The National ASA Poster Competition

By Linda Quinn

A statistical poster is a display containing two or more related graphics that summarize a set of data, look at the data from different points of view, and answer specific questions about the data.

John Tukey, in a 1990 *Statistical Science* article said, "Much of what we want to know about the world is naturally expressed as phenomena, as potentially interesting things that can be described in non-numerical words." We collect data to describe and answer questions about phenomena. We present data to communicate our ideas to others. The purpose of a statistical poster, then, is to tell a story visually from the data about some phenomena, revealing the conclusions that can be drawn.

A poster has one major disadvantage, however. Because there is no narrator to tell the story, nor an accompanying report to discuss the data, the poster must be able to stand on its own; it should not have to be explained. For this reason, special care must be taken to present ideas clearly. Not only must viewers understand the individual graphics, but they must also understand the relationships among the graphics and how the graphics address the question(s) being studied.

The American Statistical Association sponsors a competition open to students in four grade categories: K–3, 4–6, 7–9, and 10–12. The first year for competition was 1990. Additionally, a regional structure, similar to science fair judging, was established. This regional structure allows more students to be recognized for their efforts. Each region awards student posters, and then the top posters are advanced to the national competition.

Winning posters, suggestions for improving graphs, and registration information can be found at *www.amstat.org/education/posterprojects*.

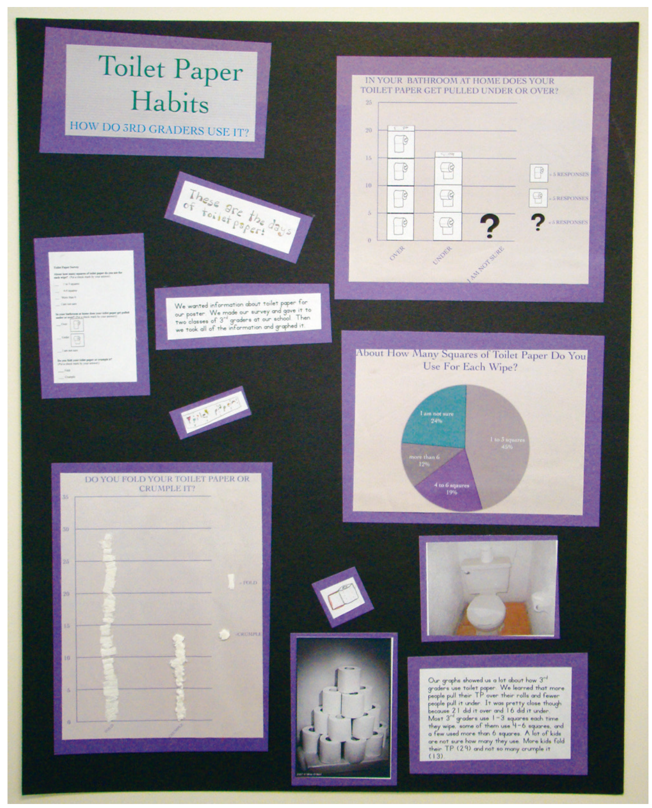
Why a Poster?

Statisticians are typically trained to graph data as the first step in a data analysis. Indeed, some would say the second step is to graph it again—in a different way. The National ASA Poster Competition is an expression of this philosophy. Sometimes, it is only in looking at data graphically that observations can be made. Graphs also complement more traditional statistical inference procedures. Posters reflect the **authors' view** of the data. Two posters using the exact same data could look very different.

Planning, Designing, and Constructing a Statistics Poster

The *Guidelines for Assessment and Instruction in Statistics Education (GAISE)* has a framework that can be aligned to the design, construction, and judging of an entry in the National ASA Poster Competition. The GAISE framework components are formulate a statistical question, design and implement a plan to collect data, analyze the data, and interpret the results in the context of the original question.

The first step in planning a statistical poster is emphasized in the first component of the GAISE framework; students should formulate a major question that can be addressed with data. Think of a major question you would like to answer, along with 3–5 sub-questions that will help answer the main question.

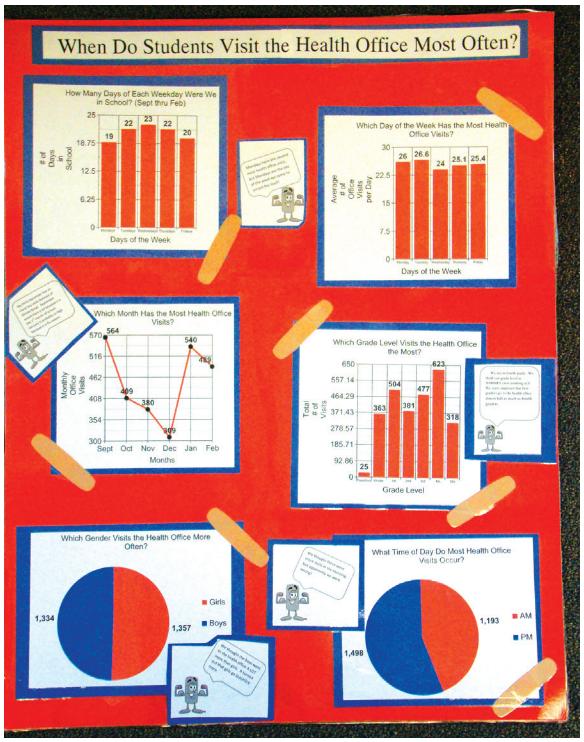


2011 first-place winner for the grades K–3 category

Remember that the central idea of the study should be the most prominent feature of the poster. This is almost always the most difficult part of entering the poster competition. So, brainstorm about things that interest your students, things they are learning about, and things they have read or seen. Statistics posters really are an interdisciplinary application. Try to encourage them to ask questions of importance to a wider audience. Avoid questions that are variations of "What's your favorite …?" and "What do you like?" You want your students to generate a question that can be answered with data, and there is variability in the data.

The second step involves the second element of the GAISE framework. Students should decide how to collect, organize, and display relevant data to answer the questions formulated. This really starts by deciding where to get the data. There are several possibilities that would result in a successful poster.

Surveys are a common source. For surveys, topics such as a person's favorite or what a person likes most are not distinguishing. Also, with surveys, students should conduct enough to make reasonable inferences. If you are collecting data from different groups of subjects, use about the same sample size, or report the data using percentages. Secondary data from the Internet or books are also possible to use. Just be careful not to simply reproduce graphs; that would be plagiarism. Students who design and collect data from an experiment are given a slight nod for creativity in data collection.



2011 first-place winner for the grades 4–6 category

Maybe just as important as where to get data is to define what type of data to collect and how that would lend itself to graphical presentation. For simplicity, consider a breakdown of either categorical or quantitative data. Is the data categorical, like favorite pizza topping, or quantitative, like how tall someone is? The type of data will dictate the type of graphs that can be constructed during the next step. The former lends itself to bar and pie charts, and the latter lends itself to histograms, dotplots, scatterplots, and boxplots. For categorical variables, try not to have too many or too few categories. Too many categories make for too many slices of a pie chart or too many bars. For quantitative data, make sure to remember the units on all graphs and in all legends.

In the third step—the third component of the GAISE framework—students will select and use appropriate statistical methods to analyze data. The analysis for the competition is mainly graphical. Graphs should be able to be read easily from six feet away and take up at least 75% of the poster space. Create appropriate graphs that show the many dimensions of the data. Usually, 3–5 graphs work best.

Each graph should display a new aspect, not be a different representation of the same data. That is, students should not simply graph the same data using a pie chart, bar chart, and pictogram. A bar chart and pie chart are both appropriate for categorical data. A pie chart is best used when the author wants the reader to compare each category as a portion of the total. A bar chart is best used when the author wants the reader to compare categories against each other. A pictogram is a special form of bar chart.

Finally, the last step—and last component of the GAISE framework—is to evaluate inferences and predictions based on data presented in the poster. Students should place themselves in an observer or reader's place. The graphs should tell the story. Students may want to annotate the poster with some of the conclusions drawn, but the answer to the questions should be clear from the graphs alone. If the judges cannot interpret it, the poster may not have told a clear story. Most of the exemplary entries in the competition pose a research question as the title or the title tells the conclusion of the research.

A **statistical poster** is a display containing two or more related graphics that summarize a set of data, look at the data from different points of view, and answer specific questions about the data.

What Makes a Winning Poster?

Judging the posters is based on the following five criteria:

1. Clarity of message
2. Appropriateness of the graphics
3. Details of the graphs
4. Creativity
5. Overall impact

Following are comments from judges and suggestions for creating a winning poster:

The **clarity of the message** is ensuring that the overall message is clear. One way for students to see opportunities for improvements here is in a peer review. Ask students to constructively critique each other's posters. What do they think the message is, and is that the story the author was trying to tell? Finally, are the conclusions obvious from the graphs, or do you want more information?

> *Dimensionality of the question* – A good poster addresses multiple dimensions of the main title or main title question. While some entries have used multiple graph types (such as bar and pie), this is not sufficient. One improvement would be to graph the overall data and then graph the data broken down by subcategories in other graphs. This is acceptable; however, in the higher grades, the graphs should be based on different questions that are all related to the main topic and not merely a breakdown by categories. This will help the reader understand the multiple facets of the question.

Considerations concerning the **appropriateness of the graphics** are about choosing the appropriate graph and doing it well. Here are some considerations for specific graph types.

Pie charts – Be careful using pie charts when the number of categories is large. The pie segments become difficult to distinguish and more difficult to interpret. When possible, consider placing the labels around the pie. It allows the viewer to make conclusions with one less step (going to the legend to figure out what category is represented by which pie segment). This also helps when the colors or patterns are really close and difficult to distinguish on the legend. If you want to make comparisons between several pies, the segments should be ordered in an identical fashion and start at a similar angle. Sometimes, a segment has zero frequency or percentage and does not show up in a pie; this can be an important fact.

Line plots – Line plots, where data points are connected with a line, are usually appropriate when the horizontal axis is ordinal or quantitative, like time. When the horizontal axis has nominal data, consider the impression the graph would have if the categories were re-ordered. The incorrect use of this graph with a categorical axis usually eliminates the poster from award consideration.

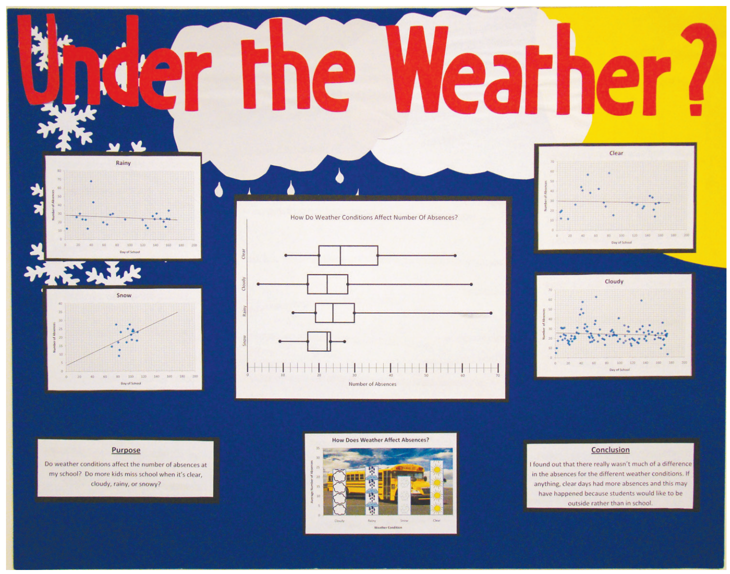
Bar charts – While bar charts can be used to show averages, you are really only conveying one number for each group. If the raw data are available, a far richer presentation in the form of boxplots or dotplots should be considered. These types of presentations let the reader see not only a measure of center, but spread and shape as well.

Stacked or cluster bar charts – Stacked or cluster bar charts are used when trying to capture the relationship between two categorical variables. A cluster bar chart is also considered a side-by-side bar chart, and a stacked bar chart would take those side-by-side bars and stack them into one bar with segments (i.e., segmented bar chart). Either graph could work. Try looking at the graph both ways. Pick the way that is clearest in its presentation. Consider the number of categories. Consider if you are using frequency counts or percents (and which kind of percents). If sample sizes are different, this is an important consideration.

Boxplots – Multiple boxplots on the same graph, or with the same scale, can be used to effectively make comparisons between groups based on center, spread, and shape. It should be remembered that a boxplot is a five-number summary and should not be used when the sample sizes are small (a rule of thumb might be fewer than 20). Also, it would be helpful to annotate the sample size on the boxplot graph. Boxplots that are meant to be compared, but are in different graphs, use different axis scales, or are drawn in differing sizes, are difficult to compare.

Scatterplots – Scatterplots do not need to have a best fit line on them, but they could. Scatterplots should be designed so the independent variable is on the horizontal axis. Sometimes, scatterplots are underused in the early grades. If the point of your questions is to show a relationship between two quantitative variables, use a scatterplot. For example, if you wanted to compare the length of your name to the value of your name in Scrabble, a scatterplot is more useful that two dotplots showing each variable individually, because now you can see the association.

Pictograms – A pictogram is a graph used often in the elementary grades. Make sure the graph is aligned to start at the same spot and each symbol is the same size. Consider putting a bar outline over your pictogram; does it make a reasonable bar chart? If not, something needs to be modified. Judges are looking for the connections between the graphs and the **details of the graphs** that enhance understanding.



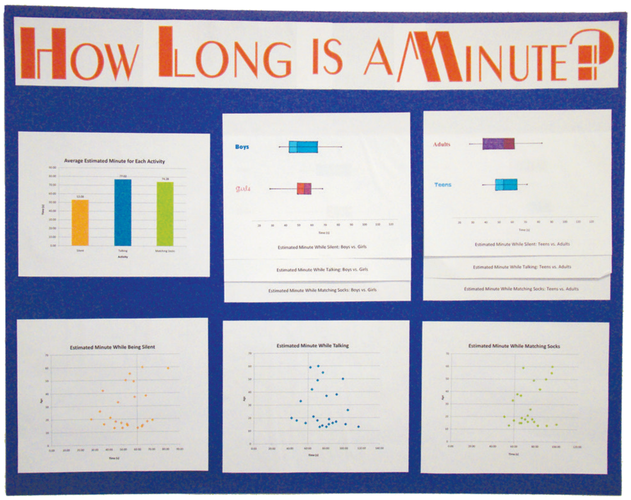
2011 first-place winner for the grades 7–9 category

Frequency versus percentage – When sample sizes are not equal and the author is trying to draw comparisons, it is better to report percentages. Percentages may not be appropriate for some K–3 students; therefore, it is important to try to make the comparison groups as equal in sample size as possible.

Axes and labeling – To ensure accurate interpretation, graphs should be adequately labeled, including labels for all axes. These labels should be spelled correctly and be large enough to read. Most times, it is preferable for quantitative axes to begin at zero. When that cannot be the case because the scale of the data is too prohibitive, there should at least be an axis break symbol used.

Grouped data – On occasion, the totals of subgroups have not made sense when compared to the total group. For example, if there were 30 males and 20 females, shouldn't the total be 50? If the grade 2 level mean is 3.4 and the grade 3 level mean is 3.6, how can the total mean be 3.9? Sometimes, this can be caught by simply proofreading the poster. If there is a reason this occurs and it wouldn't be obvious to the reader, it should be better presented.

Adjusting rates by population – Some data found in other resources or on the web, particularly geographic data, should be adjusted by the population size to make fair comparisons. If the U.S. population is larger than Cuba's, most of the statistics in terms of poverty, deaths, marriages, etc., will be, too.



2011 first-place winner for the grades 10–12 category

The use of color – Posters do not have to use color, although color may add to the overall impact. Color is a powerful way to help graphs make conclusions more readily seen. If color is an actual response, use that as the color in the graph. A graph was submitted that described the eye color preferred in a mate by the respondent. The students chose to use the color green to represent blue eyes, yellow for green eyes, and red to represent brown eyes. Imagine the judges' first impression of this data—"Red and yellow eyes!" Using a natural color, say blue for boys and pink for girls, helps the viewer quickly see relationships. Another example for choice of colors might be stop light colors—red for no, yellow for maybe, and green for yes. When possible, identical colors should be used to identify identical categories. Colors can enhance recognition of the conclusions by tying graphs with the same groups together. For example, use blue for boys in all three graphs that are showing boys.

Categorical data – When making use of the same categories across multiple graphs, show the categories in the same order if the reader is supposed to make a comparison between graphs.

Creativity is based on the creativity of the topic, creativity in the data collection, and creativity in the presentation.

Considerations concerning the **overall impact** relate to the entire poster. A winning poster, if viewed as an entire product, is readable from six feet away, is neat, has proper spelling and grammar, and tells a compelling story with four to six graphs (fewer in the K–3 category).

Use of 3-D graphics – 3-D graphics have proliferated with the use of computer-generated graphs and their use in the popular media. However, they can be deceiving to read and interpret. Where does the bar end? Is the pie piece bigger because we are comparing volume? Three dimensions should be used only when the third dimension means something other than aesthetics.

Computer versus hand-drawn – Both types of graphs are acceptable and have won. The tradeoff seems to be readability versus neatness. Computer-generated graphs may be neater and more accurate, but they may use smaller fonts that sacrifice readability. Hand-drawn graphs allow for larger fonts, but may require more time to produce in order to be neat and accurate (straight lines, coloring within bars, circular pie charts). When generating computer graphs, students should take time to make colors and scales consistent between graphs and add to the interpretation. Increase the font size of labels, axes, and titles because posters should be readable from six feet away.

Use of space – The purpose of the competition is to tell a story with graphs. Therefore, the graphs should take up the majority of the poster. Titles can be readable without taking up a third of the poster. Graphs should take up at least 75%.

Chart junk – Edward Tufte used the term "chart junk" in *The Visual Display of Quantitative Information* and defined it as the extra graphics added to a chart or graph that add no value and distract viewers with information that isn't vital to communicate. Chart junk happens when the creativity and aesthetics of the poster become more important than the information the graph is meant to convey. There are annotations to graphs that help readers understand the data better—reference lines, for example. However, pictures used to decorate, glitter that distracts as to whether it is a point or decoration, and pictures in the background of graphs that make the graph harder to read, should all be removed.

Concluding Remarks

Certainly, students can have fun working individually or using teamwork skills. Every poster is unique. There is always an animated discussion among judges since, like works of art, each judge can see different aspects. Rarely are poster winners decided by unanimous vote, but rather by discussion of the good and poor in each presentation. Then, a ranking system is used to determine the winners.

The poster competition is one way to get students actively engaged in collecting data and drawing inferences, both vital steps in critical thinking. The poster competition also integrates many subjects. While mathematics may be the most prominent, the topic can come from any subject. The same principles that guide improved writing—write, revise, and re-write—also apply to good poster products. Artistic skills come into play in knowing the line between good aesthetic layout and overkill with decorations. Overall, many varied skills are required to tell a great story with a poster.

Resources

Bereska, C., L. C. Bolster, C. A. Bolster, and R. Scheaffer. 1998. *Exploring statistics in the elementary grades: Book one, grades k–6*. White Plains, NY: Dale Seymour.

Bereska, C., L. C. Bolster, C. A. Bolster, and R. Scheaffer. 1998. *Exploring statistics in the elementary grades: Book two, grades 4–8*. White Plains, NY: Dale Seymour.

Franklin, C., G. Kader, D. Mewborn, J. Moreno, R. Peck, M. Perry, and R. Scheaffer. 2007. *Guidelines for assessment and instruction in statistics education (GAISE) report: A pre-k–12 curriculum framework.* Alexandria, VA: American Statistical Association. *www.amstat.org/education/gaise*.

National ASA poster and project competitions rules, entry form, and previous winners: *www.amstat.org/education/posterprojects/index.cfm.*

Webinar: Working with K–12 Students to Create a Statistical Poster: *www.amstat.org/education/webinars/index.cfm*.

National Council of Teachers of Mathematics publications: *www.nctm.org*

- *Navigating through Data Analysis and Probability in Pre-K–2*
- *Navigating through Data Analysis and Probability in Grades 3–5*
- *Navigating through Data Analysis Grades 6–8*

Used Numbers series (primarily data analysis): *www.pearsonschool.com/index.cfm*

STEW, Statistics Education Web, ASA peer-reviewed lesson plans: *www.amstat.org/education/stew/index.cfm.*

Note: This is from *Bridging the Gap Between Common Core State Standards and Teaching Statistics*, pp. 251-260, *ww2.amstat.org/education/btg*.