
Usability analysis of two article designs for a newspaper website

Jamie McAtee SLIS 597

Introduction

I constructed a usability test in order to understand how subjects would interact with two proposed designs for the article page of a newspaper website. The team at The Ledger Online in Lakeland, Florida provided two different versions of an article page they are considering for a site wide redesign.

The Ledger is a medium sized newspaper serving Polk County, Florida. The website gets approximately 3 million hits per month. The Ledger has a technical and content staff of two and little knowledge on how to design usability tests. This made usability testing almost impossible for this organization.

The key difference between the layouts is the placement of the additional content and the article tools. One of the layouts has everything aligned to the left of the article text. The other layout has everything to the right of the text. These two layouts will be referred to as left layout and right layout in future references.

The main question The Ledger wanted answered through usability testing was which side of the page would the reader see and understand that the additional content was related to the story. Their second question was which layout allowed readers to see and use the article tools.

Testing Setup

The test was conducted using a MacBook running Windows in Boot Camp. Video and audio were captured using the Camtasia software. The Camtasia software was used to record the audio and video of the subject as well as where they were looking on the screen. These three things were merged together on the video output that I used for the analysis.

The audio was recorded using a headset and microphone worn by the subject. The video was recorded using the iSight video camera built in to the MacBook. The Firefox browser was used to display the web pages. An external mouse was offered to the subject as an alternative to the built in touch pad. Subjects generally felt more comfortable with the external mouse while doing the test.

Evaluation

Ten subjects were tested on the each of the article pages. The subjects were all graduate students in the School of Informatics who volunteered to help out in the research.

The tests were conducted in the graduate lab of the Informatics building over a period of two weeks. Subjects were scheduled according to their availability and time preferences.

A short pre-test questionnaire was administered before each test began to gather basic information about the subjects. (Appendix I) This information was used to classify the subjects by age, time spent online and whether they had been exposed to news websites prior to the test.

In order to get fresh eyes looking at both designs the test pool was split so five subjects saw the left layout first while the rest saw the right layout first. The subjects were asked to perform four tasks on each of the article pages.

The first task asked the subject what they would do if they had trouble with the size of the text on the page. The subject was expected to enlarge the text on the screen using the on screen tool located in the article tool sets as seen in red below. The second task was to send the story to a friend. The subject was expected to use the e-mail this to a friend link located in the article tool sets as seen in red below. These two tasks were recorded as a 1 for completion or a 2 for unable to complete.



Left aligned additional content rail layout. The green area indicates the extra content area. The red boxes indicate the article tools. These were the two main items tested.



Left aligned additional content rail layout. The green area indicates the extra content area. The red boxes indicate the article tools. These were the two main items tested.

Prior to the third task the subject was asked to read a few paragraphs of the article to understand what it was about. This also gave them an opportunity to see the content on the left or right side of the page. The third task asked the subject what they learned from the content. This question was intended to find out whether the subject had seen the content on the page and to see if the subject understood that the information was related to the article. The results from this task were recorded as to whether they had seen or not seen the content on the left or right side of the page using a 1 for completion and a 2 for incomplete.

The subject was then asked how they would enlarge an image in the content rail. They had the option of either clicking the image or using the enlarge text link below the image. The reason for this question was to understand which method the subject preferred in enlarging images. The information from this task was recorded as using the graphic, 1, or the text link, 2.

After completing this task the subjects were shown the alternate design, asked to perform the same tasks discussed above on the second page. At the end of the test the subjects were asked to indicate which of the designs they preferred and to discuss why. They were also asked which they felt was clearer to see if there was a difference in their preferred design and the design they felt was clearer. The preference for either the left or right was recorded for both of these tasks.

In creating these tasks the questions I wanted to answer were:

- Q1: Which of the two layouts was easier for the subject to understand?
- Q2: Which of the designs did the subjects prefer?
- Q3: Which of the designs did the subjects find the easiest to read?
- Q4: Did the layout the subject started with influence their preferred layout choice?
- Q5: Did the layout the subject started with influence their clearer layout choice?
- Q6: Did the layout of the page influence whether the subject saw the information on the left or right side of the page?

Analysis

Using Fisher's exact test I tested for the significance of the data collected. The Fisher's test was chosen since the data set was small. The main thing I needed to determine was the statistical significance of the data.

The table contained data from the pre-test questionnaire, the results from 11 tasks encoded numerically and an indication of alignment of the page the subject started with. The data was analyzed by comparing one row of data against another row of data.

Once all five of the tests had been run the Bonferroni correction was applied to correct the alpha value for multiple tests.

I selected an initial alpha value of .05. When corrected using Bonferroni it became .01. To address the questions above I ran five tests. I compared the alpha value from each of the tests to the corrected alpha to determine the statistical significance of the test.

Tests Run				
Value 1	Value 2	P value	Selected P	Adjusted P
Starting Layout	Enlarge Text Task	0.5238	0.05	0.001
Starting Layout	E-mail a Friend Task	0.1667	0.05	0.001
Starting Layout	Preferred Layout	1	0.05	0.001
Starting Layout	Clearest Layout	1	0.05	0.001
Starting Layout	Noticing of Additional Content	0.04762	0.05	0.001

Results

Four of the subjects were women the other six subjects were men. All of the subjects read news online and spend at least 4-6 hours online per day.

In order to understand which of the two layouts was easier for the subject to understand I used the results from two separate tests. The test consisted of the ease of finding the text size tool and the send to a friend tool on the first page they were exposed and compared to the starting page. During the data collection process I found all of the subjects were able to quickly figure out how to use the article tools on the second page they were exposed to since they saw same symbols on both pages.

The analysis of these two tests found that neither test was statistically significant. The comparison of the starting page to the locating of the text enlarge button had p value of 0.5238. The p value of the comparison of the starting page to the locating of the email this to a friend button was 0.1667. The data in the comparison of the starting page to the locating of the text enlarge button is fairly independent. The data in the comparison of the starting page to the locating of the e-mail a friend button is less independent than the previous test. This follows what I saw in the testing where subjects who located the text enlarge were able to easily spot the e-mail a friend link since it was grouped with that symbol.

The data I collected did not allow me to answer the questions of which of the designs the subjects preferred or which of the designs the subjects found easiest to read from a statistically significant standpoint.

Regarding the question of whether the layout of the subject started with influenced the layout they chose in the end as their preferred layout. I chose the column that recorded the subjects starting page and the column that recorded the subject's preferred layout. The p value for this data was 1 which was statistically insignificant and highly independent. This indicates in a way that the page the subject started with did not influence the page they chose in the end.

The question of whether the layout the subject started with influenced the subject's choice of the clearer layout was also statistically insignificant. The p value for this test also resulted in a value of 1 which was highly independent and therefore statistically insignificant. This also indicates to a certain degree that these two variables had nothing to do with the other one.

The layout of the page did have an interesting influence on whether the subject saw the information the right or the left. The p value of 0.04762 could have been considered statistically significant prior to Bonferroni correction being applied. With the correction of the alpha to .01 this item is not statistically significant. The test is the most dependent test I ran which also contributes to the possibility that given a different situation there might be significance to these results. This indicates that the layout of the page could have had an influence on whether the subject saw the additional information.

Conclusion and Reflection

Based on the data from this test I was not able to answer any of the questions from a statistically significant point of view. Finding a statistically significant result using this method of testing may not be possible.

The tests I ran had more questions than subjects which may have contributed to the difficulty of finding statistically significant results. If I could have recruited a larger subject pool it might have been possible to find a statistically significant result. It is also a reasonable assumption that even with a larger subject pool I may have still found statistically insignificant results.

I also chose to run a large number of tests, which diluted the p value when applying the Bonferroni correction. A smaller number of tests with a more targeted test looking at one aspect may have produced statistically significant results.

I experimented with some numbers in order to find what the data would need to look like in order for one of the tests to be significant under the pre-corrected condition. After running a few tests I found that at least six subjects would have had to agree in order to find a statistical significance of 0.04762. This indicates that six out of my ten subjects would have had to agree in one column and five in the other column in order to find significance in any of my tests.

As I was setting up my tests a suggestion was made to have different articles for each of the pages. It was suggested that the subject would not be familiar with the information on the page therefore giving a better read on the subject's true reaction to the alternate design. For the first test I ran I thought I had found a way to do this. When the subject got to the second layout the article tool area was the same as the previous page. I realized the two designs were set up to be used on only one story.

Based on some of the feedback I received from subjects I believe it might have been a good idea to use two different stories. A few subjects told me they noticed the left or right aligned information on the second page because they were looking for it in a way since they had seen it previously. If the story and image would have been different it might have been possible to get a clearer reaction to the alternate designs from a single subject.

Not all of the data that was collected for this study was quantifiable. Although this data couldn't add to the statistical significance of the test it is valuable to the design process. The questions that were asked and not used in this analysis may still yield important insights and trigger questions about the designs.

Confusion about what a symbol or information on the page could mean is valuable non-quantifiable information. Delayed location of the enlarge text size button, not locating it at all and using a keyboard shortcut instead of the tool or finding the button at the bottom of the page are all valuable yet not really quantifiable insights. These things can contribute greatly to the user interface design of a site.

The results of this test contribute to the understanding of the way the subject interacts with the designs. Even though the test did not lead to statistically significant results it still gives the design team at The Ledger an idea of how usable both of their interfaces are. The videos of the subjects contain a wealth of information that could be used to inform this design as well as other things they may be considering for the site.

Appendix 1 Pre-Test Questionnaire

Article Usability Test Pre-Test Questionnaire

Subject # _____

1. What age group do you fall into?
online?

___ 20-24

___ 25-29

___ 30-34

___ 35-39

2. How much time do you spend per day

___ 1-3 hours

___ 4-6 hours

___ 7-9 hours

___ 10+ hours

3. Have you visited news web sites before?

___ Yes

___ No

4. If yes, what news websites do you frequently visit? (list as many as you would like)

5. Of the web sites above which one is your favorite and why? (can be brief)

Appendix 2 Testing Script

Thank you for participating in my usability test. Today you will be looking at two designs that are being considered for a redesign of a news website. I will be using these results to do analysis for my statistics class. The results will also be given to the news website to inform their design.

Please let me know what you are thinking by speaking into the headset.

Please take a minute to look at this page.

Let's say you are having trouble reading the text on the page. What would you do?

You really enjoyed this story and want to send it a friend how would you do that?

Now take a look at the article. Take a few minutes to read the article. When you have done this let me know.

What can you tell me about the content on the (left/right) side of the article?

Did you notice this information before I asked about it? When did you look at this information? Was it the first thing you saw?

What information did you gather from this content?

Let's say you want to get a closer look at the photo, how would you do this?

Please take a minute to look at this second page.

Let's say you are having trouble reading the text on the page. What would you do?

You really enjoyed this story and want to send it a friend how would you do that?

Now take a look at this second article. Take a few minutes to read the article. When you have done this let me know.

What can you tell me about the content on the (left/right) side of the article?

Did you notice this information before I asked about it? When did you look at this information? Was it the first thing you saw? Would you have noticed it on this page if I hadn't asked about it previously?

What information did you gather from this content?

Let's say you want to get a closer look at the photo, how would you do this?

Now that you have seen both of these layouts, which one were you more comfortable with and why?

Was one of these easier to read and understand than the other one?

Do you have any additional thoughts or questions that you would like to express?