Active Learning Assignments for Student Acquisition of Design Principles

Marti A. Hearst*

University of California, Berkeley

Abstract

Active peer learning has been shown in many studies to increase student engagement and learning. We are designing, developing, and assessing several new assignments for encouraging student acquisition of design fundamentals and gaining practice in creating visualization designs.

1 Introduction

For the last three years, I have been developing methods for active peer learning [5, 12, 17, 20, 19, 14, 3, 21] and introducing them into my Masters-level courses. I presented the foundations of this attitude towards teaching, which I call *Teaching as Coaching*, at a panel at Infoviz 2015 [10]. In this abstract I describe some novel assignments that we are developing to present an alternative active learning approach to the teaching of core information visualization principles. I plan very soon to introduce these as well into undergraduate and online courses.

2 New Methods for Viz Design Feedback

Principles of good design of standard information graphics are usually taught in terms of the visual variables of Bertin [1], enhanced by the scales put forward by Cleveland and McGill [4] and Mackinlay [16]. Few and Cairo describe these lucidly in their books [8, 2].

However, in my experience, students regularly violate these principles when designing their own visualizations. Therefore, it is important to give students a rigorous foundation in the design of basic charts and graphs before setting them loose on more innovative and creative forms of information visualization.

How to teach these principles? What can we do beyond the standard pedagogy of learning the known principles via readings and lecture material? In my courses I have been attempting to reinforce the ideas via active learning, using a combination of studying existing examples and active expression of the concepts by creating new examples. However, to get the best results from active learning, it is important for students to receive rapid, actionable feedback on these exercises.

For responding to existing examples, problems can be devised with known solutions and students can be shown expert critiques as well as answers to multiple choice and other kinds of quiz questions to help them assess their understanding. But high-quality feedback is more difficult to produce for designs that students produce de novo.

There are several challenges to providing feedback on students' designs. The first is of course the time and effort this requires. But a less obvious issue is that it can be difficult to give helpful and accurate feedback on designs. This is not always a matter of opinion; what makes for a good design critique of course depends in part on what the criteria for a good design have been defined to be. For design of human-computer interfaces, in some cases, the criteria related to aesthetics, and in others to functionality, or understandability, or efficiency with which users can accomplish tasks using the design or conformance to standards, or ability to be implemented, or feelings of well-being and satisfaction, or some combination of all of these.

*e-mail: hearst@berkeley.edu

For a high-dimensional design space, as in the case where students are asked to plot a large table of data, sometimes the instructor is just going to be *incorrect*, and in other cases, feedback from more than one informed perspective is valuable.

2.1 Objectively Evaluated Visualization Assignment

A standard approach in the world of human-computer interaction is to solicit feedback from several different experts, and a common technique is the expert or heuristic evaluation, which is part of the discount usability model made popular by Jacob Nielsen and others [18]. More recently, researchers have been studying ways to leverage crowdsourcing to provide additional critiques [22, 15].

As an alternative, I was interested in developing a way to give feedback to students about their chart designs that was straightforwardly *objective*. I wanted at least one assignment where an external group told the students how good their designs were, rather than having my or my TA's subjective responses determine their feedback alone.

The approach I took was motivated in part by Dow et al. [6]. I built on three important aspects of that work. First, Dow et al. [6] experimented with whether it is better for designers to start with one design and iterate on that one design, or to start with several different ideas and iterate on those in parallel. They found empirically that the latter was better; that iterating on designs in parallel led to better designs. Second, they compared having designers work alone on their ideas or work on small teams, and found that the best results came from people first generating ideas on their own and then coming together to combine the best of their ideas to create better ideas together (and in parallel).

Finally, rather than evaluating the designs via heuristic evaluation or some other critiquing method, they came up with the clever solution of letting the objective world determine which designs were "best" along some criteria. The task they assigned participants was to design a poster to advertise a community service announcement. Then they placed the final designs for each group on the web as an advertisement, and calculated how many click-throughs the advertisement generated. The better designed the poster, by these criteria, the more likely the target audience to click on the ad to learn more about the announcement.

I found these ideas to be compelling and wanted to incorporate them into my information visualization course. Although I felt the use of advertisements was innovative in that it reflected the response of uninvolved third parties who were truly reacting to the design, I sought a method that more closely reflected the principles I was trying to teach about fundamentals of usable design and visualization principles.

From this desire was born the idea for the Objectively Evaluated Visualization Assignment (OEV). The main idea of the assignment is that a good visualization allows for members of its target audience to be able to interpret it correctly and answer questions about the underlying data. In class, one of our mottos came from Cairo's book — clarify, don't simplify! The goal of the assignment was to objectively assess how well students' designs allowed people viewing those designs to answer questions about the underlying data.

This kind of evaluation falls under the Evaluating Visualization Algorithms (EVA) category of Lam et al's taxonomy of evaluation techniques [13].

To accomplish this, I selected a dataset and told the students to create designs that would make it possible for a lay person to answer questions about that data. I gave a general idea about the types of questions that would be asked, but did not reveal the exact questions in advance. When preparing the questions, I made them challenging; I tried to make them cover a wide range of information types, so that the best visualizations would have to show both details and trends simultaneously and compactly. (Designs had to be static and were limited to 800x600 pixels.)

The reason for not telling the students the actual questions in advance is analogical to not telling programming student the actual test cases in advance – to avoid allowing them to "hard-code" the ways the visualizations will be interpreted, thus requiring the students to make designs that can be interpreted in manners beyond their initial intentions.

I then showed the students' designs to neutral representative parties — crowd workers on Mechanical Turk — and asked these people to answer questions about the underlying data. The assumption was that the better the design, the more questions would be answered correctly.

In most cases, those designs I would have predicted were good designs did well, but I was most definitely surprised in some cases. The students' designs had a good spread from high-scoring to low-scoring. I had the students whose designs did especially well discuss those designs in front of the class. I did find that I needed to give individual feedback for the designs that did not do well.

Another unexpected benefit of the assignment is that in some cases high-scoring solutions suggest novel (to me) design approaches that I know work well as people were able to answer questions successfully with them.

After doing this exercise twice, I feel that it worked well at achieving its goals and decided to assess it in a formal study.

2.2 Visualization Assessed Two Ways

In Summer 2015 my collaborators and I ran a study assessing the Objectively Evaluated Visualization idea, but with two significant changes [11]. Because students whose designs did not score well had asked for qualitative feedback, we decided to compare the results of a standard qualitative assessment method – heuristic evaluation – with the objective measure of seeing how many questions about the underlying data could be successfully answered using each design. The other significant change was to have students do the assessment rather than crowd workers. However, due to time constraints (the study was doing as part of a 8 week summer school course on HCI, with only one lecture on Infoviz), the designs themselves were made by the experimenters.

The results were surprising – the students' average assessments of quality using three heuristic guidelines (derived from [7], on a 5 point Likert scale) correlated rather well with their average scores on answering questions about the data underlying the designs (students did not answer questions about the same data sets that they did heuristic evaluations of). We realized we had a fundamentally new method of evaluating designs, and call the technique Visualization Assessed Two Ways.

This suggests that peer evaluation could be run at scale, and the two different methods used to offset one another. A common complaint from students about peer assessments is that peers might not recognize a truly innovative design. An advantage of the Visualization Assessed Two Ways approach is that a discrepancy between the two measures may indicate that something unusual is going on. We can think of the approach described here as a way for instructors, or for anyone interested in finding alternative ways to assess a design, to be warned if there is a major discrepancy between the two methods. An analogy can be drawn voting algorithms or ensemble techniques in machine learning, which combine the output of several algorithms. These tend to produce a better result than any algorithm alone, but only if the algorithms hail from different underlying distributions [9].

A big advantage of this approach is that many different assessments can be obtained, and a database of different designs for the same dataset created, each with feedback on both the qualitative and quantitative performance. New designs can be mapped into this space, providing a way for a student to compare their design to other designs that have also been assessed, to help ascertain which aspects work well and which need improvements. This data base will become a valuable asset for the instruction of Infoviz both in person and online.

In Spring 2016, I successfully reproduced the results from the Summer 2015 course in my masters level Infoviz course, using three different datasets, and having students design the visualizations in pairs using the Highcharts javascript library. Students designed the visualizations as well as doing the heuristic evaluations and answering the questions about the data sets (three datasets were needed so students were not answering questions about known data values). The correlations were quite similar to those of the original study, thus lending further credence to this approach.

The next step will be to repeat this in a larger course. However, curating the data sets and creating the questions for the quantitative part of the analysis is quite time consuming for the instructor, so a further step will be to have students do this part as well, after training them in the creation of such questions and the analysis of data for this purpose.

3 Games for Teaching Perceptual and Design Principles

In order to prepare students for the OEV and VATW, they need to first learn the central ideas about visual perception and pre-attentive processing, selection of visual marks, Gestalt principles, consequences of color selection, and so on. A major goal of the effort is to have students acquire ownership of usability principles of information visualization, rather than having it dictated to them.

Some common approaches to making charts have poor usability but are visually appealing. These include charts containing circular marks (pie charts, bubble charts), some kinds of choropleths, tag clouds, and using color to indicate quantitative values. However, for many students, simply presenting them with the scientific evidence against these approaches is not sufficient to convince them about their poor usability.

We are developing game-like activities to see if exposing the students first hand to trying to solve problems using different variations of visualizations, paired with reflection after the fact, will lead to discovery of and acceptance of these visual principles in a more convincing manner than standard instructional approaches. We are currently devising several different game play scenarios and plan to compare these to more conventional methods. We should have results to report by the time of the workshop.

Acknowledgements

This research was sponsored in part by a Google Social Interactions Grant.

References

[1] J. Bertin. Semiology of graphics: diagrams, networks, maps. 1983.

[2] A. Cairo. *The Functional Art: An introduction to information graphics and visualization*. New Riders, 2012.

[3] J. Chase and E. G. Okie. Combining cooperative learning and peer instruction in introductory computer science. *ACM SIGCSE Bulletin*, 32(1):372–376, 2000.

[4] W. S. Cleveland and R. McGill. Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American statistical association*, 79(387):531–554, 1984.

[5] C. H. Crouch and E. Mazur. Peer instruction: Ten years of experience and results. *American Journal of Physics*, 69(9):970–977, 2001.

[6] S. P. Dow, J. Fortuna, D. Schwartz, B. Altringer, D. L. Schwartz, and S. R. Klemmer. Prototyping dynamics: sharing multiple designs improves exploration, group rapport, and results. In *CHI*. ACM, 2011.

[7] S. Few. Visual communication: Design principles for displaying quantitative information. *Cognos Innovation Center*, September 2006. http://www.perceptualedge.com/articles/Whitepapers/ Visual_Communication.pdf.

[8] S. Few. *Now you see it: simple visualization techniques for quantitative analysis*. Analytics Press, 2009.

[9] Y. Freund and R. E. Schapire. Experiments with a new boosting algorithm. In *ICML*, volume 96, pages 148–156, 1996.

[10] M. A. Hearst, E. Adar, R. Kosara, T. Munzner, J. Schwabish, and B. Shneiderman. Vis, the next generation: Teaching across the researcher-practitioner gap (ieee vis panel). 2015.

[11] M. A. Hearst, P. Laskowski, and L. Silva. Evaluating information visualization via the interplay of heuristic evaluation and question-based scoring. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 5028–5033. ACM, 2016.

[12] D. W. Johnson, R. T. Johnson, and K. A. Smith. *Active learning: Cooperation in the college classroom*. Interaction Book Company Edina, MN, 1991.

[13] H. Lam, E. Bertini, P. Isenberg, C. Plaisant, and S. Carpendale. Empirical studies in information visualization: Seven scenarios. *Visualization and Computer Graphics, IEEE Transactions on*, 18(9):1520–1536, 2012.

[14] T. Lord. Cooperative learning that really works in biology teaching: using constructivist-based activities to challenge student teams. *The American Biology Teacher*, 60(8):580–588, 1998.

[15] K. Luther, J. Tolentino, W. Wu, A. Pavel, B. Bailey, M. Agrawala, B. Hartmann, and S. P. Dow. Structuring, aggregating, and evaluating crowdsourced design critique. CSCW, 2015.

[16] J. Mackinlay. Applying a theory of graphical presentation to the graphic design of user interfaces. In *Proceedings of the 1st annual ACM SIGGRAPH symposium on User Interface Software*, pages 179–189. ACM, 1988.

[17] B. J. Millis and P. G. Cottell. *Cooperative learning for higher education faculty*. Oryx Press (Phoenix, Ariz.), 1998.

[18] J. Nielsen. *Usability engineering*. Elsevier, 1994.

[19] M. K. Smith, W. B. Wood, W. K. Adams, C. Wieman, J. K. Knight, N. Guild, and T. T. Su. Why peer discussion improves student performance on in-class concept questions. *Science*, 323(5910):122–124, 2009.

[20] L. Springer, M. E. Stanne, and S. S. Donovan. Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. *Review of educational research*, 69(1):21–51, 1999.

[21] D. A. Trytten. Progressing from small group work to cooperative learning: A case study from computer science. *Journal of Engineering Education*, 90(1):85–91, 2001.

[22] A. Xu, H. Rao, S. P. Dow, and B. P. Bailey. A classroom study of using crowd feedback in the iterative design process. 2015.