

## Study Design Changes

There were a few changes that were made from the original plan of implementation of the study to what was implemented. Originally, it was thought that some of the solicitation for participants would take place in a face-to-face manner. Due to personal and scholarly limitations to time, it was decided that solicitation for participation would occur online. In addition, the waiting a day between participating in experimental conditions was waived. This meant that practice effects could confound the results, however reliability in experience could be increased through consistent use of websites if participants remembered what they were trying to do in one condition and applied it to the other condition.

To compensate for no in-person directions, well thought out and simple directions were established to guide participants through the study. Because the total number of questions was large and the number of participants needed to be at least 15, the use of Qualtrics and other survey websites could not be used. To decrease difficulty, use of a Google Form was made for participants to use throughout the study.

The Google Form started with the informed consent, then introduced the first set of instructions, with links to the experimental condition and a way to sign up for the website without using a personal email address to protect the privacy of participants as stated in the informed consent. Following the instructions, participants completed the first survey. On the next page, they were instructed to use the second website and then completed the survey on the following pages.

## Analysis

The study's purpose was to examine if there were any differences in system usability and interface quality between an existing TV scheduling website, [www.episodecalendar.com](http://www.episodecalendar.com), and a newly created website, [www.exithere.org/tvschedule](http://www.exithere.org/tvschedule). The new website was developed to decrease cognitive load and simplify the creation and display of a weekly TV schedule. A survey developed by IBM was used to assess system usability and interface quality.

All participants experienced both conditions (both websites). Therefore, a repeated measures t-test was used to assess if there were differences between scores on system usability and interface quality scales.

The mean score of system usability on condition one (Episode Calendar) was 3.19 with a standard deviation of 1.15 for the 15 participants in this study. The mean score of system usability on condition two (TV Schedule) was 2.94 with a standard deviation of 1.57. There was no evidence to suggest that there was a statistically significant difference in usability scores between these conditions,  $t(14) = 0.47$ ,  $p = 0.68$ . Therefore, the null hypothesis was accepted. Furthermore, there was no correlation found between the two system usability scores which indicated independence of reporting and a lack of a relationship between the two scores.

The mean score of interface quality on condition one (Episode Calendar) was 3.40 with a standard deviation of 1.39 for the 15 participants in this study. The mean score of interface quality on condition two (TV Schedule) was 3.82 with a standard deviation of 1.73. There was no evidence to suggest that there was a statistically significant difference in interface quality scores between these conditions,  $t(14) = -0.88, p = 0.40$ . Furthermore, there was no correlation found between the two system usability scores which indicated independence of reporting and a lack of a relationship between the two scores.

## **Discussion**

Overall, my study found that there was no statistical difference in system usability or interface quality. Users responded that they found the systems approximately equal in ease of use, timeliness, efficiency, and comfort as well as pleasure of use and expected functions. Unfortunately, this means I can draw few conclusions regarding my interface and its implications in the future for developers, designers, or users. In general, we know it is important to create interfaces, which are simple to navigate, understand, and learn and also provide enough information without overwhelming the user. Feedback received indicates a preference for image (vs. text) schedules as well as simplified search functionality. However, the statistics did not reflect this. It could be a case that the study did not properly assess these certain qualities. Perhaps the interface users are looking for is in fact more than TV Schedule gave them, and less than Episode Calendar gives. A larger sample size could also help in uncovering other trends.

One of the anticipations of creating this new website was that users would find it simple and that would be a good thing. Some of the comments made during the trial and even during the full study indicated that the interface was clean and simple. Compared to the commercial website, Episode Calendar, TV Schedule includes very little outside the schedule itself; that is what the focus is. While Episode Calendar provides more information to the user, it takes away from the schedule and becomes more about the information of television shows. The measure used to quantify interface quality and system usability also includes a scale for information quality. Analysis of these results could possibly indicate that Episode Calendar had better scores in information quality if users wanted to know all about their shows and not just create a schedule to follow week to week. This would be an interesting follow-up study that could include information on what it is that users are looking for in entertainment planning websites.

Since some of the more technical programming aspects of TV Schedule were not implemented by the time the study took place, it was anticipated that scores on interface quality were not significantly better than those from Episode Calendar. Again, further insight into what it is users want out of a website will yield strong correlations to what kind of scores they indicate in usability and quality. This highlights the importance of developers and designers understanding their user base while still providing a simple enough product that will attract even the novice user.

One thing that was different than anticipated was the scores being so similar. It could be biased selfishness that led to thoughts that the newly created site was more of an improvement than it really was. In reality, it could be kind of a compliment to have scores be similar when

compared to a commercial website like Episode Calendar. In addition to this though, it was not expected to have such little user feedback provided by comments throughout the study. During the trial, there were lots of comments. This could be due to the fact that participants knew that this was for an evaluation of the effectiveness of the study and had participated with verbal comments in addition to providing comments in text boxes throughout the online survey. More qualitative feedback from participants would help guide interpretation given the lack of significance found in the results of the analysis.

Given the experiences from this study, it can be hard to produce a solid set of design guidelines due to the limited nature of participant information during the full study and low sample size. With the information that is available, the best guidelines will be delivered.

Knowing the thought process of users is important. Start with a goal in mind and then broaden out. In many cases, people start by going from big picture to small picture. But when designing a website for something specialized, it is important to develop that specialty to the best that it can be. All other information is superfluous at that point. Once a consistent, reliable, and usable method is developed to achieve that goal, the extra information should be added in a way that is consistent with that initial input. It should all build upon itself to create an experience that is relatable to the purpose that drew the user to that service in the first place.

For example, a user will want to develop a television schedule to organize their weekly entertainment. Once within the system to develop this schedule, they quickly get to work accomplishing their goal. Along the way, the user might have a thought about other various information pertaining to the shows they were adding. From there, it could be possible to obtain more information on those, but the implementation of it should be something that does not distract from the development of a television schedule. Since participants noted that on Episode Calendar, they were just clicking around, it is important to maintain clarity even though more layers are being added on into the service.

Another aspect to keep in mind is user input preference. The development of TV Schedule's system included a bug that would clear the search bar if "enter" was pressed. This behavior is inconsistent with what some users were used to and inhibited their connection and use for the service. They had to re-train themselves to not press "enter" when concluding a search so as to access the show list to incorporate into their schedule. Keeping in mind what the user is used to help reduce the effort to engage in the system and to produce a more natural feeling experience.

## Appendix

Table 1

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	SysUse1	3.19167	15	1.145514	.295770
	SysUse2	2.94167	15	1.573970	.406397
Pair 2	InterQual1	3.40000000000000	15	1.3870146083620	.35812563194225
	InterQual2	3.82222222222223	15	1.7268499014132	.44587072730496
		99		75	6
		22		43	8

Table 2

		N	Correlation	Sig.
Pair 1	SysUse1 & SysUse2	15	-.147	.601
Pair 2	InterQual1 & InterQual2	15	.297	.283

Table 3

	Paired Differences					t	d	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Pair 1	SysUse1 - SysUse2	.250000	2.078268	.536606	-.900906	1.400906	.466	.648
Pair 2	InterQual1 - InterQual2	.422222222222222	1.866439895522	.481912708803	1.455822184802	-.611377740358	.874	.396
		323	558	806	913	469	6	

# TV SCHEDULING

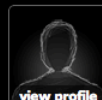
---

Reducing cognitive load to increase usability  
and quality

Lauren Gray

# Existing System



[Calendar](#)
[Unwatched](#)
[My shows](#)

[All shows](#)
[Trends](#)
[FAQ](#)
[Settings](#)
[Logout \(test4\)](#)
[« March](#)
**April**  
2013

[May »](#)

Mon	Tue	Wed	Thu	Fri	Sat	Sun
Apr. 1st	Apr. 2nd	Apr. 3rd	Apr. 4th Grey's Anatomy: 9x20 <input type="checkbox"/> New Girl: 2x21 <input type="checkbox"/> <a href="#">The Big Bang Theory: 6x20</a> ✓	Apr. 5th	Apr. 6th	Apr. 7th The Only Way is Essex: 8x12 <input type="checkbox"/>
Apr. 8th	Apr. 9th New Girl: 2x22 <input type="checkbox"/>	Apr. 10th	Apr. 11th	Apr. 12th	Apr. 13th	Apr. 14th Once Upon a Time (2011): 0x2 <input type="checkbox"/>
Apr. 15th	Apr. 16th	Apr. 17th	Apr. 18th	Apr. 19th	Apr. 20th	Apr. 21st Once Upon a Time (2011): 2x19 <input type="checkbox"/>
Apr. 22nd	Apr. 23rd	Apr. 24th	Apr. 25th Grey's Anatomy: 9x21 <a href="#">The Big Bang Theory: 6x21</a>	Apr. 26th	Apr. 27th	Apr. 28th Once Upon a Time (2011): 2x20
Apr. 29th	Apr. 30th New Girl: 2x23	May. 1st <a href="#">New Girl: 2x23</a>	May. 2nd Grey's Anatomy: 9x22 <a href="#">The Big Bang Theory: 6x22</a>	May. 3rd	May. 4th	May. 5th <a href="#">Once Upon a Time (2011): 2x21</a>

# My System

Once upon a time

Once Upon a Time  
Once Upon A Time In Cabramatta  
Once Upon A Time In Wonderland  
Adventure Time  
Question Time  
Dirty Dancing: The Time of Your Life

Show/Hide Network [Print](#)

Time	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
5:00 PM							
8:00 PM	Once Upon a Time ☒						
9:00 PM							

# User Study Conditions

- 15 Participants
- Within Subjects
  - Each person participates in each condition
  - Strength in internal reliability
  - Reduces issues in within subjects variability
  - Allows same user to comment with same perspective on both systems
    - Did not receive enough comments



# Results and Analysis

**Paired Samples Statistics**

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	SysUse1	3.19	15	1.15	.296
	SysUse2	2.94	15	1.57	.406
Pair 2	InterQual1	3.40	15	1.39	.358
	InterQual2	3.82	15	1.73	.446

**Paired Samples Test**

		Paired Differences				t	df	Sig. (2-tailed)	
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower				Upper
Pair 1	SysUse1 - SysUse2	.250	2.078	.537	-.901	1.401	.466	14	.648
Pair 2	InterQual1 - InterQual2	-.422	1.866	.482	-1.456	.611	-.876	14	.396

# Conclusions

- No statistical significant differences in usability or interface quality
- Mean for usability lower (better) for my system than existing system
- Mean for interface quality lower (better) for existing system than my system
- Larger sample to identify if this becomes significant with a study that has more statistical power