Open Research Online

The Open University's repository of research publications
and other research outputs

Prototyping Visual Learning Analytics Guided by an Educational Theory Informed Goal

Journal Item

For guidance on citations see FAQs.

Version: Version of Record

Link(s) to article on publisher's website:
http://dx.doi.org/doi:10.18608/jla.2016.33.7

oro.open.ac.uk

Prototyping Visual Learning Analytics Guided by an Educational Theory Informed Goal

Garron Hillaire
CAST, Inc. USA
The Open University, IET, UK
Garron.Hillaire@open.ac.uk

Gabrielle Rappolt-Schlichtmann
The Harvard Graduate School of Education
Ed Together[1], USA

Kim Ducharme
Northeastern University
CAST, Inc. USA

ABSTRACT: Prototype work can support the creation of data visualizations throughout the research and development process through paper prototypes with sketching, designed prototypes with graphic design tools, and functional prototypes to explore how the implementation will work. One challenging aspect of data visualization work is coordinating the expertise of people from a variety of roles to produce data visualizations guided by an educational theory informed goal (ETIG) in order to better support research. When collaborating, concessions must be made: typically, everyone seeks to follow the best practices established within their own disciplines. This paper attempts to illustrate how to rethink this interdisciplinary approach to adhere more strictly to educational research goals and consider how we may need to, at times, break away from best practices with the intent to evaluate the novel decisions resulting from this approach. A case study of the creation of a self-reported emotional measure is used to illustrate this type of collaboration. By taking this approach, a clear departure from best practices occurs in the scale selection for the visualization in order to support the ETIG.

Keywords: Data visualization, visual learning analytics, universal design for learning (UDL), paper prototype, formative evaluation, summative evaluation

1   INTRODUCTION

At the Learning Analytics Summer Institute 2014, we held a workshop on using the paper prototype process to create data visualizations. In this workshop, we illustrated a process that started with clarifying the educational theory to be used for the data visualization and worked with participants to pursue prototyping data visualizations for their own projects. Productive feedback fell into two

---

[1] www.ed-together.org

categories. The first category of feedback was that the workshop would have moved much more smoothly if we first provided a worked example that illustrated the process so that participants could have a positive model of the approach. The second was a desire to understand more behind the rationale of starting with educational theory first when creating visual learning analytics. In this paper, we will address both of these concerns raised during our workshop. First, we will articulate why we think starting by establishing an educational theory can be beneficial and then use a worked example as a case study of how to start from educational theory and work toward creating a data visualization. As with all processes, the steps are a guideline and; each phase of the work can be conducted as by an individual or a group. In the worked example, we will illustrate how throughout the process we converged and diverged between group work and individual work.

One general framework for thinking about how to create data visualizations is computational information design that suggests that it takes a multidisciplinary approach to acquire, parse, filter, mine, represent, refine, and interact with data (Fry, 2004). Given the interdisciplinary nature of the work, one can imagine that coordinating those efforts can be challenging. At times, this type of work feels like the various experts can attempt to guide the direction of the group toward their familiar territory of best practices. Fry articulates the challenge of this type of inter-disciplinary work by calling it "a 'telephone game' for context, where each step of the process diminishes aspects of the initial question under consideration" (Fry, 2004). It may be that in order to collaborate outside of their comfort zones, many experts would need to have some reason to challenge their conventional thinking. In the case of Visual Learning Analytics, it may be possible for pedagogy to guide decisions from data collection through interactive data visualizations. While the boundaries outlined in Fry's model can be debated, we will use the relationships for the work of computer science, mathematics, statistics, data mining, graphic design, infovis, and human computer interaction (HCI) as an illustration of how pedagogy can coordinate the efforts for the purposes of creating data visualizations that align with educational goals. The important questions are these: Where are standard practices challenged by taking this approach? How can those innovative decisions be evaluated? Given that the educational goal will be the driving force behind the innovative behaviour, it can always be used when establishing an evaluation strategy. In the following case study of an iterative process, we will follow a series of steps that any team or individual working on educational data visualization could adopt and use.

2   CONTEXT FOR THE WORKED EXAMPLE

In the United States, a significant challenge persists in closing the gap for struggling readers at the middle-school level, especially among those with disabilities (CCAAL, 2010; Lesaux & Siegel, 2003). The National Assessment of Educational Progress reported in 2013 that 60% of eighth grade students with disabilities performed below basic level in reading (note that these statistics reflect only students who can be assessed without accommodations (NAEP, 2013). More specifically, students with learning disabilities read 3.4 grade levels below their peers without disabilities (NJCLD, 2008; NLTS II, 2003; Wagner et al., 2003; Edmonds et al., 2009; Gajria, Jitendra, Sood, & Sacks, 2007). In addition, many students with learning disabilities experience significant social and emotional strain in school, hold

negative emotions about reading, and exhibit low motivation for reading including text avoidant behaviour (Bryan & Bryan, 1983; Maag & Behrens, 1989; Margalit & Raviv, 1984; Sideridis, Mouzaki, Simos, & Protopapas, 2006; Sutherland & Singh, 2004). These findings are concerning because, by middle school, academic outcomes become inexorably linked to students' independent ability to learn from text and then express what they know through text (Murnane & Levy, 1996; Biancarosa & Snow, 2006; Perfetti, Landi, & Oakhill, 2005).

To address these challenges, the Center on Emergent Technology, Disability, and Middle School Reading (CET; cet.cast.org) is researching and developing a technology-rich learning environment that will make it possible for schools to provide *all* students with universally designed for learning (UDL) literacy experiences as they read independently across the curriculum. UDL offers a means to provide opportunities for deep learning through the design of highly flexible methods, materials, and assessments (Meyer, Rose, & Gordon, 2013; Rose & Meyer, 2002; Rose, Meyer, & Hitchcock, 2005). The basic premise of UDL is that barriers to learning occur in the interaction with curriculum — they are not inherent solely in the capacities of the learner. Just as universally designed buildings provide options that accommodate a broad spectrum of users, tools and curricula designed using the UDL framework offer a range of options for accessing and engaging with learning materials. When developed under the UDL framework, digital environments provide the necessary infrastructure and flexibility to allow for the creation of accessible, highly effective apprenticeship environments where students are actively guided in the process of constructing meaning through the provision of just-in-time feedback and contextual supports that can be gradually withdrawn as student expertise increases (Collins, Brown, & Newman, 1989; Palincsar, 1986, 1998; Palincsar & Brown, 1984).

CET's goals are to leverage emergent technology in the service of reading comprehension, skill development, and improved motivation for reading, so that whenever students are reading they will be able to practice and grow their reading comprehension skills through "just-right" reading challenges. More specifically, CET is leveraging the capabilities of the Web itself to create a more distributed ecosystem for adolescent literacy, termed "Udio," where the goals are to 1) foster a passionate interest and investment in reading for students who have traditionally been uninterested in, or disenfranchised by, traditional classroom literacy practices; and 2) to substantially improve the reading comprehension skills of middle-school students who have experienced recurrent failure in the domain of reading.

Udio is a dynamic web of authentic, current, and widely distributed UDL reading environments connected by a central Hub. Reading materials connected through Udio represent a collaborative network of partners whose online content is already on the web and of high interest to struggling adolescent readers and their teachers. Current network partners include Scholastic, Facing History and Ourselves, the SERP Institute/Word Generation, the Poses Family Foundation/Story Share, The Courage Curriculum, Science Friday, and *Scientific American*. Network partners are working with CET to render their online content compliant with the Web Content Accessibility Guidelines (WCAG) version 2.0 and to allow for the provision of UDL supports and scaffolds to facilitate reading comprehension development. In this way, reading materials within Udio are "tuned" to students' specific strengths and weaknesses

around reading comprehension, effectively turning the Web into a place where anyone can learn to comprehend text effectively. At the core of the reading network, Udio's Hub provides a home base or platform from which students and teachers can explore the wide network of reading experiences provided and say or write something about what they have read. The Hub provides opportunities for students to monitor their progress and growth through data analytics and is designed explicitly to support the development of metacognition and positive feelings about reading.

3   STEP ONE: DEFINE THE EDUCATIONAL GOAL

We suggest starting by defining the educational goal of a data visualization because it can be leveraged through the process to ensure that all efforts are aligned toward achieving that goal. As you revisit the goal for guidance throughout the process, it is common to refine your goal. For the worked example, we have completed the initial implementation and are moving toward piloting the work and gearing up for some future work so the following goal has been refined multiple times.

The educational goal that we will use as an example of visual learning analytics informed by educational theory is to explore a universal mechanism by which students could reflect on their emotional reactions to learning material. This is framed from the perspective that the role of emotion in learning is complex. Based in social interaction, emotions are biological processes that organize human behaviour by constraining thought and action as the environment is appraised as either beneficial or detrimental (Fischer & Bidell, 1998; Frijda, 1986; Lazarus, 1991). Importantly, positive and negative emotions can, each on their own accord, activate or deactivate motivation states, and enhance or detract from productive learning (Bower, 1981; Eich, Kihlstrom, Bower, Forgas, & Niedenthal, 2000; Serrell, 2006; Steidl, Mohi-uddin, & Anderson, 2006). The goal in researching this topic was to explore if we could provide students with an interactive data visualization where they could reflect in useful ways on their experience of reading. Given the goal of a universal mechanism related to reflecting on emotion, we decided to focus on the concept of core affect from the research literature.

Core affect is a preconceptual process, a neurophysiological state that actively reflects one's evaluation of the environment in two dimensions: feeling good or bad, feeling lethargic or energized (Russell, 1980, 2005, 2009). Core affect is continuous in that people are evaluating the environment relative to these dimensions constantly and subconsciously. The concept is similar to mood, but independent in that core affect is more variable and responsive to stimuli in the environment, like content one might read online. You can imagine reading an article that does not align with your point of view or where you feel the author is not providing adequate evidence; as you read, your core affect changes in relation to the content. You feel more energized, and perhaps negative.

Often core affect is free-floating — you might feel down, slightly negative and lethargic, but not know why. However, there is significant evidence that the experience of core affect can become conscious and intentional (Searle, 1992; Russell, 2009). As one considers or is prompted to consider how they feel relative to these two dimensions of core affect (pleasure-displeasure, activation-deactivation) one can

apply "feeling" words to their experiences (happy, sad, annoyed, interested, bored), thus raising core affect to conscious awareness, allowing you to become intentional about your experiences: "This topic was really boring for me. I'm going to try another topic area" or "This author's opinion was really annoying. I am going to create a project and contest his ideas with other evidence."

When providing high-interest content, such as in environments like Udio, we noted that "Liking" an artifact (e.g., article, post, video, etc.) provides very little information to others about why that article was useful, interesting, or how difficult or frustrating the artifact was per se. Effectively the "like" feature cannot provide a mechanism to reflect on the complex role that emotions play in learning. There is evidence that "likes" on web content are primarily a reflection of popularity — an online article that has already been "liked" is more likely to attract additional "likes" regardless of usefulness or quality.

To address the weaknesses of the "like" feature, our team sought to provide a mechanism by which students would be able to reflect emotionally on their experience of articles and reading within Udio. This mechanism would need to provide useful information to other readers simultaneously when browsing content to inform their choices about what to read, and support the development of student motivation for reading as they finish one article and move on to another. To do this, we took the idea of "liking" and, using core affect as a guide, worked to align a new design of this concept to the research literature describing emotion in learning.

4   STEP TWO: DEFINE TARGET USERS

Udio's primary target audience is middle-school students with learning disabilities. To effectively reach this target audience, however, we recognize that we cannot further segregate or stigmatize that group by creating a primarily remedial or "special needs" program that effectively separates them, cognitively or socially, from their peers without disabilities. At this especially vulnerable age, where peer relations take on a dominating influence and where the development of autonomy and self-regulation are critical developmental challenges, it will be vital to reach our target audience through an approach that is inclusive rather than exclusive, designed to challenge, motivate, support, and educate not only students with learning disabilities, but their peers (and teachers) too.

Furthermore, as CET is rooted in the universal design movement it is important for us to reach our target audience using approaches that include the full range of their peers, including students with "other" disabilities. From a UDL perspective, students with all kinds of disabilities have a unique role to play in design because they are particularly vulnerable to inflexible, "one-size-fits-all" solutions; they represent the edge of variability within the population that guides us toward more successful, and inclusive, solutions. By attending to the challenges faced by individuals with many disabilities in the design process, learning environments can be made more accessible, engaging, and effective for the widest posible array of learners. For these reasons, we include struggling readers without identified learning disabilities and students with low-incidence disabilities (specifically, autism and intellectual disabilities)

in our target research populations (the latter because so little is known about reading comprehension at the middle-school level for these groups).

5 STEP THREE: INTERDISCIPLINARY PAPER PROTOTYPING PROCESS

When considering the design of visual learning analytics, the paper prototyping process within a design thinking methodology has many benefits (Brown, 2009; Brown & Katz 2009; Brown & Katz 2011). First, it supports the collaboration of a variety of disciplines in the co-design process. We recruit input from the technology, design, and research departments representing skills across computer science, mathematics, statistics, graphic design, and information visualization. Second, it is ideal for communicating and testing ideas early on and iterating faster. Finally, it allows you to focus the feedback on core intention and interaction before putting too many resources into higher fidelity visual design and implementation.

The basic idea is to engage in several rounds of individual and group ideation and synthesis. This optimizes the playing field to allow for variability in contribution styles, individual and group think, and to open the creative gates to the potential for cross-pollination and mixing of ideas. In practice, each phase can be done as a group or as an individual. For our team, the process often looks like a variation on Tim Brown's IDEO diverge/converge (Brown & Katz, 2009). Many times the selection of individual versus group work is a matter of convenience and availability. However, there are times when a difficult problem has been identified and getting everyone together to explore creative solutions helps to move the work along.

5.1 Defining the Brainstorm Problem

Defining the problem, and keeping it top-of-mind while ideating solutions, is crucial. This top-level problem statement was generated based on the ETIG to explore a universal mechanism by which students could reflect on their emotional reactions to learning material. We use a useful "POV Madlib" format from design thinking methodology: (Person) needs a way to (person's need), because (insight) (Barry, 2010). It is very important to articulate the problem, but not yet ideate solutions. That means filling in the blanks, but resisting the temptation to say "how" just yet. Our problem statement was to get at how we might facilitate a student's self-reporting of emotional metrics: *Student-A needs a way to express how they feel about the article they have just read, and have the experience that their opinion matters and is being counted.* The inclusion of having the experience that their opinion matters is a refinement of the ETIG to include that the experience should be intrinsically rewarding. By connecting our ETIG to the design thinking process, we gained further clarification of the goal.

Once we are all on the same page about the problem to solve, we brainstorm around a series of "how might we…" seed questions:

- How might we create something that allows the student to self-report an emotional reaction to an article?

- How might we use data visualization to support and encourage the student to weigh in?
- What activity data set might we leverage?
- What social aspects might make a feature/functionality more attractive or compelling to use?
- How might we devise a mechanism that allows students to provide useful information to other readers?

Individually, people write down ideas, one per post-it, and then put them on the wall. Ideas can then be viewed, discussed, added to, and grouped. This allows for multiple modes of input — individual and group. The final step here is to decide as a group on the 1–2 best ideas (we sometimes pull out the dots for this, voting against various criteria to help us decide). When working with bigger groups, post-its are put up centrally and each group reports on their 1–2 ideas, in brief, to the whole room.

5.2   Diverge: Sketch and Solutions

With an ETIG to test in hand, ideas for corresponding data visualizations to leverage, and activity data set to use, there is sufficient groundwork in place to support some rapid idea generation. A technique referred to as "crazy 8's" (Evan, 2011) is good for quickly generating many good, original ideas, and helps avoid stopping at the first good idea. In order to participate in a crazy 8 activity, each participant folds a paper in half three times to create eight sections. The participants then take five minutes to make eight quick sketches on high-level ideas, and variations of ideas. One technique to support the rapid nature of this method is to use a fat sharpie as it keeps one from getting too detailed. Constraining the time also helps to keep participants from getting too detailed. Figure 1 illustrates two examples of crazy 8 sketches.
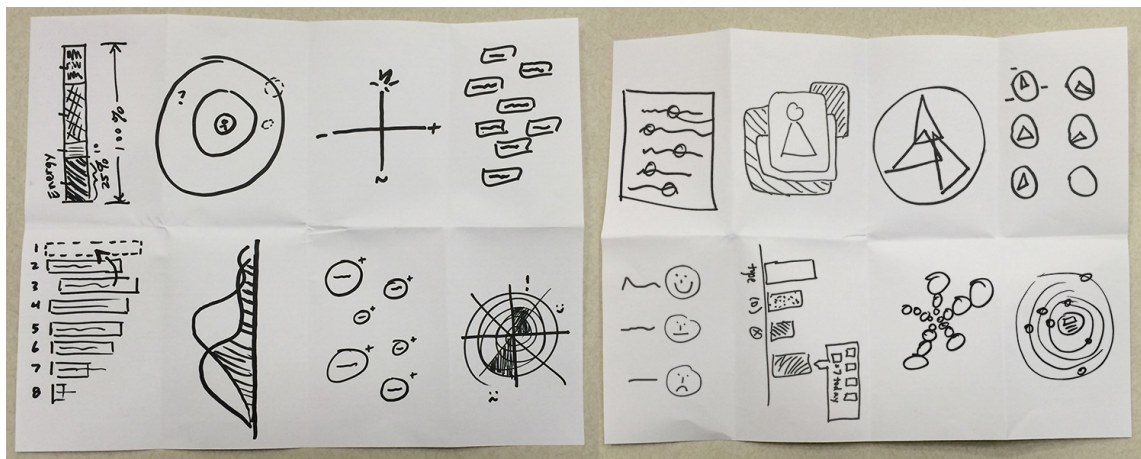

Figure 1. Two examples of crazy 8 sketches.

A facilitator (who has been keeping track of time) can then ask people to stop drawing and explain their ideas. A second round of crazy 8's supports individuals to iterate, and borrow (steal!) from a teammate's good ideas, mix, re-mix, or think up entirely new ideas. After a few rounds of crazy 8's, the group then selects one or two best ideas to refine.

Another approach to achieve the same outcome is to collaborate on a whiteboard. While the crazy 8 structure helps to ensure multiple perspectives are expressed in the prototyping phase, it can be helpful at times to follow up that exercise with a group whiteboarding exercise to further refine sketches. For the worked example of the React implementation, we skipped the crazy 8 step and started with a collective whiteboard exercise. In Figure 2, we honed in on the idea of representing emotional dimensions on axes and a windrose as a nicely expressive visualization type.
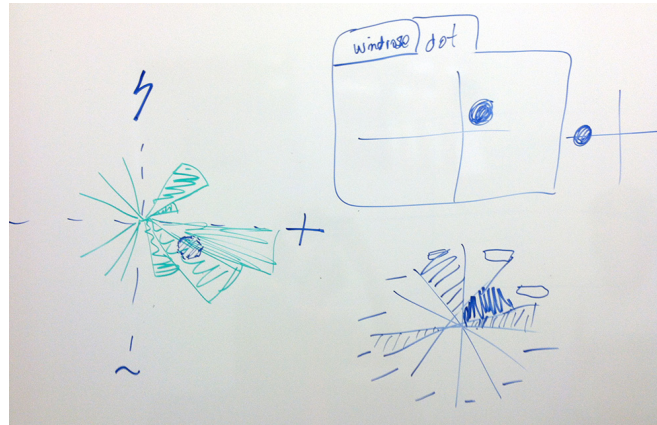
Figure 2. Example of a whiteboard outcome of converging collaboration.

5.3   Converge: Refine

One of the first steps in refining the data visualization is to clarify the realistic expectations. As a group, teams can make a list of elements clarifying those expectations. For example:

- Content requirements (what your prototype needs to do, how it functions, what information it contains, and the nature of the data set)
- Primary info/experience (e.g., displays/graphs, data points, interactions with, etc.)
- Secondary info (e.g., intro text, menus, dates, labels, links, etc.)
- Navigation — within, and global
- Make preliminary notes on what you will test or ask of a user

Sketching with a real data set is ideal. When this is not an option, it can work well to create an imagined or fabricated data set, either before initial sketching or before the refining stages (e.g., inside of a given session, my user will click into one article and self-report an emotional reaction to that article). Data sketching — trying out various visual outputs using charting software — is a great way to get a sense of the potential shapes of your data set and the necessary real estate it may require.

Individually, we create more detailed sketches of 1–2 main screens or displays, full size, adding annotations. After backing up all the work, we present refined screens for discussion and synthesis, then

decide on a single direction, and settle on the details. This can still involve a bit of mixing and matching, bringing together the best ideas to work as one solution.

This step includes additional screens. Each person takes a screen to refine, illustrating and/or annotating functionality, interactions, and navigation within and between screens, and display states to simulate animation (change over time). We also use post-its or tracing paper to overlay information.

Another person creates annotated storyboards of the flow and navigation within and between screens and displays, and gets more specific on a testing protocol. In our case, we wanted to test interface functionality and a few different displays of the data set.
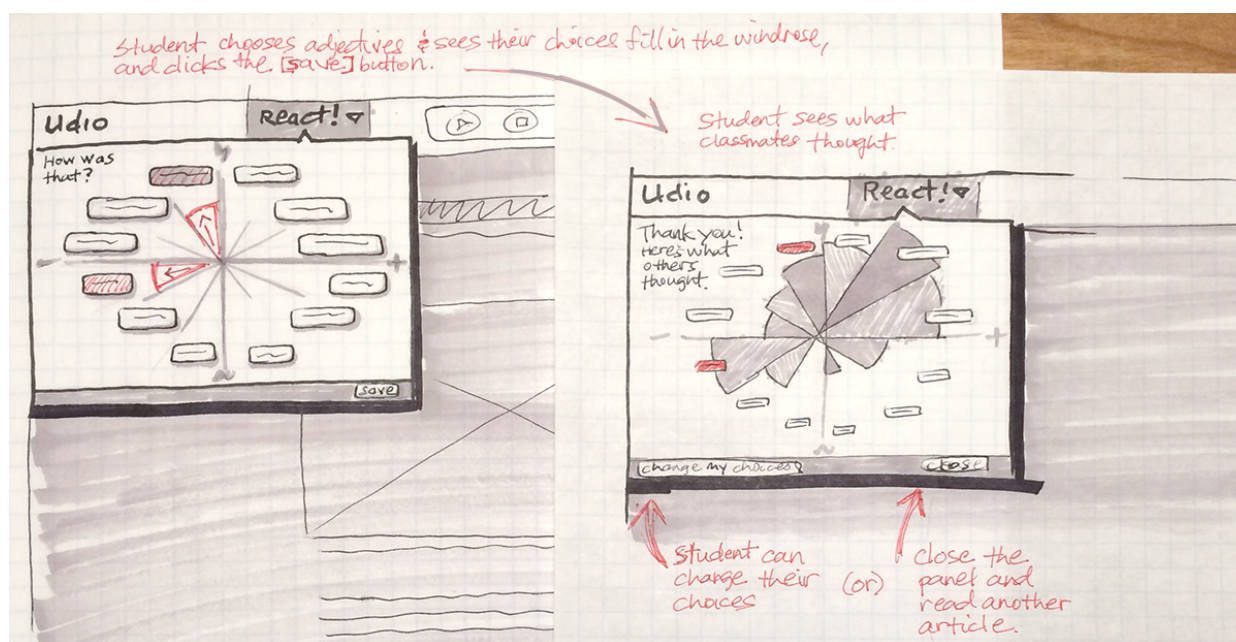

Figure 3. Example of refined sketch of "React" data visualization.

Refined sketches, as illustrated in Figure 3, can be used as-is for testing only when they have been created with sufficient clarity. Otherwise, they can serve as a basis for creating preliminary visual designs or functional wireframes for testing. When working on the implementation of the interactions of students selecting their reactions and viewing others' reactions, we wanted to accomplish two things. First, we wanted to make sure that the interactive design supported participation; second, we wanted to make sure it was easy for a student to toggle between seeing their reactions and seeing others' reactions, as the comparison would create an opportunity for students to reflect. The goal of supporting a comparison between an individual response and the responses of their peers was intended to support the ETIG aim of encouraging reflection on their emotional reaction. The exploration of comparing your response to the responses of your peers is an attempt to acknowledge the social nature of emotions outlined in the ETIG. The example of the buttons highlighted in red illustrated a design support that allowed students to compare their responses to others' reactions. One of the first design challenges that

arose from this wireframe was asking the question, Should a student be allowed to see others' reactions before providing their own? While it could influence the student's response, it would support the student by benefiting from reflecting on others' reactions even without responding. If it can help spark further questions about implementation, then you have a successful sketching process.

6 FORMATIVE EVALUATION

We conducted a few different rounds of testing. Through early paper prototyping, we gathered a preliminary data set by asking students to match adjectives to how they felt about an article they had just read.
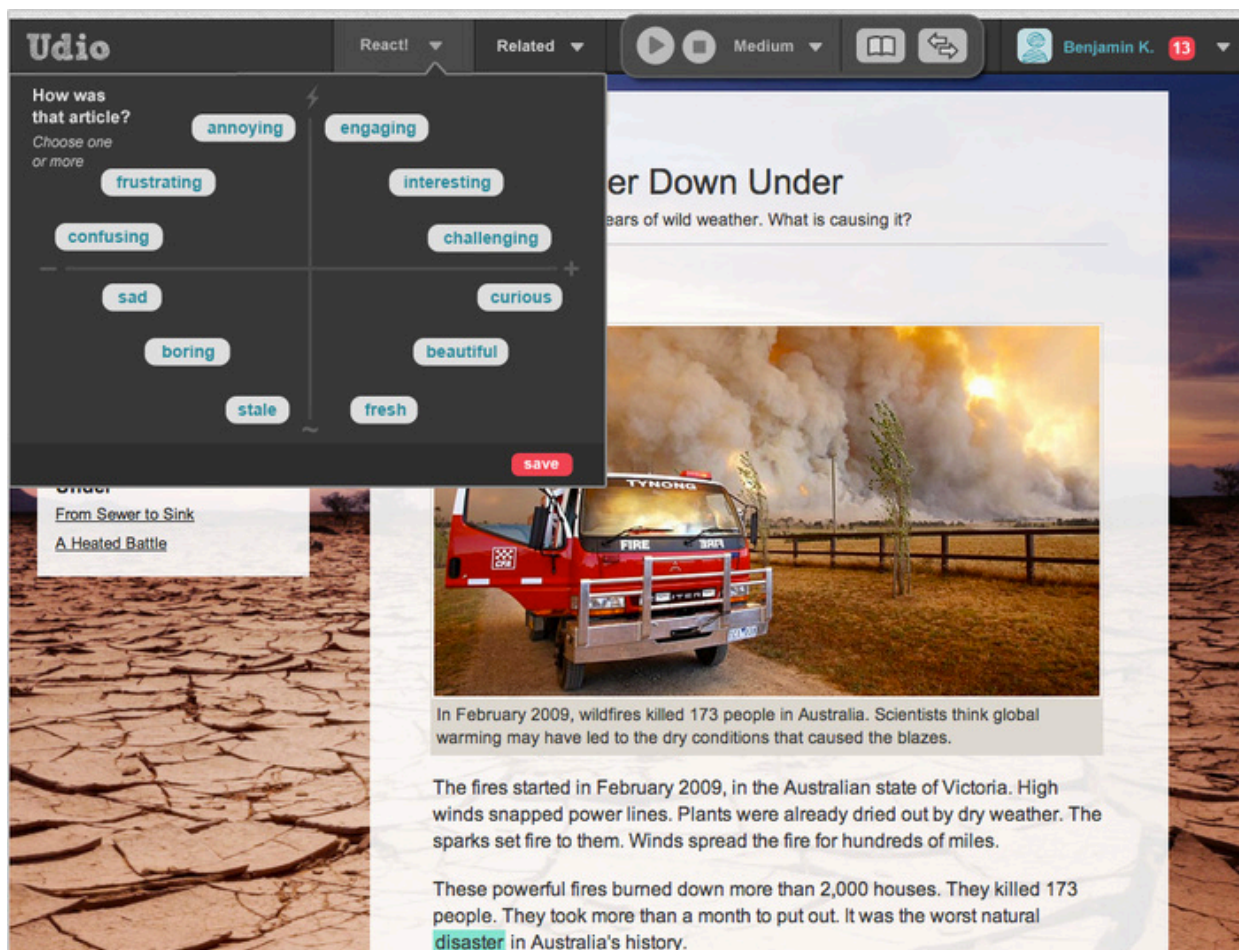

Figure 4. Example paper prototype for formative evaluation.

The tester showed students a printout of Figure 4 and asked them to tap their finger on the React button in the screen's printed interface. Then the tester overlaid a version of the screen with the response adjectives. The tester read the words as they appeared in the interface, "How was that article? Choose one or more (by pointing with your finger)," then asked the student to tap the save button. The

student was then asked what they thought would happen next. The expectations were recorded for future analysis but not examined in depth at this stage.
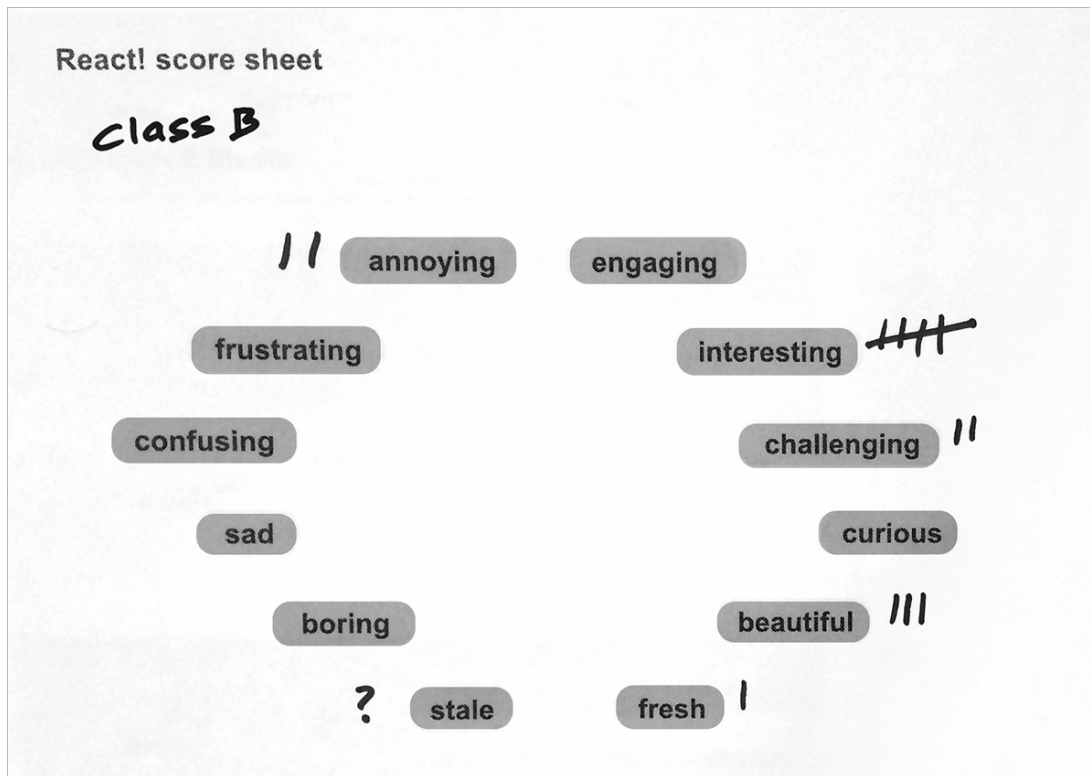

Figure 5. Example paper prototype for note taking during formative evaluation.

The tester made notes and tallied the responses with tick marks on a score sheet, as illustrated in Figure 5, so that we could collect a real data set from our users. From multiple rounds of testing, these tally marks were aggregated to generate a preliminary data set. With the preliminary data set in hand as a reference, we created visually refined screens and data visualization variations to test. We printed out the visual designs in Figure 6 for another round of paper testing in the classroom. Prototype evaluation is typically done with several people. If you want a specific number, Nielsen (2012) has suggested that five people would be sufficient for a usability study.

This data collection also allowed for the identification of words unfamiliar to the target population. When looking for shortcomings, small sample sizes are fine since a single person can identify a problem with the design so long as that individual has similar characteristics to the target population (Nieveen & Folmer, 2013, p. 163). In Figure 5, you can see a question mark next to the word stale because a student asked what the word meant during the testing. Unfamiliar words were removed from the implementation.

We tested the interface elements in a single classroom with individual interviews of students with the same characteristics as our target population. The evaluation usability testing focused on a few different

displays of the data set for clarity and preference around different visualization types and colour variations.

Figure 6. Paper prototype of four options for viewing others' reactions.

Again, the testers went through the same initial steps, and the adjectives' selected states were marked with mini post-its. Once the student tapped the save button, a data visualization was overlaid while the tester read the words as they appeared in the interface "Thank you! Here's what others thought."

The following testing protocol outlines a series of questions that students were asked while looking at the options outlined in Figure 6:

- Start with one visualization type, and ask "Can you describe what you see?" and "What does it mean?"
- Give the user some time, and then if needed, point to the help text and read "Bigger circles mean more people chose those words" or "Bigger shapes mean more people chose those words." Then, prompt them to try to describe what they see again.
- Show the other variations to get at preferences around visualization types and colour variation, asking questions such as, "Is this one easier to understand," "Do you like the shapes in different colours, or all in the same colour." "Why?" Always ask why!
- How would you get back to change your answer?
- How would you move on to read another article?

When running the screens by users, students appeared to understand the concept. They were all able to interpret the information correctly. Moreover, while they understood the information associated with the adjectives when presented as differently sized bubbles, they preferred the windrose data representation. They found it more expressive while still being informative for representing the character of the article. Our small focus group of five students gave us the confidence that we could put together an implementation for a pilot.

7 MOCK DATA

When working on a data visualization, mock data can help in considering if your implementation can realistically display the possible range of values. For the "your responses" view, the first thing we define is the range of potential values; in this case it is 0 or 1, with 0 representing not selecting the option while 1 represents selecting the option. From this range, we consider a variety of potential sets of values for a given student. It is possible that a student has not responded at all, has selected one response, or has selected a few different responses. It helps to put a table of data together as a reference. See Table 1, which illustrates that Student 1 has no reaction, Student 2 has selected Engaging as her reaction, and Student 3 has selected Engaging, Curious, and Confusing as his reaction.

Table 1. Mock Data for Three Students

	Student 1	Student 2	Student 3
Engaging	0	1	1
Interesting	0	0	0
Challenging	0	0	0
Curious	0	0	1
Calming	0	0	0
Good	0	0	0
Dull	0	0	0
Boring	0	0	0
Sad	0	0	0
Confusing	0	0	1
Frustrating	0	0	0
Annoying	0	0	0

In generating mock data, we decided to use three classrooms (see Table 2). Classroom 1 was created with all React values set to 0 in order to simulate a classroom where none of your peers has provided a reaction to the article. For Classroom 2, we randomly chose numbers ranging from 1 to 5 for five React words to simulate a situation where some peers have provided a reaction to this article. For Classroom 3, we chose random numbers between 1 and 25 for nine React words to simulate a classroom where all of your classmates have provided a reaction to an article.

By referring to Tables 1 and 2, we can create nine potential combinations of responses. For example, Student 3/ condition 2 would create the following data set, as illustrated in Table 3.

Table 2. Mock Data for Three Classrooms

	Classroom 1	Classroom 2	Classroom 3
Engaging	0	0	1
Interesting	0	0	0
Challenging	0	1	12
Curious	0	0	5
Calming	0	5	0
Good	0	0	3
Dull	0	3	6
Boring	0	0	0
Sad	0	0	25
Confusing	0	1	20
Frustrating	0	0	18
Annoying	0	0	4

Table 3. Mock Data for the Two States of React (your response and others' response)

	Student 3	Classroom 2	Total Reactions
Engaging	1	0	1
Interesting	0	0	0
Challenging	0	1	1
Curious	1	0	1
Calming	0	5	5
Good	0	0	0
Dull	0	3	3
Boring	0	0	0
Sad	0	0	0
Confusing	1	1	2
Frustrating	0	0	0
Annoying	0	0	0

8   IMPLEMENTATION

Now that we have a clear educational goal, a paper prototype, and some mock data, we check in again with our target users for more insight into some particular implementation questions around chart type, scales and axis, colour, interactive design, and accessibility. These categories illustrate worked examples of a specific case implementation but we believe that the categories are useful for a variety of contexts. Each section will begin with the key question and then illustrate how we answered that question for the worked example.

8.1   Research the Chart Type

Key Question: What are the known strengths and weaknesses that others have identified about your chart type or similar chart types?

In our prototype, we decided to build from a windrose chart type. Leone Lelanne is commonly credited with introducing the windrose diagram in 1843 though it was first introduced by Guerry in 1829, and a related chart type, introduced by Florence Nightingale in 1858, is called the polar area chart (Friendly, 2008). Nightingale's chart was similar both in design and goal, as Nightingale explained that the diagram should "affect thro' the eyes what we fail to convey to the public through their word-proof ears" (Ehyde & Seneta, 2001). In a recent *Economist* article (2007), Nightingale's polar area chart was named one of the three best charts of all time. Tufte (2001) suggests that for smaller data sets, tables would outperform these graphics and "the only worse design than a pie chart is several of them." For the case study of React, we compare the implementation to a more traditional data visualization. Suda (2010) criticized charts like the windrose because it is hard to draw an imaginary arc or circle when comparing values whereas it is easy to compare values in a bar chart with an imaginary line. When visual learning analytics depart from standardized chart types and those charts have known weaknesses, it is important to understand the potential strength of the decision based on the ETIG so that the team can determine if the benefits outweigh the costs.

In this case study, it is possible that the cost of making precise interpretations of the chart actually supports the ETIG. A recent *Science News* article points out that when using the square root of the value for the radius in order to maintain the appropriate relationship between area and value through the scale, it also de-emphasizes the contrast between the small areas and the large ones (Rehmeyer, 2008). While this point is identified as a weakness for Nightingale's original use of the chart, as it could potentially diminish the impact of larger numbers (Rehmeyer, 2008; Willers, 2011; Lewi, 2006), the de-emphasis in another context could be considered a potential strength given that it is theoretically possible to use area encoding to purposely de-emphasize larger data (Willers, 2011). In the case of React, this could support the goal that the student should "*have the experience that their opinion matters and is being counted.*"

The original diagram by Nightingale showed two years of data; by using two side-by-side charts, it was easy to compare the two years as well as look for seasonal impact on the data (Lewi, 2006). This suggests that a potential strength of this chart type might be in using small multiples. Given this information, we can examine the design decisions with some additional information about the potential strengths and weaknesses others have considered for this chart. While this illustrates research on the chart type for the case study, almost all chart types have debatable strengths and weaknesses. Spending time getting familiar with these attributes helps to evaluate whether or not the visual learning analytics are achieving the ETIG.

8.2   Scale and Axis

Key Question: What scale clearly communicates the goal of your visualization?

When establishing the scales and axis of a graph, there is guidance on how to communicate information clearly. While we know that the selection of scales and axis shapes information (Stevens, 1946; Hand, 1996; Velleman & Wilkinson, 1993), we must consider the context under which the data are collected and use the substantive scientific knowledge available to select appropriate scales for our charts (Wilkinson, 2005). What does it mean when learning science is what informs the selection of scale? And what role should quantitative displays play in emotional measures?

When building a functional prototype, we had to decide how to assign values to the chart. Given that the average class size in 2007–08 was 20.0 pupils for public elementary schools (Snyder & Dillow, 2013) by having the upper limit of the range of values at around 40, the design would have ample room to display reactions appropriately. In order to accommodate those upper limits, we used a Fibonacci sequence of twelve numbers 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144 for the respective reactions: engaging, interesting, challenging, curious, calming, good, dull, boring, sad, confusing, frustrating, and annoying. We generated a linear scale (see Figure 7) that would allow us to use Boring at 21, Sad at 34, and Confusing at 55 to gauge the upper limit of responses while Engaging at 1 and Interesting at 1 would represent what it would look like to a student viewing only their own reactions. In connecting back to our educational theory informed goal, we noted that the single response in comparison to the upper bound looked insignificant. This did not satisfy the latter half of the goal that the student should *"have the experience that their opinion matters and is being counted."*
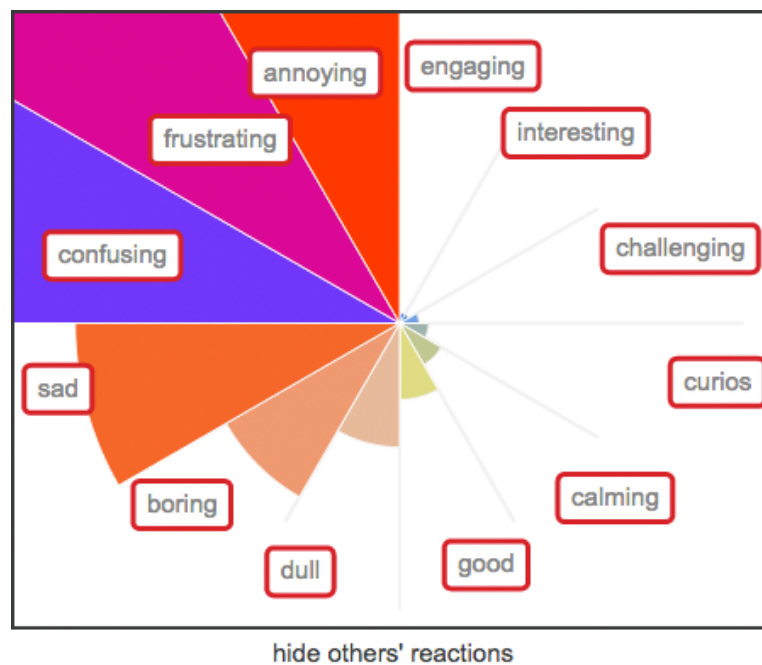

Figure 7. Linear scale for React.

After generating a functional prototype, it became obvious that some transformation of the data would be necessary. We asked the designer to devise a graphical version that showed roughly how one reaction would look when every word was selected and where 5, 20, and 40 reactions would look by outlining the size of the circle with grey outlines. The purpose of this illustration was to demonstrate visually to the end user that their opinion matters and is being counted. The resulting sketch in Figure 8 was used as a guide in creating an approximate scale. We selected 40 as the maximum as our goal was to show only the reactions of other students in the same classroom. The general sense from the design department was that 20 would just touch the labels while 40 would extend beneath the labels. One response looked somewhat inflated and five responses looked to be about twice the radius of one.
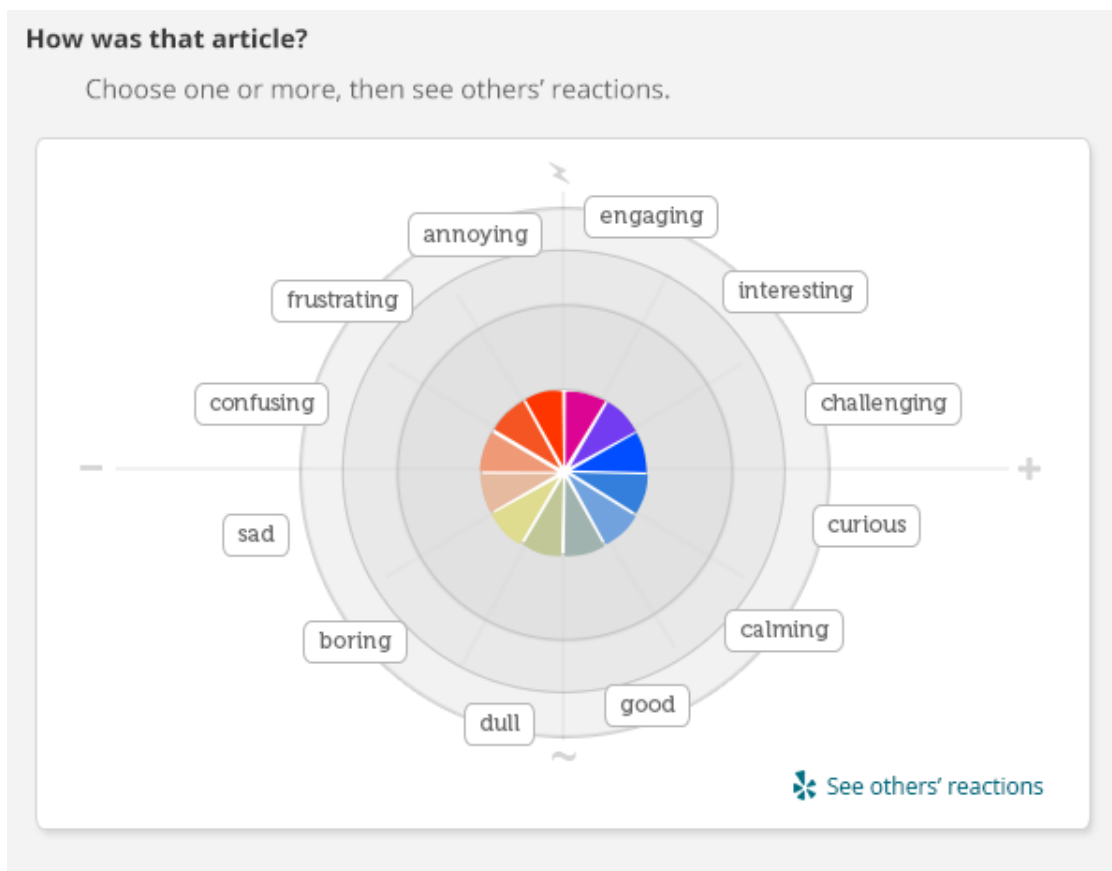

Figure 8. Target scale for React.

Figure 8 also outlines how we decided to treat the axis of the chart. We placed a minus sign on the left, a plus sign on the right a tilde on the bottom, and a lightning bolt on the top. These symbols were intended to convey low energy on the bottom, high energy on the top, negative on the left, and positive on the right. While the concentric circles were used to help establish the scale, they were not intended to be included in the final outcome.

The goals of the design sketch were to make sure that the individual reaction should take up enough screen real estate to indicate to the user that their reactions are important. We wanted to continue to

communicate that each reaction was a part of the overall reactions from classmates. Based on the design sketch, we explored a variety of transformations on the linear scale. Using the same 12 numbers as the initial linear scale, we functionally prototyped three transformations on the linear scale: squaring the value, getting the square root of values, and the log transformation (see Figure 9).
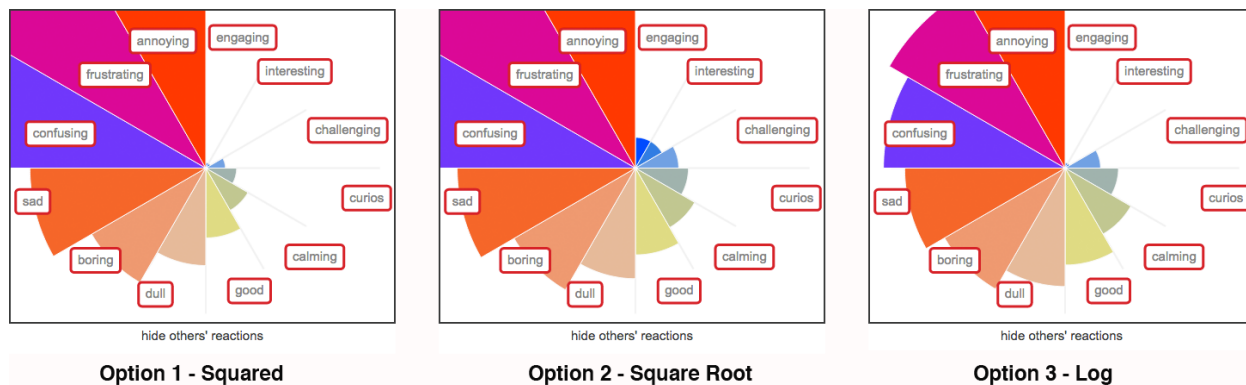

Figure 9. Three options of transformations.

By referencing Boring, Sad, and Confusing as the potential largest values, the log scale looked the most promising and closest to the intended design on the upper end of the scale while the square root appeared closer to the intended design on the lower end of the scale. The major flaw with the Log transformation on the linear scale was that the single response was not visible, so we decided to see what it would look like if we set the minimum value based on some fraction of the total graphic real estate. While the square root transformation produced the most mathematically accurate display of the data set, we wanted to explore our options on how we could shape the display to provide the target experience desired by the design team.

Again, we built a functional prototype of what React would look like with every reaction having a value of 1 using a fraction of the graphic real estate. Figure 10 shows the unit circle at three options: an eighth, a quarter, and half of the graphic real estate of React.
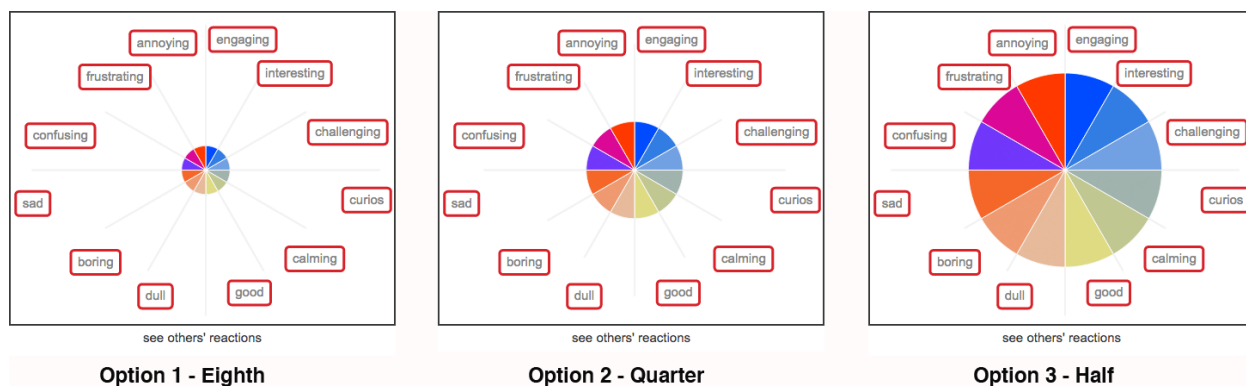

Figure 10. Three options to illustrate radius of "Your Reactions."

When discussing what to use, we wanted to select a representation of an individual reaction that looked meaningful, we decided that the quarter option would make sure that the individual did not feel marginalized by the potential scale of data. The rationale behind doing so was that, given our target population of struggling students, an experience that felt like their contribution was important and meaningful was critical.

When bringing the minimum value of a quarter together with the log transformation on the linear scale, we produced what looked like a satisfactory outcome using the 1,1,2,3,5,8,13,21,34,55,89,144 sequence, as illustrated in Figure 11.
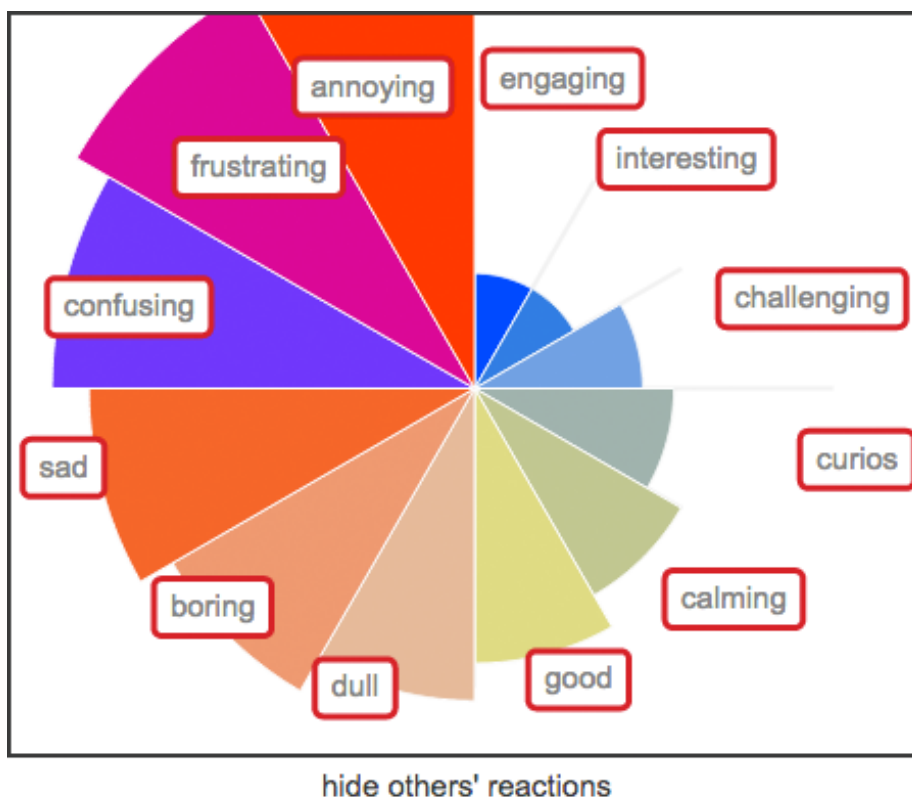

Figure 11. Log with minimum value of a quarter.

This made the single response a closer approximation of the target illustrated in Figure 11 and lined the target values closer to the goal. This clear contrast on selection of scale is a departure from clearly communicating the data, which raises the concern of limiting discovery by the user (Velleman & Wilkinson, 1993). However, it was informed by deliberate choice and, more importantly, represents a theory that we are testing in attempting to convey something that would give the user a general sense of the emotional landscape of reactions people have had toward the reading without marginalizing a single voice. In effect, we are seeking to take the weakness of de-emphasizing the difference between the small areas and the larger areas and pushing that de-emphasis further in order to prevent the sense of a single voice being marginalized in the context of data for the entire classroom. Guided by our

educational theory informed goals, this illustrates how such goals can coordinate an interdisciplinary team.

8.3 Interactive Design

Key Question: Does the interactive design support the goal of your visualization?

After establishing our scale and axis, we could use our mock data set of three users and three articles to generate nine different versions of the chart to get a general sense of what this would look like realistically based on our mock data set. In addition, given that one of the strengths of this chart type is that some consider it good for small multiples, Figure 12 illustrates nine small multiples based on our mock data set.
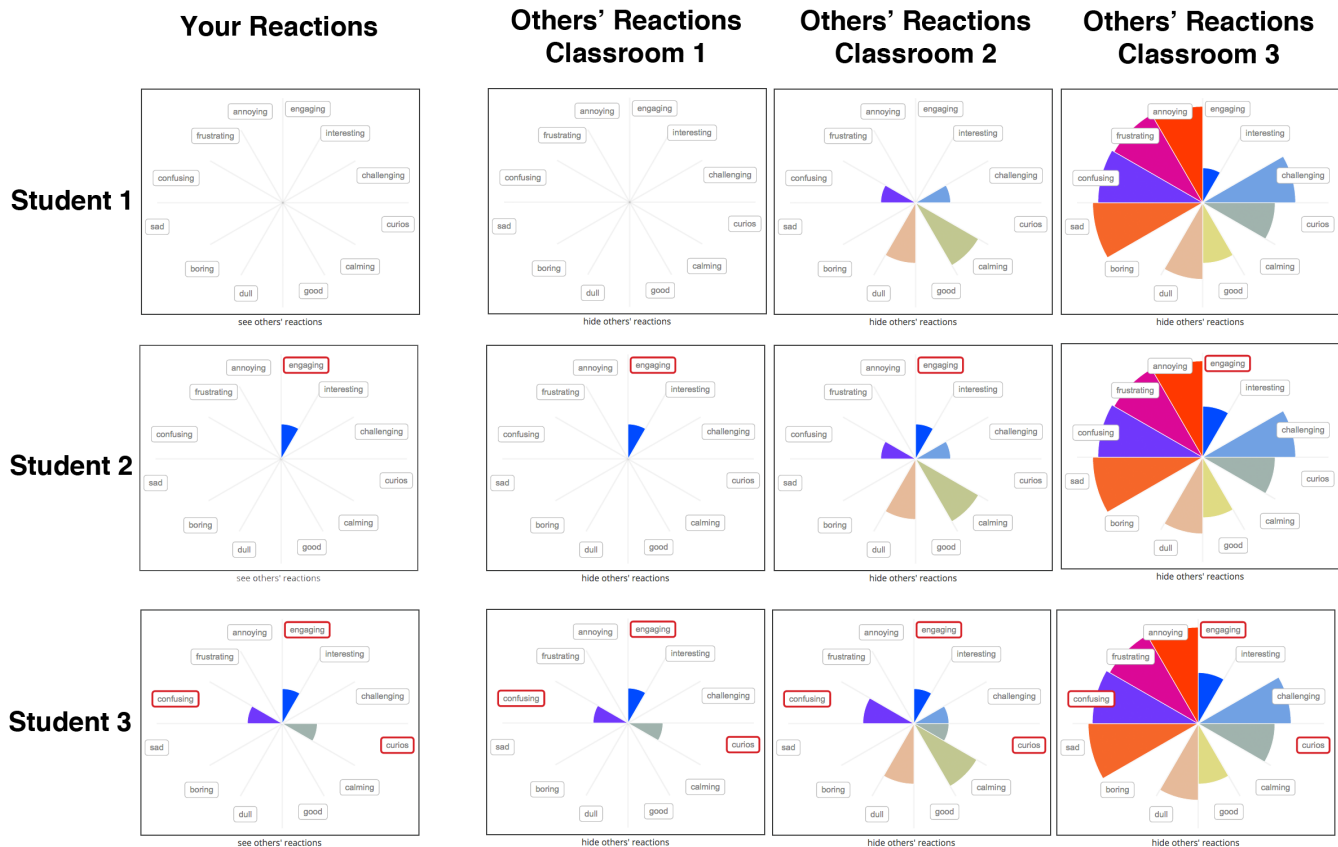

Figure 12. Small multiples of React from mock data.

The purpose of using the mock data set to generate some small multiples is that it gives you a sense of what students will experience. Small multiples allow you to see how your visualization looks in multiple data sets side by side. This clearly depicts the data specifically, but moreover shows many possible views of the visualization look like side by side. In the small multiple form, it is clear that there is no graphical cue indicating that you are viewing an alternate display. While you cannot see the words choices of React here clearly, you can see that no words are selected for Student 1, one word is selected for

Student 2, and three words are selected for Student 3, as indicated by the red border around the word choice, as well as the area chart that indicates that the user made selection. As you scan to the right, you can see what the student would see if they clicked on "see others' reactions." The three potential outcomes are generated from our mock data for Classroom 1, Classroom 2, and Classroom 3. In that comparison, you can already see some examples where the student had similar views to those of the other students and where they had divergent views. This appeared to support the goal of student reflection, as a student could compare their response to those of their peers.

In order to explore the goal of using the data visualization to help peers when exploring content, we explored how it could be incorporated into the explore page where students selected an article to read. The following mock-ups (Figures 13 and 14) have been generated to paper test the potential feature on the explore page of toggling between the cover art of articles and the React visualization of the articles displaying class reactions. Figure 14 illustrates the explore page where a student can select something to read from multiple articles.
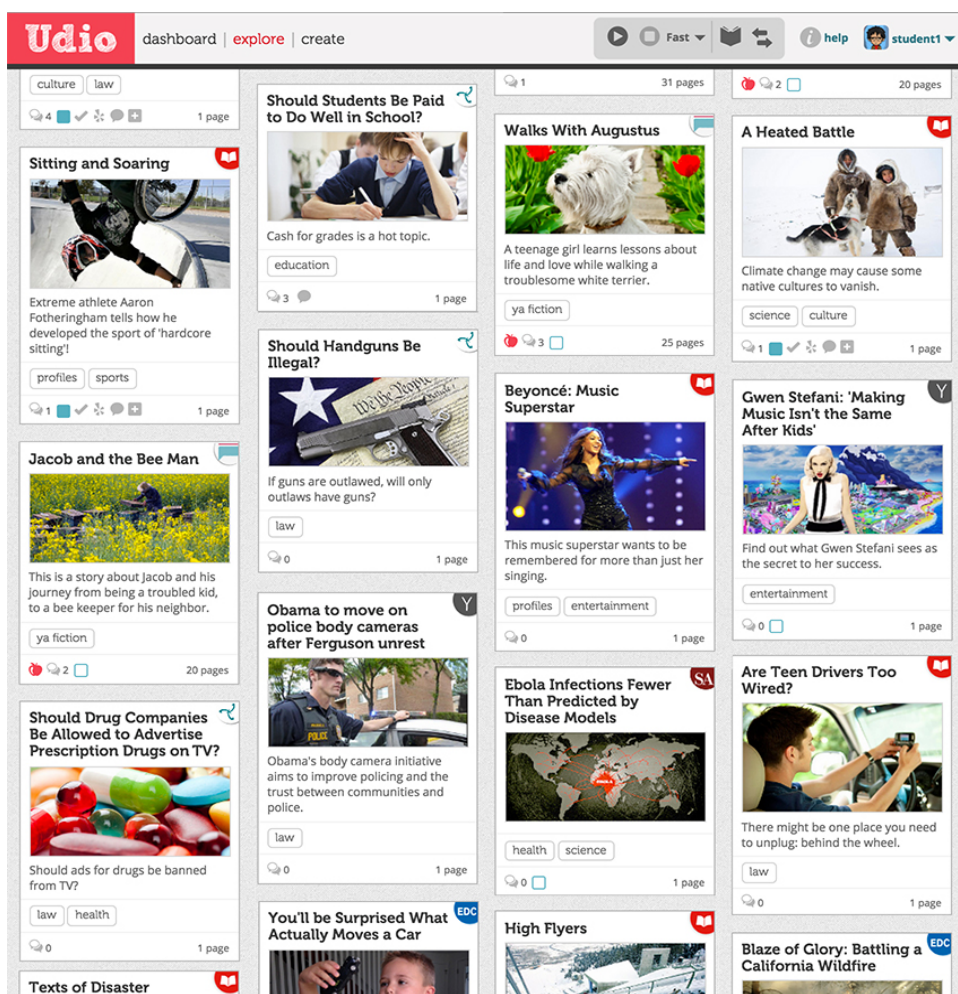

Figure 13. Explore page.

Figure 14 is a depiction of what it might look like on the explore page if a student could see the list of articles where the cover art is replaced with React data visualization that shows the reactions of their classmates to the reading.

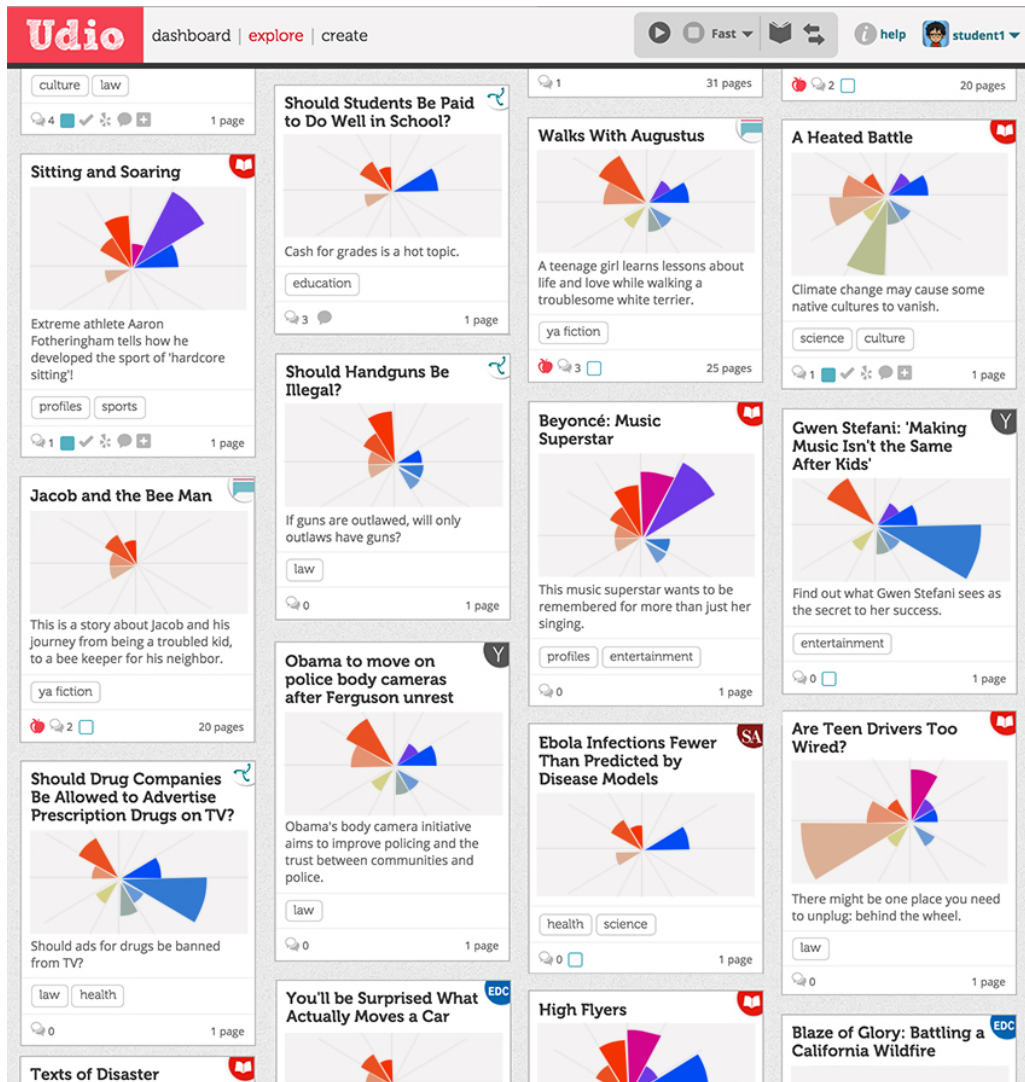**Figure 14. Explore page with React view.**

When taking this out for further formative evaluation we can check with our user population to see if the feature would support students in selecting an article to read. At this phase of the work, this feature has not been tested beyond the generation of the paper prototype.

8.4   Colour

Key Question: How does your choice of colour support your goal?

In early designs, we had tested a rainbow of colours and a monochromatic treatment where size was the only indicator. While these earlier designs were understandable, adding in the dimensions of colour theory and saturation made the visual display more meaningful. We reinforced the relative emotional energies of the words by choosing a spectrum of saturations of colour in conjunction with the relative intensities of the axes.

8.5   Accessibility

Key Question: How does the accessibility of your data visualization clarify or challenge your goal?

8.5.1   Keyboard Accessibility

Accessibility was examined using Google Chrome version 40.0.2214.115 (64-bit). Further testing is needed across browsers. Keyboard accessibility was achieved by making sure that the D3 JavaScript implementation of the buttons that allowed the user to select words for their reaction would respond to both *keydown* and *click* event handlers so that both the keyboard and the mouse could interact with the data visualization. The SVG links were used for the buttons so that they would be included in the tab order. Keyboard accessibility did not clarify or challenge our goal.

8.5.2   Colour-blindness

The goal is for everyone to be able to understand and access the visualization. In general, it is good practice not to have colour as the only indicator of the information. The results must be decipherable for those with colour-blindness. Simulation tools allow you to check for common colour-blind conditions such as Protonopia and Deuteranopia. Photoshop offers a simulation tool through the proof colour feature. The rationale at this stage is that we would like to test the 12 colours we have and see if they help support those with colour vision to interpret the emotional categories of words. Figure 15 shows what React looks like in those two simulations.
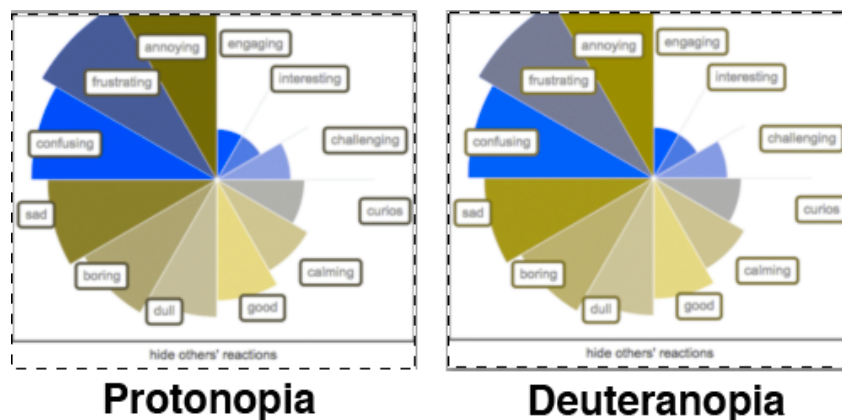

Figure 15. React in Photoshop colour-blindness simulation.

These simulations illustrated in Figure 15 are clearly no longer in 12 distinct colours. This did challenge our understanding of the role of colour in the data visualization as outlined in section 8.4. While colour and saturation were intended to help convey the intensities of the axis, they would not necessarily

convey the same ideas to those with colour blindness. When deciding whether or not we needed to change the colour selection we elected to keep the initial selection. Given that, in section 8.2, the axes were intentionally labelled to convey the same information this was sufficient treatment to support the goal of having a universal mechanism of self-reporting. The alternative decision would have been to use the colour-blind barrier-free palette. Okabe and Kei (2002) suggest using the following eight colours: Black rgb (0,0,0), Orange rgb (230,159,0), Sky Blue rgb (86,180,233), Bluish Green rgb (0,158,115), Yellow rgb (240,228,66), Blue rgb (0,114,178), Vermillion rgb (214,94,0), and Reddish Purple rgb (204,121,167). This simulation shows why multiple representations are critical.

8.5.3 Screen Reader

Accessibility was examined using Google Chrome version 40.0.2214.115 (64-bit) and Mac VoiceOver Utility version 6.0. In the implementation, the buttons were adjusted by changing the class attribute, which highlighted the buttons in red when they were selected. An equivalent experience for the screen reader was achieved through ARIA standards by indicating that each was a toggle button and updating the attributes so that the screen reader could tell if the button was selected. While we had hopes of integrating ARIA-Live tags into this implementation in our initial environment, we were unable to get the buttons to leverage the ARIA-Control feature so that the screen reader did not live update the transition to the See Others' Response toggle. A separate proposal was put together to investigate accessible D3 data transitions.

The most interesting aspect of exploring how to make an equivalent experience with the screen reader related to our implementation around the graphic elements of the visualization. We established a non-standard scale and chose to omit numeric labelling on the axis. It was difficult to determine what an equivalent experience would be from the screen reader as this chart relied on the viewer's interpretation without providing explicit numeric indicators. The easiest alternative text to provide would be the corresponding numeric values for each reaction option, but that seemed more like an equivalent experience of a bar chart displaying the information, not this windrose. Alternatives included the option of describing responses in categorical ways such as "some people reacted Engaging and a lot of people reacted Boring," but this was layering interpretation onto the chart since the screen reader would not allow them to draw their own conclusions, as they might not use the phrases "some" or "a lot" to describe the data visualization. Ultimately, we decided to start with numeric values for alternative text because we were not yet sure if our axis was achieving the intended goal. The next steps will be conducting some further focus groups and testing to better understand how our users are interpreting the graph to see if a possible alternative text could provide an equivalent experience.

9   FUTURE WORK

One of the benefits of organizing your work around an ETIG is that it can help establish how to evaluate the decisions made. This worked example was selected because of the frequent interactions with focus groups and the iterative nature of the pursuit of an ETIG; however, decisions are still made without user feedback.

One design element that requires further evaluation is our choice of scale and axis. First, we need to see if it generated the effect we were seeking. Given that our target user is the struggling reader who has often met with failure in the classroom already, we hoped our design would increase participation with React by making sure that their responses did not appear to have marginal value. We wanted to inflate the value of an individual visually in the hope that users would provide their reactions over the course of their usage of Udio. This means that we would want to see both a high level of use and sustained usage over time. One of the first comparisons we could make without further development is how the React feature looks in comparison to other supports in the system. However, that would be influenced by the novelty of the data visualization. In future pilots, it would be interesting to try an A–B test simply adjusting the scale of React. One comparison could be the square root scale as opposed to our inflated log scale or possibly some linear scale. We could also adjust the lowest value. For instance, keep the log scale but compare the impact of setting a single response to an eighth of the width of React as opposed to the current setting of a quarter. While these strategies could be employed to see if our visualization had the intended impact, we should also test for some potential unintended consequences.

By deviating from the goal of clearly communicating the data set, we know that one potential pitfall is in preventing users from discovering their own insights. Given that we did not put labels on the axis of this chart, it would be interesting to know if students could interpret the chart for numeric value accurately and if they have the desire to know the specific numeric values of the chart or if the visual representation is valuable without knowing the specific numbers. For this purpose, we could conduct interviews with students and use a variety of static images of the React visualization to see if they can interpret numeric values and if they can describe qualitatively the characteristics of the article based on the React visualization.

REFERENCES

Barry, C. (2010, March 10). POV Madlibs [Weblog post]. Retrieved from https://dschool.stanford.edu/groups/k12/wiki/22e39/POV_Madlibs.html

Biancarosa, G., & Snow, C. E. (2006). *Reading next: A vision for action and research in middle and high school literacy: A Report to Carnegie Corporation of New York* (2nd edition). Washington, DC: Alliance for Excellent Education.

Bower, G. (1981). Mood and memory. *American Psychologist*, *36*(2), 129–148.

Brown, T. (2009, September). Change by Design. IDEO. Retrieved from https://www.ideo.com/by-ideo/change-by-design

Brown, T., & Katz, B. (2009**)**. *Change by design: How design thinking can transform organizations and inspire innovatio**n**. New York: HarperCollins.

Brown, T., & Katz, B. (2011), Change by design. *Journal of Product Innovation Management, 28*, 381–383. http://dx.doi.org/10.1111/j.1540-5885.2011.00806.x

Bryan, J., & Bryan, T. (1983). The social life of the learning disabled youngster. In J. D. McKinney & L. Feagans (Eds.), *Current topics in learning disabilities* (Vol. I), (pp.57–85). New York: Ablex.

139

CCAAL (Carnegie Council on Advancing Adolescent Literacy). (2010). *Time to act: An agenda for advancing adolescent literacy for college and career success.* New York: Carnegie Corporation of New York.

Collins, A., Brown, J. S., & Newman, S. (1989). Cognitive apprenticeship: Teaching the craft of reading, writing, and mathematics. L. Resnick (Ed.), *Knowing, learning and instruction: Essays in honor of Robert Glaser* (pp. 453–494). Hillsdale, NJ: Erlbaum.

Economist. (2007, December 19). Worth a thousand words: A good graphic can tell a story, bring a lump to the throat, even change policies. Here are three of history's best. *The Economist*. Retrieved from http://www.economist.com/node/10278643

Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Tackett, K. K., & Schnakenberg, J. W. (2009). A synthesis of reading interventions and effects on reading comprehension outcomes on older struggling readers. *Review of Educational Research*, *79*(1), 262–287. http://dx.doi.org/10.3102/0034654308325998

Ehyde, C. C., & Seneta, E. (2001). *Statisticians of the centuries*. New York: Springer

Eich, E., Kihlstrom, J., Bower, G., Forgas, J., & Niedenthal, P. (2000). *Cognition and emotion*. New York: Oxford University Press.

Evan, B. (2011, May 17). "6-8-5." *Gamestorming*. Retrieved from www.gamestorming.com/games-for-fresh-thinking-and-ideas/6-8-5s

Fischer, K. W., & Bidell, T. R. (1998). Dynamic development of psychological structures in action and thought. In R. M. Lerner (Ed.), *Handbook of child psychology. Vol 1: Theoretical models of human development* (5th Ed.; pp. 467–561). New York: Wiley.

Friendly, M. (2008). A brief history of data visualization. In C. Chen, W. Härdle, & A. Unwin (Eds.), *Handbook of computational statistics: Data visualization III* (pp. 1–34). Heidelberg: Springer.

Frijda, N. (1986). *The emotions*. Cambridge, UK: Cambridge University Press.

Fry, B. (2004). *Computational information design*. PhD dissertation, MIT, Cambridge, MA.

Gajria, M., Jitendra, A. K., Sood, S., & Sacks, G. (2007). Improving comprehension of expository text in students with LD: A research synthesis. *Journal of Learning Disabilities*, *40*(3), 210–225. http://dx.doi.org/10.1177/00222194070400030301

Hand, D. J. (1996). Statistics and the theory of measurement. Journal of the Royal Statistical Society, Series A, 159, 445–492.

Lazarus, R. (1991). *Emotion and adaptation*. New York: Oxford University Press.

Lesaux, N. K, & Siegel, L. S. (2003). The development of reading in children who speak English as a second language (ESL). *Developmental Psychology, 39*(6), 1005–1019. http://dx.doi.org/10.1037/0012-1649.39.6.1005

Lewi, P. J. (2006). Florence Nightingale and polar area diagrams. Speaking of Graphics: An Essay on Graphicacy in Science, Technology and Business. Retrieved from http://www.datascope.be/sog.htm

Maag, J. W., & Behrens, J. T. (1989). Depression and cognitive self-statements of learning disabled and seriously emotionally disturbed adolescents. *The Journal of Special Education*, *23*(1), 17–27. http://dx.doi.org/10.1177/002246698902300103

Margalit, M., & Raviv, A. (1984). LDs' expressions of anxiety in terms of minor somatic complaints. *Journal of Learning Disabilities*, *17*(4), 226–228. http://dx.doi.org/10.1177/002221948401700408

Meyer, A., Rose, D. H., & Gordon, D. (2013). *Universal design for learning: Theory and practice*. Wakefield, MA: National Center on Universal Design for Learning.

Murnane, R. J., & Levy, F. (1996). *Teaching the new basic skills: Principles for educating children to thrive in a changing economy*. New York: Free Press.

NAEP **(**U.S. Department of Education**,** National Center for Education Statistics**).** Retrieved **from** http://www.nationsreportcard.gov/reading_math_2013/files/Results_Appendix_Reading.pdf

Nielsen, J. (2012, June 4). How many test users in a usability study? Norman Nielsen Group. Retrieved from www.nngroup.com/articles/how-many-test-users

Nieveen, N., & Folmer, E. (2013). Formative evaluation in educational design research. *Educational Design Research*, pp. 152–169. Netherlands Institute for Curriculum Development. Retrieved from http://downloads.slo.nl/Documenten/educational-design-research-part-a.pdf

NJCLD (National Joint Committee for Learning Disabilities). (2008). Specific learning disabilities: A national review. New York: National Joint Committee for Learning Disabilities.

NLTS II (National Longitudinal Transition Study II). (2003). National Center for Special Education Research at the Institute of Education Sciences. Washington, DC: U.S. Department of Education.

Okabe, M., & Kei, I. (2002). *Color universal design (CUD): How to make figures and presentations that are friendly to colorblind people*. [Web page] Retrieved from http://jfly.iam.u-tokyo.ac.jp/color

Palincsar, A. S. (1986). Metacognitive strategy instruction. *Exceptional Children*, *53*, 118–124.

Palincsar, A. S. (1998). Social constructivist perspectives on teaching and learning. *Annual Review of Psychology, 49*, 345–375.

Palincsar, A., & Brown, A. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, *1*, 117–175.

Perfetti, C. A., Landi, N., & Oakhill, J. (2005). The acquisition of reading comprehension skill. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 227–247). Oxford, UK: Blackwell.

Rehmeyer, J. (2008). Florence Nightingale: The passionate statistician. *Science News.* Retrieved from https://www.sciencenews.org/article/florence-nightingale-passionate-statistician

Rose, D. H., & Meyer, A. (2002). Teaching every student in the digital age: Universal design for learning. Alexandria, VA: Association for Supervision and Curriculum Development.

Rose, D. H., Meyer, A., & Hitchcock, C. (2005). *The universally designed classroom: Accessible curriculum and digital technologies.* Cambridge, MA: Harvard Education Press.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*, 1161–1178.

Russell, J. A. (2005). Emotion in human consciousness is built on core affect. *Journal of Consciousness Studies, 12*, 26–42.

Russell, J. A. (2009). Emotion, core affect and psychological construction. *Cognition and Emotion, 23*(7), 1259–1283. http://dx.doi.org/10.1080/02699930902809375

Searle, J. R. (1992). *The rediscovery of the mind*. Cambridge, MA: MIT Press.

Serrell, B. (2006). *Judging exhibitions: A framework for assessing excellence*. Walnut Creek, CA: Left Coast Press.

Sideridis, G. D., Mouzaki, A., Simos, P., & Protopapas, A. (2006). Classification of students with reading comprehension difficulties: The roles of motivation, affect, and psychopathology. *Learning Disability Quarterly*, *29*(3) 159–180. http://dx.doi.org/10.2307/30035505

Snyder, T. D., & Dillow, S. A. (2013). Digest of Education Statistics 2012 (NCES 2014-015). National Center for Education Statistics, Institute of Education Sciences, Washington, DC: U.S. Department of Education.

Steidl, S., Mohi-Uddin, S., & Anderson, A. (2006). Effects of emotional arousal on multiple memory systems: Evidence from declarative and procedural learning. *Learning & Memory*, *13*(5), 650–658. http://dx.doi.org/10.1101/lm.324406

Stevens, S. S. (1946). On the theory of scales of measurement. *Science, 103*, 677–680.

Suda, B. (2010). *A practical guide to designing with data*. London: Five Simple Steps.

Sutherland, K. S., & Singh, N. N. (2004). Learned helplessness and students with emotional or behavioral disorders: Deprivation in the classroom. *Behavioral Disorders*, *29*(2), 169–181.

Tufte, E. R. (2001). *The visual display of quantitative information*, 2nd ed. Cheshire, UK: Graphics Press.

Velleman, P. F., & Wilkinson, P. F. (1993). Nominal, ordinal, interval, and ratio typologies are misleading. *American Statistician, 47*, 65–72.

Wagner, M., Marder, C., Blackorby, J., Cameto, R., Newman, L., Levine, P., … Davies-Mercer, E. (2003). *The achievements of youth with disabilities during secondary school: A report from the National Longitudinal Transition Study-2 (NLTS2).* Menlo Park, CA: SRI International. Retrieved from http://www.nlts2.org/reports/2003_11/nlts2_report_2003_11_complete.pdf

Wilkinson, L. (2005). *The grammar of graphics*, 2nd ed. New York: Springer.

Willers, B. (2011). Numbers and area. Retrieved from http://www.benwillers.com/assets/willers_numbers_and_area.pdf