Statistical graphics are important during every stage of the statistical process: exploratory analysis, examining model assumptions, and presenting statistical results. Graphics summarize the data visually, freeing cognitive resources to consider the implications of the data. As cognitive aids, it is important that graphics are carefully designed to display important features of the data in a way that facilitates visual inference. I study this interplay between statistical graphics and the brain, seeking to understand how the brain processes statistical graphics and applying that knowledge to create intuitive and effective graphs.

Perception of Statistical Graphics My dissertation examines the impact of several facets of human perception of statistical graphics: optical illusions, visual aptitude, and the hierarchy of features in graph perception, arguing that it is important to consider human perception in the design and evaluation of statistical graphics.

In the first chapter, I explore the sine illusion, which affects the graphical perception of data along nonlinear trend lines, in two separate user studies. The sine illusion research highlights a neglected issue in statistical graphics: how do we ensure that information is perceived by the brain as it is presented graphically? This work has been well received - the first study has been submitted to and accepted by JCGS; the second study was selected for the ASA graphics section student paper award at JSM in 2014, and an expanded paper is in preparation for publication. The papers are largely designed to raise awareness of the illusion and its effect on a wide array of oft-used graphics, such as scatter plots, time-series charts, hammock plots, stream graphs, ribbon charts, and candlestick plots.

The remaining chapters are more focused on visual inference, exploring the use of statistical lineups (sets of 20 plots: one containing real data and 19 generated by permutation) to conduct formal *graphical inference*. In the second chapter, I explore the interaction between tests of spatial ability and visual inference using lineups; the results suggest lineups are ultimately a classification task, though performance is also a function of the type of chart, overall visual ability, and demographic factors. An article exploring the results from this study is currently in preparation for publication. In the final chapter, I examine the hierarchy of visual features in statistical graphics in the context of the lineup protocol; that is, which graphical features are most visually salient. While this experiment is still in the design stage, I expect the results to reinforce the importance of carefully highlighting notable features of the data using aesthetics such as color, shape, and size.

Interdisciplinary Collaboration In addition to my research on statistical graphics, I also collaborate with researchers in other disciplines to create effective graphics for large data sets or multidimensional data. Recently, I have worked with soybean researchers at the USDA to analyze and visualize soybean genetic data at the population and individual level; the result is a series of topic-specific interactive applets which integrate the data and analysis for use by biologists. Working with co-authors, I have also created interactive applets for use in the statistics classroom; these applets were designed to intuitively demonstrate difficult statistical concepts. During my masters research, I worked with engineers to model and identify peaks in mass spectroscopy data using Bayesian and nonparametric statistical techniques.

Future Research As data sets grow larger, interactive visualizations become more critical because user interaction can be used to display subsets of larger data, adding complexity and utility. I would like to investigate the relationship between the user experience and different types of interactivity, utilizing tools such as eye and cursor tracking, recording user interactions, and analyzing participant descriptions of the charts and data; these tools provide insight into both the perceptual process and the information communicated by the visualization. This research would build on existing literature examining static graphics, but would also incorporate sensory integration research, as the motion used in interactive graphics triggers specific perceptual organization schemes. In parallel with this research, I would also like to explore methods for reducing the complexity of graphics; this research is more directly applicable to visualization in the field, as current interactive plots typically rely on JavaScript and do not scale well to “big data".

Statisticians can assemble huge databases of information, but in order to communicate findings (particularly outside of the field), one of our most effective tools is still a well-designed chart. Understanding the interaction between human perception and statistical graphics provides an essential foundation for optimizing graphics to more effectively communicate with each other and with those outside of the field.

Statistics courses often make a bad first impression: students walk away from introductory classes with the idea that statistics is hard, extremely theoretical, or not particularly relevant to everyday life (outside of election season polls and choosing colored balls from a box). The rise of "big data" and "data science" have created a climate where statistics is vital to many different areas of business, government, and science, but only if it masquerades as something "cool". It is important to counter this trend by making statistics accessible, fun, and relevant to students.

Course Structure In my experience, the best courses set students up for success with clear objectives, well-organized reference materials, and numerous sample problems. Ideally, the textbook should complement the lectures; in particular, the lectures and the textbook should provide different approaches to the material, so that students who do not understand one explanation have alternatives which may be more suited to their learning style. Lecture notes, outlines, and code allow students to prepare for class ahead of time, so that lectures can focus on assessing and reinforcing students' comprehension. For each topic, the lectures and examples should begin with a basic overview, then provide details that facilitate a nuanced understanding, encouraging exploration of open-ended problems.

Feedback At every stage of the learning process, mutual feedback is important. Feedback from students should shape the course structure and presentation, so that lectures and written materials help as many students as possible; feedback to students should clarify misconceptions, identify problems, and direct students to additional resources. Instructors should also be prepared to assist students with situations that may not be directly related to the course material: disabilities, medical problems, or personal issues may affect student performance in class and their ability to engage with the material; accommodating these students positively impacts the learning environment.

Course Design Statistics courses are typically designed for a specific audience; introductory classes may be targeted toward students in engineering, business, or scientific disciplines, while more advanced courses may be designed for students with a background in statistics. Introductory classes tend to focus on literacy (understanding analyses) while encouraging students to develop competency (the ability to design and interpret their own analyses); students in these classes do not have time to develop fluency (the ability to solve a problem, explain, and justify the solution), while advanced classes usually encourage students to fluency as well.

Literacy is a prerequisite for statistical competency and fluency; literate students can read and assess statistical analyses and conclusions. For students in introductory courses, statistical literacy is often the most important goal: students need to be able to think critically about statistical claims. In computational courses, literate students can understand well-structured code and make simple modifications. Breaking lectures up with demonstrations, worked examples, and group work reinforces a literate approach to the material, and short assessments (true/false, multiple choice, or short answer questions) provide mutual feedback.

Competency, the ability to correctly execute and interpret a statistical analysis, requires a more thorough understanding of the material. Students must engage the topic abstractly and may need to understand some theory; this is often where students with sparse math backgrounds become hopelessly confused. In my experience, group discussions, hands-on problems, and individual exploration are valuable tools to facilitate competency. Outrageous and fun examples may also motivate students to attempt problems that would otherwise seem too difficult. In computational courses, competent students can write their own code and solve new problems using an established set of tools. Open-ended test and homework problems can be used for assessment and feedback.

Fluency, the ability to apply course material to novel problems independently, requires time and exposure to a wide variety of problems. Open ended questions, discussions, and projects encourage students to develop an understanding of the material and to think critically about the subject. The ultimate goal for most teachers at the end of a course is that students can be trusted to use their knowledge in the outside world: they can discuss a problem coherently, apply "textbook" knowledge appropriately, communicate the logic behind their approach, and interpret the results correctly.

Courses and learning environments which are well-designed, engaging, and responsive encourage development of a more nuanced understanding of the subject matter, whether the goal is literacy, competency, or fluency. As a student, I have experienced courses which exhibited all of these traits; as a teacher, I work to engage students, provide frequent, mutual feedback, and illustrate the subject matter with fun, engaging, memorable, and relevant examples.