

<meta name="viewport" content="width=device-width; initial-scale=1.0; \>Event ID:
1627170

Event Started: 10/22/2010 5:00:00 PM

>> Please stand by for real time captions.

>> Hi, folks, this is Sheila Campbell. Welcome to this afternoon's webinar. We're thrilled that you could join us this afternoon. We've got a great presentation ahead of us. I would like to introduce Stephen Few, the founder and principal of Perceptual Edge, he will be talking with us today about how to make data meaningful for the average person in terms of creating a great user experience. This is really important to all of us. The open government initiative, one of the pillars is transparency and making data transparent to the public. And so a lot of us are hearing about this, some of us are directly involved with putting up our data sets on data.org. I think we're at a phase where the data is out there, the question is how are creating the data so it means something to the average person? What we're hoping to do is give folks some really good background and education and show you some examples of where we can really do a better job in terms of visualizing that data. I think many of us should work with different people in our agencies, we have data architects and managers, we can come together and create a better experience so the data is meaningful to people in their daily lives.

>> Stephen, we're thrilled you could join us today. Stephen is a leading expert.

He's written a book called "show me the numbers." Welcome, Stephen. Take it away.

>> Thank you, Sheila. I would like to begin with a poem that I think sets the stage for what we want to talk about today. This was written back in 1939 by Millay. Given the fact that she wrote this back in 1939 she wasn't talk being the information age today. Yet I think she expresses the situation that we find ourselves in, living in a day when this is the experience of most people that deal with data. So many of us feel overwhelmed by data. One of the things that the information age has done for us is it's made data readily available, what it has yet to do is make it meaningful for us. For data to be useful we need to be able -- to be able to discover the stories and find ways to tell the stories to others that need that information to make better decisions. One of the ways that I believe in doing that is to present data visually, to use pictures to bring to life the stories. We need to find ways to tap into that data in ways that bring it alive. Pictures are one of the ways to do that. We need to use the right pictures, turn the light bulbs of understanding on. A picture is worth a thousand words, as we all know. And this I think is an example of one such picture. When we use the right picture it really does make the story impactful in a way that other means won't do. When words and numbers fail us pictures often come to the rescue. Not just any picture will do.

>> Here's an example of a picture that is worth a thousand words. This map is a data visualization. It was showing circumstances during an epidemic in a section of London. A medical doctor created this map. During this epidemic he had a theory that cholera was spread in a different way for that time. He believed that it was the case that it was spread through the water supply. To test that he went door-to-door to find out where people had died during the epidemic. On this map every one of the little dots that you see represents a place where someone had died. Then he also identified the locations of the wells where people got their water, such as the one I have marked here. And then looking at the relationship between the deaths and the various pumps he found that there was a pump on Broad Street that seemed to be at the center of the distribution of deaths in the area. He wondered about some of the deaths that were far away from that pump, he actually went and investigated situations like some of those dots that are far away. He found stories like Mrs. Smith used to live on Broad Street, but she moved to a further location. She didn't like the taste of the water close to her, she would have her sons bring her buckets full of water once or a twice a week. He was able to make a very strong case using this picture, which was really necessary to bring this story alive. He persuaded the powers that be in the neighborhood to remove the handle from the well and within just a short time the epidemic subsided. Here's a case where a picture really was effective. But not all pictures are.

>> The example that I'm showing you right now is a picture that appeared in Bloomberg not long ago. It told the story of the results of the recent financial melt down in

relation to banking assets. It was comparing the losses of various banking institutions. In looking at this picture it is probably not at all obvious to you that the purpose of this particular article in Bloomberg was to point out that JP Morgan suffered a loss that was minor compared to most of the banking institutions out there. The only other that suffer the less was Santander. Comparing the circles is not something that people do well. This didn't tell the story well. I went in and redid this picture. This is the beginning, rather than using circles I'm using bars. We can very easily compare the heights of the bars see the differences from Quarter 2 to January 20 on the right-hand side. We can see that JP Morgan suffered a loss of about half of its assets. The part that is still not clear is the fact that their loss was second in terms of the least amount compared to Santander. I'm adding this lower graph now where the percentage of loss is presented directly here. Very clearly we can see that JP Morgan lost loss.

>> Now we've been telling quantitative stories for a long time. The real pioneer is William Playfair, a Scottish gentleman. He was significant. He was the first person to use a line to show how values change through time. He invented the bar chart that we're all familiar with. He invented the pie chart and so on. We owe a lot to him in terms of popularizing the use of graphics to display information.

>> Here we are today, graphs have been commonplace in our world. Yet the sad truth is that most of the graphs that are used today actually don't work as well as the pioneering efforts of Playfair long ago. Here's a common example today, fox news pie charts that adds up to 193%. We've lost the skills that are required to do this well along the way.

>> Let me tell you a story. I want you to imagine that you have just walked into a conference room for the last meeting of the week. You sit down, it's a Friday afternoon at 3:00, you sit down and about five minutes after the hour a young fellow jumps up and says "I am so glad that all here today, all of you IT managers, I'm so thrilled because I'm going to introduce to you today a new report that you will find sitting on your desk as you arrive every morning that will tell you what is going on in the systems that you manage, you will be receiving data in near real time, you can receive instant insight and go out and take effective action immediately." Here it is, the new report. At about this moment you notice that you are the only person in the room who is not nodding enthusiastically. What can they see that I can't see? I just can't see anything meaningful in this data. Rest assured that you are not alone. This graph doesn't say anything of use. It doesn't say anything in a way that really clarifies what is going on in these systems. I did not invent this example to make this point today. This is an example of a graph that was distributed at my last place of corporate employment. People did get excited as if they were getting something useful. As you can tell, they were getting something that couldn't be used in any meaningful fashion. I think this illustrates the sad state of affairs.

>> At this time I became aware of the fact that we do have a problem. I looked for resources at the time to teach my staff, I managed a large staff of professionals whose full-time work was analyzing data and building reports. I realized that nobody in that team had the skills to be able to communicate data graphically. I looked for the resource to teach them, I could find some books out there about the design of graphs, but I didn't find anything that communicated these principles in a practical way and was accessible. Because I couldn't find the necessary resources I decided to be the person to develop the resources. I quit my job and wrote my first book "show me the numbers" which was designed to do what I just explained. To teach very simple skills in a practical way that is access to a broad audience to help them display quantitative data. So the stories and data come through in a clear and compelling fashion.

>> Most of the time we use a table or a graph or some combination of the two. Most of those that are produced in the workplace today do not communicate effectively. In the year that I was born this book was published, back in 1954, "how to lie with statistics." He had a concern that people were using data to mislead, to disreprove. That problem still exists today. But I think that today an even greater problem exists in that many of us with good intentions, with the intention of telling the truth with data end up misleading people because we don't know how to do this affectively. We've never learned the skills for doing this well. When I was first involved in the computer industry if you -- which was back in the 1980s, before the proliferation of the PC if you wanted to create a graph that you might be using in a meeting you had to get out your special triangles and T squares and labor over the

creation. If you were putting that much time into it you were probably taking the time to learn how to do it well. But today now with the ready availability of the PC we began to believe that if we knew how to use software that was designed to do something that we automatically knew how to do that thing. And so if we had a copy of Excel we automatically knew how to present data, but that's no more true than if we have a copy of Word that we know how to write effectively. There are skills that we need to bring to this process to use those tools. Because most of us have not learned these skills we tend to rely on the lead of software vendors to figure out how to display data. We end up creating horrible displays, like this example right here.

>> It is a very, very simple message, but it's almost impossible to get that simple message from the data here. What many vendors are encouraging us to do is to dress up our displays, to decorate them. But they obscure the messages that live in the data. We need to learn the basic principles and avoid the directions that some vendors tend to be leading us into.

>> Now I have talked about the story of how the PC is responsible for much of this, we want to unlearn some of these patterns, some of these techniques that the PC has been leading us to develop. I want to make my point here by showing you examples. This is a typical example of a report that you might see in the workplace. Let's say the purpose is for us to compare how our democratic is doing compared to the other departments out there in the organization. Can you do that with this graph? You might look at it and realize with a bit of work I can make that comparison. But if you had been the person that created this you might realize this is not compelling enough, what can I do to better present this? If you follow the lead of the software products you might want to make this change to it. Which, of course, is accomplishing the opposite. By making the pie chart 3D and rotating it we've now made it much more difficult to get the simple story. Now because of the 3D, department D is much more weighty, the angle is causing it be much more difficult to compare the slices of the pie. Without any struggle here we can see where our department falls in comparison to the other departments out there. We're second in line, just slightly higher in expenses than department A. We're looking for simple and clear pictures, nothing fancy. Simple and clear pictures that draw people into the meanings that live in the data.

>> Let's look at another. I pulled this off of the website of PBS. This was from a program called "now," which used to be hosted by Bill Moyers. This is important information, this is from 2004. This shows how various countries in the world feel about the United States. It's very difficult by looking at this presentation to get a sense of how the world feels about America. I re-created this display in this form. Now quickly we can get a picture of how these countries feel about the United States. The green bars to the left are showing us favorable opinions of the U.S. The red bars are showing us the unfavorable opinions. To the right of that is the gray, which is undecided and neutral. We quickly get a sense, we also can spot details that might not have ever jumped out in that tabular display. For example, look at the country of Jordan, third up from the bottom. Few people at the time felt favorably. We can see looking at the gray bars that nobody in Jordan was without an opinion. That's a fact that pops out to us in this display, but was hidden in the previous display. This is what we're looking for, ways to bring these stories to life.

>> Here's another example. This one I pulled from a government website, this is from the state of Maine looking at healthcare data. There's a comparison between AIDS diagnosis and deaths related to AIDS across several years. Here we're looking at change through time, displayed in this way it's much harder to see the pattern of how these things are changing through time. Yet here the message is simplified. Our attention is drawn to the patterns of change in a way that makes it very easy to compare the patterns of change of diagnosis versus deaths due to AIDS across these years. Another government example here.

>> Here we're supposed to be using these two charts to compare H.I.V. diagnosis for males on the left and females on the right during a period of time. We're also supposed to compare the methods of transmission. As you can see that would be very, very difficult to do. You would have to spend a lot of time making those comparisons. Whereas when displayed in this fashion we can easily compare the modes of transmission within males and within females and by comparing the bars across the two charts we can compare those modes between males and females. Very, very simple

by making the data accessible to people. Our eyes can do most of the work in making the comparisons.

>> What are the fundamental challenges that we face? I believe that it boils down to two. The first is determining the medium of expression that will tell the story best. At the very most fundamental level it boils down to a choice of using a table or a graph or some combination of the two. If you decide to use a graph, what kind of graph? A bar chart, line chart, a scatterplot? That's determining the medium of display. Once we've made that decision then we have to design that display, the components, in a way that really do feature the story clearly. How that often starts is we have to undo some of the horrible defaults of the software to simplify the presentation. Let's say in this particular example I want to feature what is happening in the west region. I might do something to make that region stand out, maybe making the green bars much brighter to focus attention on that aspect of the story. The two challenges are determining the medium of display and designing it in a way that featured the story clearly and accurately.

>> Now talking about determining the medium for a moment, let's talk about the strengths of tables. Tables are marvelous devices for doing a couple of things. First of all, for giving people an easy way of looking up individual values. If I want to look up the value in the month of July for the year 1996 I can easily find that when the data is displayed in a table like this. The other strength is in tables we can express the data precisely. When people need a high degree of precision we can do that on a table. Those are the basic things that tables do well. Beyond that tables are not generally all that useful.

>> Let's talk about the strengths of graphs. Graphs feature patterns, trends, and exceptions in the data and make it possible for us to examine series of values at once. In this example we're seeing the series of values for contributions domestically versus international. We can compare the lines to one another in a way we could have never have done with the table previously. We can see such facts as that in the month of July international sales took a significant -- not sales, but contributions took a significant dip. In domestic contributions we can see that contributions trended upward. There was a pattern in the contributions, up, up, down, up, up, down. It's repeating quarterly. We've seen exemptions in the data, we've seen trends and this repeating pattern. We can also compare the pattern of contributions domestically to international contributions. All of that is accessible to our eyes. This is the great strength of graphical displays.

>> Now just to illustrate that, looking at this example, which I pulled off of a website, we're looking at populations by age. Can we through time compare these patterns and see how age distribution across the population is changing by looking at this table? There's no way we can construct that picture of the distribution of ages in our heads. But if we display it graphically, using this line graph, we can see the patterns and the changes in those patterns of age distribution across the population quite easily. So this is the great strength of graphs, to bring these stories that exist alive and available to our eyes.

>> Just to illustrate further this point, tables don't make those kinds of features available to us, here's a very simple table. It's a very small table. It has eight values. These number represents employee job satisfaction. The higher the number the more satisfied the employees are. Employees are broken into groups, two rows that break them down into income levels. We also see columns that break employees into college and noncollege degrees. And also age categories. Yet just eight values, there's not a lot of data. There's a pattern here that is not jumping out. There's a group whose job satisfaction is quite different. Yet you are probably not seeing that particular pattern. But if I show you this information presented graphically here.

>> In this case I'm using lines to show you the data, you can see there's that one group of people, the black line, that has a slope going downwards, that is the people with no college degrees that make over \$50,000. When they're reaching the older age category their job satisfaction is taking a dive. I just want to point out this is not real data. I did not base this example on actual data out there. If you don't have a college degree and you make a lot of money and you are about ready to turn 50, don't worry. You see the point that I'm making here. Even in a small table that pattern which is so easy to see here was not accessible to our eyes before. It's not just because I used a graph.

>> Here we see in the upper left the graph, in the bottom right we see another graph

of the same data, yet that fact that was so clear in the line graph is no longer as easily accessible. So it's not just a matter of displaying the data graphically, it's a matter of doing it in a way that really features the aspects of the story that we're trying to tell people.

>> Alluded to the fact earlier that a lot of the software that we use tends to mislead us into bad practices. Here I have an example of a graph where we're comparing actual expenses across 12 months of the year to the budget. All I have done is used some of the features that are available in Excel, in the 2007 version and beyond. I hope in looking at this you can see that when displayed in this way it's much more difficult than it ought to be able to make the comparisons and see the patterns here that we want to see. Let's improve this display, just take it one step at a time to design it in a way that communicates better.

>> To start the 3D aspect of this graph, I hope you see there's no reason to do it that way. It's making the comparisons much harder. Let's get rid of that. Let's have the bars sitting side by side. At this point I hope you also see that the background color is actually drawing our eyes to something that has nothing to do with the data, it's distracting. Let's get rid of that. Okay. Now with that gone it's much easier to focus on the data. Now looking at the bars of actual to budget, notice how much more difficult it is to compare the heights. These spires are making the part of the bar that is most important, the top of the bar, the hardest to see, which is the opposite of what we want. Let's change it to just a standard bar chart. Now we can easily compare the heights of those bars. Let's shift to things other than the bars that represent the data. Let's notice the fact that along the X axis and along the Y we have tick marks. We can see the months without them, they're unnecessary. The ones along the Y axis are not necessary at all. We have grid lines in the graph itself, we can get rid of them. We're removing content that doesn't do anything for us. That is simplifying the display. Now looking again where the dollar values are, notice that in a case like this where maybe you are presenting this in a meeting we want to make sure that the values are expressed large enough so people can see them easily. Let's increase the size of the font to make it easier to read. Let's a level of precision now. Point zero zero, this is useless to us. Let's get rid of the cents. Now also notice that the dollar signs is redundant. Let's just get rid of them and just say that once, U.S. dollars. Let's orient that right up above the axis and make it very easy to read. Notice that the legend on the right-hand side, the items are one on top of the others. The bars in the graph are side by side. Let's arrange the legend to make the connection easier to make.

>> When looking at actual and comparing it to budget we care more about the actual values than the budget. The purpose of the budget values are a point of comparison. Yet the red bars are really standing out in a way that is greater than is necessary. Let's cause those bars to not stand out so much. Let's change the color to one that doesn't pop as much. I think it's still standing out too much. Let's replace the bars with just little symbols for the budget figures. Now eyes are drawn to the actual expense figures with the blue bars, we can still make the comparison to the budget figures just as easily as before. This is changed through time, by using bars we're making that pattern of change through time a lot harder to see than it would be than if it were lines. Here we have a blue line for actual, gray for budget. Now notice that because actual expenses and the budget for expenses are pretty close to each other throughout the year the differences in the patterns between them is not as easy to see. They're not spread across enough space. We can spread the values by narrowing the scale. We can start just a little bit below the low estival in the data -- lowest value in the data set. Now it's across more space, it's easier to compare them in detail now.

>> Okay. The next thing is remember when we changed the legend that worked for the bars, it doesn't work that well for the lines. Rather than having a legend let's just label the lines directly, which is what I have done here. That's taking some of the complexity out of the process. Now I'm wondering if people print this on a black and white printer if they will still be able to see the difference between the actual -- easily be able to see the difference between the two lines? Let's switch to colors that I know will look different even if you print it in black and white, which is what I have done here.

>> Let's consider how people are using the data. I happen to know the folks that need this report what they care most about the variance between the two. It's hard to see that by comparing the two lines. Let's present the data, the difference,

directly. Which is what I have done here. Here that single line is showing us the variance in terms of the actual dollars. As I think about it, there's one last change, I happen to know this group rather than seeing the dollar difference is more concerned with the percentage difference. Let's make one last change to express the differences between actual and budgeted expenses as percentage difference. Here we started off with a graph that looked like the one up here in the left corner, and through a series of changes made a graph that did not present the data affectively work quite well now. It's makes it very, very easy to understand. The good news is that all of those steps, there were about 15 to get to this final graph, you don't have to start with a bad graph and then go through a series of steps to get to a good graph that works well. If you know what you are doing you can get to that final graph immediately. That's what I think is important, to develop the basic skills so you can get to these graphs that work well without a lost extra work.

>> -- a lot of extra work.

>> When we display information graphically I believe that it really pairs down to about six relationships that we are featuring in the data, that is relationships between the values. If we know what those relationships are and how to best display them we have a nice set of principles to follow. I want to take you quickly through these six basic relationships that we display in graphs and identify them and show some examples of how we can display these affectively.

>> Starting with this relationship. The relationship between these values is how these values are changing through time. That is how they're going up and down through time. I call this a time series relationship. It's a very common relationship that we often use and feature in graphs. Now here is an example of a graph I found on the web that did not feature that time series relationship well. Bars don't do a particularly good job here. If I redo it here. You can see that the pattern of change is much more readily accessible if we express it as a line rather than a series of bars. That's one of the basic rules to keep in mind for time series values. To show the patterns and trends use lines rather than bars.

>> Here's the second relationship. In this case we're looking at the head count, the number of people in these departments. The arrangement of this is being used to show a particular relationship between these values. If you notice the arrangement is from the highest value to the low estival. And one of the benefits of arranging these values by size is we are making it easier to see the differences between them. When the values are closest to one another are arranged that way it makes the comparison much more easy to make. I can call this relationship a ranking relationship. We're featuring the rank order of values. Here's an example of a graph that I re-created from some data I saw for the state of Utah regarding trauma injuries county by county. Here by displaying the data alphabetically it's hard to make the comparisons. By sorting them in order, in this case from the greatest injuries to the least, now the comparisons are easy to make. These ranking displays are often quite useful.

>> Now here's the next relationship. These bars are representing the expenses across four regions, they're in order from the highest value to the low estival. There's something more going on here. If you notice if you add these values up they total to 100%, 100% of total expenses. This is a part to whole relationship. We're looking at the parts of something and how they compare to one another. I did not use the type of graph that is used generally for type to whole relationships, a pie chart. Pie charts are really hard to read. When you look at a pie chart you know you are looking at parts of a whole, but when you try to compare the parts it's difficult. Visual perception is not well tuned to compare areas well. If I take this data and display it as bars now all of a sudden those comparisons are easy to make, by using a bar chart. This is how I almost all display part to whole relationships. It makes the comparison easily. I recommend that you save the pies for dessert. This is the one pie chart their think does make sense. It's about a pie. Otherwise pie charts tend to not work well.

>> What I'm featuring here is a display that is focusing on the difference between two sets of data. In this particular case the difference in terms of the actual head count compared to the plan or budget for head count. We see differences here. When we're featuring the differences I call this a deviation display, we're looking at the deviation between multiple sets of values. Often we look at displays like this. Here's a case where in a graph that I pulled from the web the purpose is to see the difference between the median annual household in Utah compared to the U.S. as a

whole. If we really care about the difference between the two and how that has changed through time it's hard to construct that picture by comparing the two lines. In this case it would work better if I displayed the difference between the two. Don't force people to try to construct a picture in their heads from the two lines when we can show them what they need directly. That was a deviation relationship.

>> Here's the next one. In this case what we're looking at is we're looking at the number of orders that have come into a company broken down by size. On the left-hand side we have orders for under \$10, the next bar is \$10 to \$20. The nature of the relationship that is featured here is how these values are spread across the range, where they fall, across this range. The number that I give to this display is a distribution display. We're showing how a set of values are distributed across the range. Sometimes we might do this by age groups or by grade levels. We're counting things across a series of intervals across a quantitative range.

>> Here's an example of a simple display of comparing children across various grade levels in terms of how overweight they are. But displayed in this way notice how hard it is to focus on the shape of the distribution for boys or girls or to compare them. All we can do here is within a grade level compare boys and girls. I can simplify it here by separating boys from girls and putting them one above the other. Or I can combine both distributions in a single graph and use lines to show the distributions, I make that comparison of the shape and also of the magnitude of percentages across boys and girls. Very, very easy to see and compare. That's a distribution relationship.

>> Let's go on to the last. We're looking at the relationship between how tall people are in inches and how much money they make. The nature of the relationship here, as one set of values change does the other set of values change? This is a correlation relationship. This is something we need to do and present to people occasionally. Here's a scatterplot that tells an interesting and important story. This one is based on data from the World Health Organization by country. The organization here is how literate different countries are versus fertility rate in these countries. We can see there does appear to be a relationship between these two variables. As literacy increases the fertility rate seems to decrease. There's a negative correlation. In this particular display we can see that almost all of the countries with high fertility rates were in Africa. There is one exception, the country of Yemen.

>> We looked at all six of these relationships. I hope that you can see just through those examples that there are best ways of displaying these various relationships. I just touched on it. Of course, there's a lot more to learn on this. I'm hoping to wet your appetite on this.

>> Something that's become popular is this thing called a dashboard in recent years. I help organizations with these. I want to spend time talking about dashboard displays. I want to point out that most of the dashboards out there do a miserable job of displaying data in a way that people can easily make use of. The problem is a problem of design. Most dashboards say too little. What they do say they tend to say very, very poorly. Which is a shame, these can be quite rich and telling stories that exist in the data in a powerful way. One of the problems that we're experiencing out there is that dashboards as they're promoted by software vendors are being misused. People are encouraged to dress up the data with decoration and so forth, it obscures the message. Dashboards are not for show, they're for communicating data. Looking at this example, seeing the lighting effect of that reflection of light on the pie chart, that reflection makes it look like a third slice to the pie, that's not serving a purpose. I was out on a -- if I were outside and I look on the road and there's a street sign that I'm trying to read but the light from the sun is reflecting off of it into my eyes I would consider that very annoying, it's getting in the way of seeing the information that I need. Why would I want to duplicate that on a dashboard? That's doing exactly the opposite of what we want to achieve.

>> Let me give you a definition before I go further. The definition that I use for a dashboard is: It's a visual display of the most important information needed to achieve one or more objectives that has been consolidated on a single computer screen so it can be monitored and understood at a glance. In other words a person or a group of people who share common responsibility has a job to do, they need to be aware of what is going on, they need to get a certain set of information that keeps them aware of what is going on. They're using this single screen display to quickly

and efficiently monitor what is going on, to maintain situational awareness. If I were to focus on a particular part of this definition I would put the emphasis on the word "monitored." They help us monitor what is going on. I emphasize that because the term dashboard means a lot of different things to different people. There's different reasons to combine information on a screen.

>> Now the fundamental challenge is to be able to take this somewhat large set of information that somebody needs to monitor and display it on a single screen in a way that doesn't end up looking like a cluttered mess. If you don't think that's challenging to do, you probably having tried. To do that effectively it is quite challenging. I am really glad that the folks that designed cockpit displays for these large airplanes that we fly in have taken a lot of time to understand the information that people need to monitor while they're flying, and how to display it in a way where they can rapidly assess the situation. Most dashboards for monitoring our organizations have not had that kind of care put into the design, which is unfortunate. We're trying to do important work, they deserve a lot of care in how the data is displayed.

>> Now to teach people the basic principles for how to do this I wrote "information dashboard design." It teaches simple principles for how to display information on a screen at once in a way that people can easily and efficiently use it for monitoring what is going on.

>> Visual monitoring needs to be supported through displays that make it possible to do the following: It should be possible to scan the information of what is going on quickly to get the big picture of what is going on. It should show a lot of information, we can rapidly scan and understand it. Within that big picture it should draw your eyes to particular things that look like they might need attention. About each of those specific things that need attention on the dashboard itself we ought to give people enough information about what is going on to allow them to make a decision of whether they need to respond or not. If they do need to take action we ought to make it easy for them to get to supporting information that they need to figure out how to respond. When I design dashboards this is the model that I use and try to support this kind of monitoring.

>> Now we want to display the data in a way that takes advantage of the power of people's eyes. Visual perception is by far our strongest sense. 70% of the sense receptors in the body reside in the retinas of our eyes. We want to take advantage of that in the way that we present data. If we present it just as text that involves a part of the brain that handles verbal process. That's a slow serial process. If I can display it in graphical form I take advantage of the visual processing part, I can quickly scan these values and get a rapid sense of what is going on. I'm taking advantage of my eyes.

>> Now I want to go through what I consider to be some of the common mistakes in dashboard design, in my book I talk about 13 common mistakes. I will limit that today to six. We don't have a whole lot of time. I do want to touch on some of the common problems that I find. The first one, exceeding the boundaries of a single screen. If you notice on this example in the bottom right you can see a scroll bar, there's more information located off the bottom of the screen, to see that you have to scroll down to that, which will cause information on the top to disappear. In the upper right you see a section that begins with number of transactions and profit, there's a section of key metrics there, there's several other metrics that I can't see right now. If I want to see all of them and make connections or comparisons I have to scroll down. I'm having to rely on working memory to work with all of this information, which doesn't work very well. Working memory can only hold about three chunks of information at a time. There's very little, we're limited in our capacity of making comparisons between things. The goal is to put everything that people ought to connect in front of their eyes at once. You want to avoid forcing people to have to scroll around. The other way this problem exists is forcing people to have to go to separate screens to get all of the information that they need. Here is a dashboard that has 10 different screens that a banking executive would have to look at to get all of the information that he or she needs to monitor. I worked in the banking world for ten years, I happen to know that a banking executive wants to make comparisons between the categories. There are connections that need to be made between these pieces of information that can't be made if people have to look at each piece of information separately. What you are seeing here is an example of what I call lazy man's dashboard design. You want to display all of this information, you

just break it up into tiny chunks and maybe throw a couple gadgets in.

>> Let's look at the second problem, supplying inadequate context for the data. These gauges have become the popular form of display, which is really unfortunate. It's taking this metaphor too far. This is a computer screen that we're using, we can display the information in much more effective ways. They tend to say very little. We've sold 7822 units year to date. Green is good, red is bad. Hopefully you are not color blind. The tick marks that give us a sense of scale don't mean anything on this gauge. Good in what sense is this? Compared to target? Compared to last year on this date? It's good today, but was it good yesterday? Are we getting better? Is it worse? None of that context is there to help us interpret how year to date unit sales is doing. All we're getting from this display I could display by saying YTD units and giving you a green number, you would get the same information. I'm not suggesting that's how we want to display this. We want to find ways of displaying a much richer context around the data than what these gauges display. That's the second problem.

>> Here's the third, choosing inappropriate display medium. Here's a case where a portion of the dashboard, we're comparing the number of sales leads that have come in of a various qualities. I'm guessing that looking at this it looks like to you that everybody is about the same. That's far from the truth. The smallest value here is .01%, it goes up 21.5%. But displayed in this way all of a sudden that variation that was impossible to see is very, very easy to see in this bar chart. We want to find the medium that is most appropriate for it and stick with it. That goes back to before, knowing how to choose appropriate medium for different relationships in the data.

>> Onward to problem four, ineffectively highlighting what is important. Where are your eyes drawn? That's an important question. We want to draw eyes to what needs the most attention. I'm guessing your eyes are all over the place because everything about this display visually is shouting at us. Everything is busy. Everything is shouting at us. If everything is shouting nothing in particular is going to be heard. We want to find ways to cause the things that need the most attention to really pop out so that our eyes don't have to struggle to find them. That's the fourth problem.

>> Here's the fifth out of the six we will talk about today. Cluttering the dashboard with useless decoration. This reminds me of a sound mixing board that I used to use many, many years ago. We have these LED light meters, and so forth. But this is not a sound mixing board. We don't need to make it look like some piece of equipment out there in the real world. By making it look like a piece of equipment we've made it harder to see the values, to interpret the values and to compare the values. Why did I use a sound mixing board 30 years ago? That was the best that a piece of equipment like that could do at that point in time. Here we can display data in much richer ways than this. We don't want to get caught up in this silly decoration. There's no room for decoration on a dashboard.

>> The last problem I want to mention today is misusing or overusing color. Because dashboards are visual, graphical in form, people tend to think they need to throw lots of color on to them. If you look at this particular chart you can see that we're comparing average load capacity across flights from LA international to various other airports. Every bar has a different color, they don't mean anything. We don't need colors. Here we're using color when not needed. When our eyes spot visual differences our brains look for meanings. Here's a case where our brains are wasting time looking for meanings. If we use too much color on a dashboard it will become visually overwhelming. One of the things to consider when displaying data is some people cannot differentiate particular colors. 10% of men and 1% of women are color blind. You have the inability to distinguish red and green. As you see in this example if you look at the right that's how the colors red and green would look to someone who is color blind. It would be impossible for them to distinguish. We need to understand some of those limitations in the use of color. Here in this example I'm illustrating the fact that if we use color to much we undermine the use of power. In the upper left the red icons to the left of metrics that need attention, because something is going wrong, those are not popping out because there's too much color used elsewhere on the dashboard. Yet those red icons in the bottom don't even require conscious attention. We can scan and spot them using the visual processing parts of our brain. Those are a couple of lessons regarding the use of color.

>> What we want to do is actually do something that I will illustrate here. If I

were to ask you which photo of Jessica you prefer to look at, you would answer the one on the left. If you are Jessica's dermatologist and it's your job to diagnose problems with her skin based on looking at photos which photo would you use? You would all say the one on the right. With we present data to people in dashboards or any form people are operating like doctors. They've trying to see the stories, the truth that lives in the data. We cannot hide a layer of makeup. We want to tell that story's plainly and simply as possible.

>> What we want to achieve in the design of these displays is what I call eloquence through simplicity. Eloquence of communication through simplicity of design. We want to borrow lessons from artists that know that just a simple marks on a page they can communicate a lot of information. We want to learn how to pare it down. Here's an example of a dashboard that actually does pare the message down to its essence and displays a lot of information that for someone who is trying to monitor the sales process and is an expert in monitoring sales would find this to be enlightening. Not struggling at all to find those things that need their attention. Not struggling to see the organization. For each piece of information, getting enough information about what is going on to make decisions, do I need to respond? Or do I not need to respond in this moment? This is what we're shooting for in a well-designed dashboard.

>> This brings us to the end of the content I wanted to present. I want to leave you with this basic note, you have a choice. If you have data to pass on to people, stories that live in data that you want to tell to people, you can tell it in ways that are effective and compelling. The skills that are required to do this are very easy to learn. There's nothing complicated about these things. Unfortunately, they're not obvious. We have to learn these skills. But they're easy to learn. For that reason none of us have an excuse for doing this poorly. I would invite you to take the time to take advantage of resources and learn to do it well. If the data you are passing on is important enough to take the time to pass it on, it's important enough to do it in a way that is really affective.

>> With those points I leave you. At this point we do have just a few minutes in our time for me to try to field some of your questions. I will pass control back over to now to Alicia of GSA.

>> Great. Thank you. This is such a very informative webinar. I really enjoyed it. I hope that everyone else did too. We have a few questions here. Mainly there are a couple questions that are surrounding small multiples. Would you like to talk just a minute about small multiples and your thoughts on using those and is there an easy way that you know of to produce charts in popular software?

>> Absolutely. Small multiples is a term that was coined by Edward Tufte. It's a set of graphs that have been all arranged together for purposes of making comparisons between them. Sometimes the information that you need to display cannot all be displayed in a single graph. In order to be able to display everything you can often break it down into a series of multiple graphs that each have the same quantitative scale, so you can make comparisons. It's a way of putting a lot more information in front of our eyes at once. Imagine a whole array, maybe 15 different bar charts, so we can make comparisons. The question that is asked: Is there software available that makes that easy to do? There are some products that make it incredibly easy to do. Unfortunately they're not the ones that are most readily available. Just to name a few packages that do a good job Tableau, Spot Fire, Jump, . Those do a good job of it. They're available to purchase. But they're not packages that most people have available to them. What about doing small multiples with Excel? Unfortunately it's not a standard feature of Excel. Right now you have to create each of the small multiples individually and take the time to arrange them. That's a lot of work. There are ways of doing it with Excel that takes that work out of it. There are some add -in products that you can get. There's a new one that I haven't had a chance to really play with myself, I can't say whether it's really affective or not, it's from Dimensions. This makes this capability something that is easy in Excel. Also with the latest release of Excel, 2010, there's an add -in that comes with it called Power Pivot. Not nearly as good as what you would find in packages like Tableau and Spot Fire. But they do provide the functionality. I hope that answers the question, at least in a way that gets you started.

>> Great. And have you checked these various examples of displaying information with accessibility laws and usability testing?

>> I really haven't focused on that a lot in my work. I'm very aware of the problem.

I actually feel that some of the restrictions that are required for that accessibility go beyond what is necessary. Of course, one of the things that the regulations, or the compliance guidelines are trying to do is make color, or to deal with problems of color blindness, using alternatives to color. I think that takes it too far. The vast majority of people that are color blind, it's not that they can't see colors, they can discriminate particular combinations. Most people who are color blind will not have any problem as long as we avoid using the green and red combination. There's wonderful websites, color Brewer.org, it provides suggestions for colors that will work for various purposes, including for people are color blind. I know that's not exactly the answering the question that was asked. When I deal with some of the accessibility issues I don't necessarily restrict myself in the way that the guidelines say you have to, because I don't think it's always necessary. I hope that helps a little bit.

>> Okay. What about using a screen reader? Do you test them with a screen reader? Would you recommend that?

>> I don't. Not that I don't think that would be a useful thing. In my work I have actually not gotten involved in any form of display that will translate graphical presentations to people who can't see, whose sight is impaired to the point where they could not see the representation. It's hard to translate into verbal terms a visual display in a way that provides the richness of the visual display. One of the great things about data visualization is it allows people to deal with a whole lot of data at once because of the power that is built into our eyes. When you try to translate into a form of presentation that is available to our ears it's a much slower process. People are never going to be able to deal with the richness in the data as they could if they could actually see it. Unfortunately, even though we can make some attempts to translate these visual displays into verbal or oral terms we cannot do it in a way that has the richness of a visual display. I wish we could. Partly for that reason I have not gotten involved in work or research that is focusing on that particular concern.

>> Okay. Great. Another question is how would you suggest avoiding distortion of findings when you start the Y axis at a value other than zero?

>> Really good question. Depending on the audience you may want to in many cases avoid starting at anything other than zero, or if you do in order to make the differences in values easier to see, which might be difficult to see if you started at zero, if those cases you might want to draw attention to the fact that you have narrowed the scale to make the differences easier to see and compare. You have to understand your attention and draw attention to those things when you do narrow the scale.

>> Great. We're just about out of time. I did want to ask this last question: What about social math? Do you have any examples of how social math can be communicated and have a great impact?

>> By "social math" I assume you are talk being social networks and how to display those visually? Something like that?

>> I don't --

>> I will assume that is what is being asked. That's become popular recently, how to display relationships between networks of individuals or organizations. This is done in the intelligence community quite a lot. There are wonderful forms of displaying the relationships visually. Generally in the form of node link diagrams. You have something like circles to represent the nodes with lines to connect them to indicate the relationship between them. The lines might differ in weight or width to indicate the degree of relationship.

>> Great. Thank you, Stephen.

>> Thank you.

>> So I just want to wrap up by drawing your attention to the chat box. Our survey link is in there, we do read and use the surveys to make all of our courses better. Hope that beer meeting -- hope that we're meeting your needs. Check our website, we have an upcoming metrics course that will talk about which data you should monitor and use and how to do that effectively. That's on November 2 and 3. Be sure to check that out and sign up. I thank everyone for attending. Stephen, this was wonderful. I thank you for taking time to present to us today.

>> Thanks, my pleasure.

>> Okay. That's the end of our webinar.

>> [Event Concluded]