Overview of Research Basis for *TinkerPlots® Dynamic Data™ Exploration* Software

Research

During the last ten years, statistics educators have learned quite a bit about how younger students learn to reason about data, and how experts approach data differently from ordinary people or novices. This research influenced both the design of *TinkerPlots® Dynamic Data™ Exploration* software and the type of data and activities included with it.

In this document below, we briefly describe some of what we have learned from research and how we've applied it to *TinkerPlots®*. For a more detailed description of this research, see Chapter 13 of ***A Research Companion to Principles and Standards for School Mathematics*** (2003), published by the National Council of Teachers of Mathematics (NCTM).

Developing an Aggregate View of Data

Experts look at a group of data as a collection, or aggregate. When they look at a histogram or stacked dot plot, for example, they tend to see its general shape, how spread out it is, and where the data are centered. When experts compare two groups (say two different stacked dot plots showing people's heights) they generally base those comparisons on these aggregate features. For example, an expert looking at two distributions of height for men and women might conclude that men tend to be taller than women by comparing the locations of the centers (for example, the means) of each distribution.

Students new to the study of data tend not to see these aggregate features. Rather they see individual cases or subgroups of individuals with the same, or nearly the same, value. Looking at the same distribution of height, students might focus mostly on individuals on the extreme ends of the distribution. Alternatively, they might tend to see short people, average people, and tall people, and reason about these as three different groupings. When students compare two groups, they often compare the number of cases at one of the extreme ends of the distributions. For example, a student looking at a graph showing men's and women's heights might argue that the men are taller because there are more of them above 6 feet. One problem with this comparison is that it doesn't take into account the number of cases in each group. If we had a group of 1,000 women and 10 men, we would probably have more women than men above 6 feet. This method of comparing numbers in common slices of the distributions often leads to faulty conclusions.

Researchers have also found that it usually isn't fruitful to simply tell students that what they are doing is wrong and to recommend an alternative. If students don't understand why their methods are flawed and why the alternative is better, they will do what you ask now, but later revert to what makes more sense to them. This research led us to take a different approach in *TinkerPlots*. **We have included tools**

students can use to look at and compare graphs using their intuitive methods. For example, students can use dividers and reference lines to break up distributions into parts and click a button to quickly get counts of cases in different parts. We believe that by allowing them to use the methods that make sense to them, and discussing what they do with other students and with teachers, **they will come to see the limitations in their approaches and develop more powerful methods.**

Describing Centers of Distributions

Rather than using means and medians to summarize numeric data, researchers have reported that most people, including young students, prefer to describe a distribution using a center clump: for example, “Most of the data are between 15 and 25." **Dividers and hat plots are *TinkerPlots* tools designed specifically to help students describe data in terms of these informal averages.** Students can also use the Drawing tool to show where they think the bulk of the data are. It is fine for students to summarize groups and to describe differences by saying that one group's center clump is higher up than another group's. We expect that formal averages, like means and medians, will eventually make more sense when students see that these are almost always located within what they see as the center clump.

Staying Grounded in the Data

Many researchers have reported the importance of students keeping in mind what a case is—that a data point on a graph is not simply a number or dot, but that it is a measurement for a case (for example, a person's height). Often when students (or experts) get confused looking at a graph, it is because they lose this connection between a case and a point on the graph. In *TinkerPlots*, students can, at any time, click a data point on a plot to see the data card for that individual case and to automatically highlight that case in any other plots that may be open. Again, this helps students keep in mind what a case is. It also helps them analyze their data, because when students highlight a case that has a high value for one attribute, they can see where that case is located in another graph of a different attribute. For these reasons, in all *TinkerPlots* graphs—even histograms, box plots, and pie graphs—**it is easy for students to see individual cases and how those cases work together to form a particular graph.** When one plot changes to another, *TinkerPlots* animates the transformation to help preserve the integrity of the case.