Snap-Together Visualization: usability evaluation

Goal
- To evaluate the usability and benefit of the Snap system itself and discover potential user interface improvements

Participants
- 3 data analysts—familiar with data and analysis as they were employees of the US Bureau of the Census and the study used census data
- 3 programmers—1 from the Census, and 2 CS students on campus
- Domain experts vs. novices? Part of the design?

Evaluation is required at all stages in system development
1. Initial assessments
2. Iterative design process
3. Bench-marking
4. Deployment
   - How is the system used in the wild?
   - Are people using it?
   - How does the system fit in with existing work flow? Environment?
   - Contextual studies, field studies...

Timing Results
To evaluate the usability and benefit of the Snap system itself and discover cognitive trouble spots in training and test trials, and Snap user-interface design:

- 3 exercises to construct a coordinated-visualization user interface according to a provided specification
- Exercises designed to test different aspects of the system to uncover usability issues
- First 3 exercises were interface construction according to specific criteria:
- Exercise 2 was more open-ended than required (subject thinking about"
- Did not say how these tasks were chosen. For example, is "one-to-many" join relationship (Exercise 2) suspected to be difficult prior to the study?

Results:
- Reported differences between analysts and programmers
- Analysts considered interface building to be easier than programmers as construction
- Analysts performed better
- Would more varied is to identify individuals in their report
- For example, did the Access/SQL experience of the Census programmer made a difference?

Success:
- Also hard to interpret as did not report in what form and how often the help was provided

Advanced

Interface Design and Evaluation
Evaluation is required at all stages in system development
1. Initial assessments
2. Iterative design process
3. Bench-marking
4. Deployment
   - Identify problems and go back to 1, 2, 3, or 4

Your Questions: about the snap idea

Your Questions: Usability Evaluation

Your Questions: Usability Evaluation

Your Questions: Usability Evaluation

Your Questions: Usability Evaluation

Readings
- The Perceptual Evaluation of Visualization Techniques and Systems. Ware, Appendix C.

Further Readings
- Task-Centered User Interface Design. Chapters D-S. Lewis, C. and Rummel, J.

Snap-Together Visualization: usability evaluation

Task
- 3 exercises to construct a coordinated-visualization user interface according to a provided specification
- Exercises designed to test different aspects of the system to uncover usability issues
- First 3 exercises were interface construction according to specific criteria:
- Exercise 2 was more open-ended than required (subject thinking about"
- Did not say how these tasks were chosen. For example, is "one-to-many" join relationship (Exercise 2) suspected to be difficult prior to the study?

One thing that struck me about this paper is that it appears to give very little credit to intuitive user interaction in system design. Firstly, the paper gives off the impression that users are "constructing" and "building" visualizations from scratch when in fact much of what can be done originally was done in the system itself. This is an interesting point for there is the sense that they are more likely to notice intuitive differences in different pre-defined visualization types. Additionally, many of these specific types are rather broad and could potentially fail in the wild.

What is the extra work users have to do in setting up links etc. does not appear to gain much more benefit than a custom-specific pre-constructed visualization system that affects the user interface of the system. So, for instance, if there was enough about the visualization that was available, a user could interact with the visualization enough, then they already knew what they were looking for.

In the 'snap' system, users need to drag and drop snap buttons to another window to construct visualizations and it is within the wild if the system for making interactive visualizations. The click-and-drag system requires users to predict what the visualization will look like when they drag and drop objects. There is no control over the final result.

Understanding when visualizing multidimensional data could appear to the difficult task since every visualization dimension is represented (certain techniques highlighting which could cause loss of the main visualization data). How to address this?

There is a similar system that can appear in the meaning of selecting is case of the list of items visualization and, in general, a visual representation of interactions. When we are selecting the list, should we just, visually, change the color or similarity changing colors that is not the main visualization. People say that windows are located at the top right of the main visualization because of the size of the edges of the selection but if we changed the first case then we easily view look of the product when we are located in the data globally.
The Perceptual Evaluation of Visualization Techniques and Systems

- Structural Analysis
  - Requirement analysis, task analysis
  - Structured interviews
    - Can be used almost anywhere, for open-ended questions and answers
  - Rating scales
    - Commonly used to solicit subjective feedback
      - E.g., 7-point Likert scale to assess mental workload
      - E.g., "It is frustrating to use the interface" (Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree)

- Comparative user study: Hypothesis testing
  - a practical poster presentation
    - E.g., In Snap: Participants will be faster with an interactiveomographic display than with an interactive display with no interactivity.
    - The last means making the above statement objective/subjective.
  - Factors
    - Independent variable:
      - E.g., Interface, task, participant demographics...
    - Levels
      - The number of variables in each factor
      - Limited by the length of the study and the number of participants

- Your Questions: Evaluation in practice
  - How much work is in visualizations is actually inspired by psychophysics and take into user studies of what kind we might do a good job of?
  - More tasks involved in understanding views within the research field. Is there similar focus on evaluation techniques of visualization/visualization in the standard software engineering principles apply?
  - There is nothing about the "The Perceptual Evaluation of Visualization Techniques and Systems" that puts too expensive to conduct, since this are user generated. Given that we have created a methodology to determine what is the standard," or "the interface that the standard works fine without.
  - More work can be done to either produce or indicate the standardization process, it is not uncommon to determine the standard works fine without.

- Low-Level Components of Analytic Activity in Information Visualization
  - How to select tasks for a user study?
    - Generally, use tasks that the interface is designed for
      - Use directly in the design is successful over competitor
      - But, tasks for researchers to see if the new visualization technique is useful elsewhere
      - Need a standardized task metric for comparison
    - While tasks are atomic and simple, may not reflect real-world tasks

- Low-Level Components of Analytic Activity in Information Visualization
  - Identified 10 low-level analytic tasks that largely capture people’s activities while employing information visualization tools for understanding data
    - Retrieve value
    - Filter
    - Compute derived value
    - Find extremes
    - Sort
    - Determine range
    - Distributions
    - Find anomalies
    - Cluster
    - Correlate

- Low-Level Components of Analytic Activity in Information Visualization
  - We could study tasks based on these operations
    - E.g., find extremes: Find data cases possessing an error that is far away from the rest of the data
      - In the scenario of monitoring and managing electric power in a control room...
    - Which location has the highest power surge for the given time period (extreme y-dimension)
      - A fault occurred at the beginning of this recording, and resulted in a temporary power surge. Which location is affected the earliest (extreme y-dimension)
      - Which location has the most number of power surged (extreme count)
In Summary: Two evaluation techniques
Two evaluation techniques are the fundamental components of user evaluation: (1) statistical analysis and (2) prototype usability testing. The former will provide a deeper understanding of the specific factors in the data. The latter can be used as a preliminary step to determine whether to perform the statistical analysis, although in practice it is challenging. This approach provides a systematic approach for using evaluation techniques.

The "User-Centered Components of Analytic Activity in Information Visualization" refers to the design and evaluation of information visualization. This paper discusses the evaluation of information visualization systems, including the visual and cognitive processes of users, the identification of effective visualization techniques, and the effectiveness of these techniques. The paper emphasizes the importance of evaluating information visualization systems in a user-centered manner and provides guidelines for evaluating these systems.

Usability testing
Usability testing is a critical component of user interface design. It helps designers to identify and correct usability problems early in the design process. The paper discusses the importance of usability testing and provides guidelines for conducting usability testing. The paper emphasizes the importance of involving users in the design process and using usability testing to guide design decisions.

Taxonomy use
The paper discusses the use of taxonomies in user interface design. It provides guidelines for using taxonomies to improve the usability of information visualization systems. The paper emphasizes the importance of using taxonomies that are relevant to the users and that are easy to use. The paper also discusses the importance of evaluating the effectiveness of taxonomies and provides guidelines for evaluating their effectiveness.

Questions: trivial hypothesis? Valid comparisons?
It appears that they use two-way ANOVA’s and multiple t-tests to compare pairs of conditions and states that the performance advantage was revealed by these individual tests. It is not uncommon, along the many individual tests in a task practice, as such, that researchers recommend that at least one out of ten to be a true positive. New important not statistical significance means, in general, Snapp together makes us think of all the test cards, it is necessary to test the significance of results and not report the significance level, or at least significance levels not often done.

Questions: statistical analysis
We question questions such as, ‘Which visuals did you like best?’ Instead, we left behind the ordinal test for each pair of visuals and instead used a binomial test to determine whether the cost of SNAPP on a pair of visuals was above the median cost of SNAPP on the set of visuals. We also asked participants to rate their confidence in their judgments of the cost of SNAPP.

Questions: user behaviours
I find it very strange that the authors were surprised when users found a visual prototype of a task that was not expected. I was skeptical as well, but they also used a well-known method. The idea is that the task of a visual prototype is to be highly effective, and that the actual task is to be less effective. This prototype was used as an exploratory tool in the visualization system. We will see if the visual prototype is effective (Is your novel technique actually useful?)

Questions: taxonomy use
The translated analytic types are the fundamental components of user evaluation: (1) statistical analysis and (2) prototype usability testing. The former will provide a deeper understanding of the specific factors in the data. The latter can be used as a preliminary step to determine whether to perform the statistical analysis, although in practice it is challenging. This approach provides a systematic approach for using evaluation techniques.

Questions: reproducibility
The authors introduce the study by giving the example that brightness generated from tasks (plotting numerical analysis) generally is not a problem of users, and this is one of the reasons why the results might suggest that the experts in the given dataset just didn’t need the task.

Questions: scientific community
It appears that they use two-way ANOVA’s and multiple t-tests to compare pairs of conditions and states that the performance advantage was revealed by these individual tests. It is not uncommon, along the many individual tests in a task practice, as such, that researchers recommend that at least one out of ten to be a true positive. New important not statistical significance means, in general, Snapp together makes us think of all the test cards, it is necessary to test the significance of results and not report the significance level, or at least significance levels not often done.

Questions: developers
We question questions such as, ‘Which visuals did you like best?’ Instead, we left behind the ordinal test for each pair of visuals and instead used a binomial test to determine whether the cost of SNAPP on a pair of visuals was above the median cost of SNAPP on the set of visuals. We also asked participants to rate their confidence in their judgments of the cost of SNAPP.

Questions: bottom-line
The paper discusses the use of taxonomies in user interface design. It provides guidelines for using taxonomies to improve the usability of information visualization systems. The paper emphasizes the importance of using taxonomies that are relevant to the users and that are easy to use. The paper also discusses the importance of evaluating the effectiveness of taxonomies and provides guidelines for evaluating their effectiveness.

Questions: user study
Your Questions: trivial hypothesis? Valid comparisons?
It appears that they use two-way ANOVA’s and multiple t-tests to compare pairs of conditions and states that the performance advantage was revealed by these individual tests. It is not uncommon, along the many individual tests in a task practice, as such, that researchers recommend that at least one out of ten to be a true positive. New important not statistical significance means, in general, Snapp together makes us think of all the test cards, it is necessary to test the significance of results and not report the significance level, or at least significance levels not often done.

Questions: time result analysis
The translated analytic types are the fundamental components of user evaluation: (1) statistical analysis and (2) prototype usability testing. The former will provide a deeper understanding of the specific factors in the data. The latter can be used as a preliminary step to determine whether to perform the statistical analysis, although in practice it is challenging. This approach provides a systematic approach for using evaluation techniques.

Questions: metrics
Please create a user-interface that will support users in efficiently completing their tasks. The users have different preferences for how they want to see the data and how they want to interact with the system. The user interface should be designed to support these preferences and should be easy to use.

Questions: bottom-line
The authors introduce the study by giving the example that brightness generated from tasks (plotting numerical analysis) generally is not a problem of users, and this is one of the reasons why the results might suggest that the experts in the given dataset just didn’t need the task.

Questions: data analysis
This is further reflected in the fact that they gave the users the visualizations to use before they took the test, that they ask the users ‘what’ this would like to test for (the data), that prior testing their visualizations generated from tasks (plotting numerical analysis) generally is not a problem of users, and this is one of the reasons why the results might suggest that the experts in the given dataset just didn’t need.

Questions: best practices
Time Result analysis: Hypothesis testing with ANOVA
I find it very strange that the authors were surprised when users found a visual prototype of a task that was not expected. I was skeptical as well, but they also used a well-known method. The idea is that the task of a visual prototype is to be highly effective, and that the actual task is to be less effective. This prototype was used as an exploratory tool in the visualization system. We will see if the visual prototype is effective.

Questions: best practices
Time Result analysis: Hypothesis testing with ANOVA
I find it very strange that the authors were surprised when users found a visual prototype of a task that was not expected. I was skeptical as well, but they also used a well-known method. The idea is that the task of a visual prototype is to be highly effective, and that the actual task is to be less effective. This prototype was used as an exploratory tool in the visualization system. We will see if the visual prototype is effective.

Questions: best practices
Time Result analysis: Hypothesis testing with ANOVA
I find it very strange that the authors were surprised when users found a visual prototype of a task that was not expected. I was skeptical as well, but they also used a well-known method. The idea is that the task of a visual prototype is to be highly effective, and that the actual task is to be less effective. This prototype was used as an exploratory tool in the visualization system. We will see if the visual prototype is effective.

Questions: best practices
Time Result analysis: Hypothesis testing with ANOVA
I find it very strange that the authors were surprised when users found a visual prototype of a task that was not expected. I was skeptical as well, but they also used a well-known method. The idea is that the task of a visual prototype is to be highly effective, and that the actual task is to be less effective. This prototype was used as an exploratory tool in the visualization system. We will see if the visual prototype is effective.