bar graphs are used to display characteristics of data. A case value bar graph uses the length of a bar to show magnitude of individual data elements with no data reduction, whereas a frequency bar graph involves simple data reduction by using the length of a bar to display the number of occurrences of data elements (Russell et al., 2002). Distinguishing between sets of data as discrete cases, discrete categories, or grouped numerical data along some scale, and whether frequencies or magnitude of data elements are provided, is also critical for constructing appropriate representations of the data.

Knowing the ways these types of bar graphs represent certain types of data may help teachers make decisions about the level of complexity for instruction. Transition from case value bar graphs with no data reduction to frequency bar graphs with simple data reduction may be confusing if this transition is not carefully considered and explored because the axes must be redefined (Friel et al., 2001). Three levels of graph comprehension include reading the data, reading between the data, and reading beyond the data (Friel et al., 2001). Even the simplest level of graph comprehension, reading the data, requires understanding the conventions of graph construction (Friel et al., 2001; Shah & Hoeffner, 2002).

Nature and Role of Mathematics Curriculum

Teaching inevitably involves the ways that teachers interact with students about the mathematics content (Hiebert, 2008). Mathematics curriculum includes the intended curriculum, meaning what students should learn as presented in textbooks, and the enacted curriculum, meaning the learning opportunities during instruction in classrooms that reaches students (Hiebert, 2008; Stein et al., 2000). How teachers understand and use the mathematics content in commonly used textbooks influences the instruction and learning opportunities for students (Stein et al., 2000). Teachers' knowledge, beliefs, and dispositions influence the ways they understand and implement mathematics instruction using textbook (and other) curricula materials (Remillard, 2005; Tarr et al., 2006). The textbook plays a major role in guiding mathematics teachers in India and the U.S. when making decisions about what and how to teach. In India, the textbook is literally the curriculum (from the national or state level) to help prepare students for annual exams. In the U.S., the textbook is a surrogate curriculum since many claim alignment with state curriculum standards and are approved by school districts for instruction. Instructional decisions for enacting the curriculum influence the learning opportunities for students, whether the decisions involve a

degree of fidelity to the printed textbook material or selecting examples and adapting materials for instruction.

Recognizing the potential of textbooks for supporting teaching and learning in mathematics classrooms, this paper focuses on a textbook analysis of representing and data in bar graphs.

Methodology

Ten textbooks were selected and examined for definitions, descriptions, and examples used in developing the ideas of representing data with bar graphs. Nine teacher's guides were selected from current standards-based and commonly used textbooks for teaching middle grades mathematics in the U.S. One textbook used by teachers and students, and published as a national curriculum in India, was selected based on observations of lessons taught from that text. A total of ten textbooks selected for this study ranged from fourth through eighth grades.

Definitions, descriptions, and examples of bar graphs from each textbook were compiled, sorted, and compared. These definitions, descriptions, and examples were then analysed for features related to conventions in graph construction (e.g., bar placement and representation of data). The content of definitions, descriptions, and examples was examined for content knowledge needed to make sense of data in bar graphs (e.g., type of data and type of bar graph). Communication of key ideas about the content and conventions in constructing bar graphs identified essential and sufficient information for constructing conventional bar graphs, and clarity of language to communicate meanings (e.g., using unambiguous language). Points of confusion were identified based on absence of essential features in definitions or descriptions, emphasis of non-essential information, or use of language that was ambiguous in meaning. Examples were used to test textbooks' definitions and descriptions.

Results

Definitions and descriptions of bar graphs were quite varied across grade levels, from short and vague (e.g., use bars to show data) to detailed descriptions explicitly mentioning multiple features of bar graphs. Earlier grades textbooks tended to mention fewer features of bar graphs and tended to be less descriptive than textbooks from later grades. Examples of bar graphs tended to be more alike than different across grade levels regarding elements of visual display and conventions in constructing such graphs. One textbook assumed prior knowledge of bar graphs and had no definition or description, but did include examples of bar graphs. Another textbook had a definition of bar graphs but included no examples of bar graphs.

Conventions in constructing bar graphs

Conventions in constructing bar graphs include using bars to represent data elements (e.g., categories, individual cases, or a group of numbers such as different time periods during a day). The lengths of the bars represent the frequencies of data elements in a frequency bar graph; the length of each bar indicates the magnitude (or measure or amount or value) of each individual case (e.g., the amount of money earned as profit for a particular company). Four out of nine textbooks specifically mentioned that bar length shows frequency or magnitude. The rest did not explicitly state the meaning of bar lengths, but all examples were consistent with using bar length to show frequencies or magnitude of data elements.

Additionally, conventions for constructing bar graphs include equal bar width and equal spacing between bars. All examples of bar graphs had equal bar width, but only two textbooks explicitly mentioned that bars should be of the same width. All textbooks except one had equal spaces between bars, but only one textbook mentioned this convention explicitly. One textbook included some examples with no spaces between bars even though the data was categorical (e.g., kinds of pets such as cat, dog, etc.), along with examples having equal spaces between bars.

Bar orientation was explicitly mentioned by four textbooks. Three of those indicated bars could be drawn vertically or horizontally; the fourth textbook specified vertical bars by defining the x-axis for showing categories and the y-axis for showing frequencies. Although seven out of nine textbooks with examples had at least one example each of vertical and horizontal bars, it was more common to see vertical bars than horizontal bars. Two had approximately equal numbers of examples with vertical and horizontal bars, one had a majority of horizontal bars, and the rest had a majority of examples with vertical bars. One textbook did not mention bar orientation and had only examples with vertical bars.

Two textbooks stated that bars could be placed in any order. There were no examples from textbooks in this study showing how the visual display of the data might be changed, by changing the order of bars, or whether this made a difference for some data. Categorical data typically had no obvious order for listing the categories (e.g., kinds of pets). However, examples using numerical data were usually arranged in some order (i.e., grade levels from youngest to oldest). Although there was a logical order, it

would not have changed the meaning of the data if the data elements (e.g., grade levels) had been put in a different or mixed up order. Infrequently, bars for categories were arranged in order of increasing or decreasing frequency, which facilitated comparisons of the data; however, ordering bars in that way could lead to inappropriately identifying trends across categories.

Content knowledge of bar graphs

The content of bar graphs involves recognizing the kind of data that is being represented as categorical or numerical. Additionally, it is important to note whether the data is appropriate for a frequency bar graph or a case value bar graph. Knowing the characteristics of the data is helpful for constructing an appropriate bar graph. Identifying the type of data that can be appropriately represented in a bar graph is necessary to decide or notice how the bars represent that data. Types of data were explicitly identified in definitions or descriptions of bar graphs by five textbooks: four mentioned only categorical data and one mentioned numerical data. The other five textbooks did not specify the type of data.

Eight out of nine textbooks had examples of both types of data, including two of the textbooks that specified categorical data and the textbook that identified numerical data. Four textbooks had a majority (64-100 percent) of examples using categorical data. Three others had more examples using numerical data than examples with categorical data. Examples of bar graphs representing numerical data included examples such as the number of books (frequency) read by each grade level (number of the grade level used as a label for a category) or number of children (frequency) who were a certain height (number indicated a measure). Sometimes examples representing numerical data grouped the data into discrete categories along a scaled axis such as 1-5 years, 6-10 years, etc. In some instances, examples of grouped numerical data included one value in adjacent groups such as 140-145 cm, 145-150 cm, etc. In these cases, it was not clear in which group certain data values (e.g., 145 cm) were represented. Grouping numerical data in this way for bar graphs was placed in some textbooks as a transition to a study of histograms.

Identifying the type of graph, frequency bar graph or case value bar graph, is also helpful for determining or noticing how the bars represent the data. Six out of ten textbooks explicitly mentioned the type of bar graph: four mentioned both frequency and case value bar graphs, and two mentioned only frequency bar graphs. Four textbooks did not specify or distinguish between different types of bar graphs.

Seven out of ten textbooks had examples of both types of bar graphs. Of the textbooks that had examples of both types of bar graphs, one had an equal number of examples for case value bar graphs and frequency bar graphs. The rest had a definite majority (56-86 percent) of examples showing frequency bar graphs. Two textbooks mentioned only frequency bar graphs (indicating "how many" of something); one included examples of case value bar graphs (showing "how much" of something) in addition to examples of frequency bar graphs, but without distinguishing between these types. The other textbook had no examples. Two textbooks had only examples of frequency bar graphs; one mentioned both types of bar graphs and the other did not specify the types of bar graphs. The textbooks that did not specify the different types of bar graphs and had examples of both types, did not distinguish between them.

Communicating the content and conventions of bar graphs

Seven features of bar graphs were examined: type of graph (frequency bar graph or case value bar graph); type of data (categorical or numerical); bar orientation (vertical and/or horizontal); bar length (frequency or data value); bar width (e.g., uniform width); bar order (e.g., arranged in any order); and bar spacing (equal spacing or no space between bars). None of the textbooks mentioned all seven features. Three mentioned five features, one mentioned four features, two mentioned three features, one mentioned two features, one mentioned one feature, one did not explicitly mention any of the seven features, and one did not have a definition of a bar graph. Eight of the nine textbooks that had definitions or descriptions each mentioned something specifically about the type of bar graph and/or the type of data. Four or fewer textbooks mentioned one or more of the other five features. The one that did not mention any specific features simply indicated that bars show the data without saying anything about how or what the bars show about the data.

The language for describing how the bars represented the data was very clear in definitions and descriptions of bar graphs in four textbooks. For example, one textbook described a case value bar graph where "each case is represented by a separate bar whose relative length corresponds to the magnitude or value of that case" (Lappan et al., 2006). Three other textbooks did not specify how the bars and/or length of bars represented the data and two others used ambiguous language, but all of these textbooks consistently showed examples that implicitly indicated the conventions for using bars to represent categories or cases and the lengths of the bars as representing frequencies or magnitude (i.e., case values). An example of ambiguous language included using a phrase (e.g., value of the data) to mean a

category or case represented by each bar, and then later in the description using the same phrase to refer to a frequency or case value as represented by the length of each bar. The same phrase was used with different meanings within the same description, so that difficulty might arise with sorting out which set of numbers in a numerical set of data is represented by the bars and which set of numbers represents the lengths of the bars. This confusion is less likely when using categorical data.

Implications and recommendations

It was evident that some textbook definitions, descriptions, and examples highlighted key features of bar graphs, while other definitions, descriptions, or examples reflected some but not all key features explicitly or implicitly.

Textbook support for teachers' knowledge of bar graphs

Mathematical knowledge for teaching data representation with bar graphs involves not only the ability to accurately construct such graphs, but involves identifying key features of these forms of data representation, distinguishing characteristics of data sets, and an awareness of the purposes and contexts for making sense of the data that are relevant and meaningful for students. Teachers should know that bar graphs sometimes represent relative magnitude (or case values) and sometimes represent frequencies, and be able to tell the difference between these types of bar graphs presented in textbooks. The amount of data reduction embedded with making certain types of bar graphs is an important factor to consider in choosing appropriate representations of data from textbooks for instruction. Introductory examples and textbooks from younger grade levels tended to have more examples with case value bar graphs (no data reduction) than textbooks at the later grade levels. Some textbooks used grouped numerical data represented in bar graphs as a way to transition from frequency bar graphs to histograms that defined intervals on a continuous scale.

Potential sources for confusion related to graph comprehension and conventions of constructing graphs included vague textbook definitions or descriptions, and examples that were not explicitly identified to distinguish between key features that mattered. Especially when using bar graphs to represent numerical data, teachers need to be clear about the characteristics of the data for distinguishing between numerical "values" that represent individual cases or label categories, and numerical "values" that indicate magnitude (case values or measures) or frequencies of categories. To consistently apply the convention that the length of a bar represents the magnitude or frequency of data, teachers must attend carefully to making this distinction. If each category or case had a unique frequency or magnitude, then the complexity for distinguishing between the two sets of numbers in a set of numerical data was increased. When more than one category or case had the same frequency or magnitude, it was easier to distinguish what aspect of the data was being compared, and thus represented by the lengths of bars in the bar graph.

Implications for supporting teachers' knowledge of the development of ideas related to representing data with bar graphs include: identifying textbooks that provide explicit attention to key features of bar graphs with clear language and supporting examples. Teachers of mathematics who understand the key features of bar graphs, how these key features relate to bar graph construction, and how the different kinds of bar graphs are similar and yet different may be better able to make instructional decisions that consider a sensible sequence for introducing and developing these ideas explicitly, especially when textbook information does not.

Using textbooks as resources for instruction

Hiebert (2008) defined teaching as "how the teacher and students interact about the content." Teaching data analysis requires teachers to know something about characteristics of data, conventions in constructing graphs for representing data, and be proficient in graph comprehension. Knowledge of bar graphs and data represented appropriately with bar graphs is reflected through studying curricula resources such as commonly used textbooks and teacher resource materials. "A good curriculum is important (teachers can't be expected to create their own curriculum); but a good curriculum is not enough" (Hiebert, 2008).

Decisions for when it makes sense to address case value plots, when and how to group data in categories that deal with data reduction, and general development of graph sense are important for teachers to consider when planning instruction (Friel et al., 2001). Beginning with categorical data (not numerical data) seems reasonable for introducing bar graphs. Grouping numerical data into discrete categories with bar graphs seems an intermediary step in moving to representing data in histograms.

An important implication of using textbooks as resources for instruction on representing data with bar graphs includes recognizing what the textbook highlights explicitly and what is implicit in examples. Teachers' knowledge of representing data with bar graphs has potential to influence instruc-

tional decisions based on supplementing or making explicit key features that are vague or implicit in textbook definitions, descriptions, or examples. In addition, teachers may need to clarify confusing definitions or descriptions, and emphasize characteristics of data sets to encourage students to consider how these influence constructing appropriate graphs. Considering ways to make key features of bar graphs explicit when textbook information does not, directly influences the learning opportunities for students.

Final thoughts

Why are some teachers of mathematics, who have studied mathematics for themselves as part of their teacher preparation, confused about representing data with bar graphs? Many teachers rely on textbook materials to plan mathematics instruction; therefore, trying to understand the source of teachers' confusions about these graphs naturally led to examining textbooks commonly used for teaching mathematics in the middle grades.

Definitions, descriptions, and examples of bar graphs were studied from a sample of commonly used textbooks for the purpose of sorting out key features of these graphs. Additionally, how these features were emphasized or obscured in the textbooks was considered. "Graph sense develops gradually as a result of one's creating graphs and using already designed graphs in a variety of problem contexts that require making sense of data" (Friel et al., 2001). This study focused on what teachers need to know about the characteristics of data sets and the conventions of bar graph construction presented in commonly used textbooks to develop their own and students' graph sense.

"Teaching is the one component of the system that teachers control ... unless we focus on teaching, directly, all the other changes we make won't matter to students" (Hiebert, 2008). Recommendations to teacher education include ensuring that teachers and prospective teachers understand the importance of data analysis in today's world and how the ability to interpret data representations is becoming a new basic skill. Typical representations of data using bar graphs are used frequently in all forms of media and information. Teachers need to have sufficient content knowledge to be able to supplement or clarify textbook material in order to provide learning opportunities for students that consider how the data is represented and to critically examine what visual images data representations convey about the data. Number and operations is a major emphasis in the elementary (primary) grades. However, the study of case value bar graphs may provide opportunities for younger children to work on number ideas and skills, in contexts that are interesting and relevant. In the middle grades, the study of both types of bar graphs allows for identifying characteristics of data to determine appropriate representations. The study of bar graphs for representing data has much to offer for mathematics education.

References

Ball, D. L., Hill, H. C., & Bass, H. (2005). Knowing mathematics for teaching: Who knows mathematics well enough to teach third grade, and how can we decide? *American Educator* (pp. 17). http://www-personal.umich.edu/~dball/articles/index.html.

Friel, S. N., Curcio, F. R., & Bright, G. W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education*, 32(2), 124-158.

Hiebert, J. (2008). Creating lasting improvements in classroom teaching. Presentation at the annual meeting of the National Council of Teachers of Mathematics, Salt Lake City, UT.

Lappan, G., Fey, J. T., Fitzgerald, W. M., Friel, S. N., & Phillips, E. D. (2006). *Teacher's guide, data distributions: Describing variability and comparing groups* (pp. 9). Boston, MA: Pearson Prentice Hall.

Ma, L. (1999). *Knowing and teaching elementary mathematics*. Mahwah, NJ: Lawrence Erlbaum Associates.

Remillard, J. T. (2005). Examining key concepts in research on teachers' use of mathematics curricula. *Review of Educational Research*, 75(2), 211-146.

Russell, S. J., Schifter, D., & Bastable, V. (2002). *Developing mathematical ideas: Statistics: Working with data casebook*. Parsippany, NJ: Dale Seymour Publications.

Shah, P., & Hoeffner, J. (2002). Review of graph comprehension research: Implications for instruction. *Educational Psychology Review*, 14(1), 47-69.

Stein, M. K., Smith, M. S., Henningsen, M. A., & Silver, E. A. (2000). *Implementing standards-based mathematics instruction: A casebook for professional development.* New York: Teachers College Press.

Tarr, J. E., Chavez, O., Reys, R. E., & Reys, B. J. (2006). From the written to the enacted curricula: The intermediary role of middle school mathematics teachers in shaping students' opportunity to learn. *School Science and Mathematics*, 106(4), 191-201.