Comparative Usability Evaluation for an e-Government Portal

Jason Withrow, Tom Brinck, Alfred Speredelozzi
Diamond Bullet Design
in Collaboration with
the National Information Consortium
315 W. Huron St. Suite 140
Ann Arbor, MI 48104
+1 734 665 9307
tom@diamondbullet.com

ABSTRACT
A redesigned state government world wide web portal, launched in December 2000, was compared to the previous portal site for usability issues. Twenty subjects completed ten tasks each. The results show that the redesigned version of the site is more usable than the previous version. Statistically significant improvements were found on task success, time to complete a task, task difficulty, expectations of how long tasks would take relative to actual time, and how intuitive the paths were to desired information. These findings suggest that the investment in usability was justified as users are likely to be more efficient and effective when using the new portal.

Keywords
Government, e-Government, Portal, World Wide Web, Design, Measurement, Performance, Experimentation, Verification, Human Factors, Electronic Commerce.

INTRODUCTION

A private company, which operates many state governments' websites and electronic data systems, contracted Diamond Bullet Design to re-architect the e-government portal for a particular state. Diamond Bullet, located in Ann Arbor, Michigan, is a web design firm specializing in usability. The portal is a directory of government web services (some separately provided by the e-government company) which are the responsibilities of various government agencies, departments, or non-profit entities.

A design team followed a user-centered design process [1] in building the new state e-government portal. This multidisciplinary team was composed of usability engineers from Diamond Bullet Design, an information architect from the e-Government company and graphic artists from both companies. The process began with a user needs analysis including user profiling, interviews, and card sorts with the state's citizens. Competitive analysis and task analysis were performed to identify best practices and areas for improvement in the existing site.

cite as:
Withrow, J., Brinck, T., and Speredelozzi, A. *Comparative Usability Evaluation for an e-Government Portal*, Diamond Bullet Design Report #U1-00-2, Ann Arbor, MI. Dec. 2000.

THE DESIGN PROCESS

The usability specialists examined the original site as well as the web sites of all fifty U.S. states. Differences in graphic design as well as architecture were noted. Also, the variety of possible interactive features for a state website were listed and compared to the original website. Search query logs and hit log data from the old site were also analyzed to find out what users were currently doing on the original website. Then, twenty-five common tasks were mapped out, to better understand the current process and identify improved methods of completing the tasks. The number of steps to complete each task, as well as the number of possible ways to complete a task was listed. Some tasks were redesigned, however, most of this information was used to inform user interviews and architecture revisions.

Based on the information from these analyses, four distinct site architectures were developed for the new portal. The architecture of the previous portal organized most information by government agency. For example, a "Business & Commerce" section led to a set of content mostly developed by the Department of Commerce and Housing. Figure [1] shows the main architecture of the site, with one expanded subsection.

The proposed architectures focused on other ways to approach the information, such as through the services provided by the agencies. The alphabetical lists of state agencies, which the previous site provided, often confused users. A thematically organized option was explored, in which the user navigated according to the role that they identified with. The candidate roles were terms such as "Citizen," "Government Employee," "Vacationer" and "Business Owner." Other architectures were developed, which allowed a competitive evaluation of each idea. To determine which architecture was most usable, in terms of task completion and task difficulty, six rounds of user

testing were conducted. Findings from each round of user testing were then iteratively built into all the architectures prior to the next test.

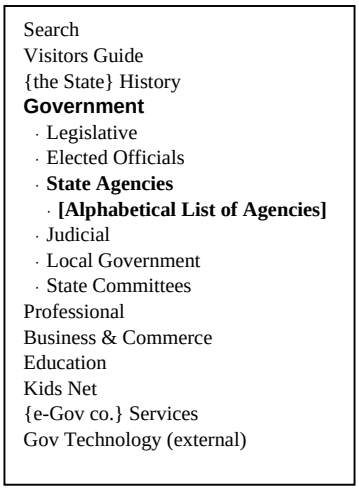

Figure 1: Architecture for the old portal, with two expanded levels

The testing was conducted using HTML wireframes, with the locations and users alternating between Diamond Bullet Design's offices in Ann Arbor (Michigan citizens) and locations in the target state (target state citizens) near the e-government company offices. The number of users tested in each round ranged from 10-13, with 8-10 tasks tested per user.

The final architecture that was selected for the new portal followed a hybrid thematic-topical approach, with thematic top-level categories such as 'Living in {the State}' and 'Working in {the State}' and topical top-level categories such as 'Government' and '{the State's} Facts and History' (where {the State} would be replaced by the state's name). The architecture of the previous portal followed a purely topical approach for its top-level categories. The lower levels of both sites also differed, as the new site organized content topically, while the old site primarily organized content by state agency. Figure [2] shows the architecture of the new portal.

Following the final round of user testing, the team built the new portal using a redesigned graphical look-and-feel and information architecture. The logos and branding of the website had also been redesigned. Special attention was also paid during site creation to accessibility concerns, XHTML compliance, and the inclusion of Dublin Core metadata attributes for better page indexing.

The present study compares the new portal to the old portal on numerous measures. This study aims to show that the improved architecture substantially improved usability. Further, the overall improvements of the site have demonstrated that the user-centered redesign approach was worthwhile. Differences favoring the new website suggest that the investment in developing the new portal has resulted in a more efficient, enjoyable, and successful experience for visitors to the site.

Figure 2: Architecture for the new portal, with two expanded levels

METHOD

Participants

Twenty subjects (12 men, 8 women) were tested. Eighteen subjects reported an ethnic background of Caucasian; two reported African American. Subjects' ages ranged from 20s to 70s, with most subjects in their 20s and 30s. All subjects were Michigan residents. Subjects were recruited through newspaper advertisements.

Materials

The new website was tested using the live version currently available on the world wide web. To test the old portal, a copy of that site was hosted on a server at Diamond Bullet. To control for possible platform and browser differences, testing was divided equally between an IBM-compatible PC and an iMac and between Netscape and Internet Explorer, respectively. Netscape 4.7 and Internet Explorer 5.5 were the browsers used on the PC. Netscape 4.7 and Internet Explorer 5.0 were the browsers used on the iMac. All testing occurred in the usability lab at Diamond Bullet Design, using a T-1 connection.

1.You are interested in renewing a {State} driver's license online.
2.How do nurses get licensed in {the State}?
3.To assist in traveling, you want to find a map of {State} highways.
4.What four-year colleges are located in {the State}?
5.What is the state bird of {the State}?
6.You are interested in registering as a voter in {the State}.
7.Who are the U.S. representatives and senators from {the State}?
8.How would someone start a business in {the State}?
9.What are the current road conditions on {State} highways?
10.You are interested in getting a hunting license.

Figure 3: Tasks Tested

Design and Procedure

A between-subjects design was used, so each subject was tested on only one of the sites. Testing occurred over a period of two weeks. Subject assignment to the various conditions (site, computer, browser) was counterbalanced. Each session lasted one hour, with the subjects completing ten tasks [Figure 3] on the websites using a think aloud protocol. Five of these tasks had been tested previously during the development of the new website, while the remaining five tasks were tested for the first time in this study. Those five tasks were drawn from a pool of candidate tasks originally developed during the design of the new site. Also, since some of the content changed in the redesign, tasks were eliminated if they could not be accomplished on both sites. Tasks were selected to be common and representative actions for a range of users based on hit log data, search query logs, and interviews conducted during the redesign effort.

Subjects were instructed to approach the tasks from the perspective of a citizen of the target state. A facilitator and data recorder were present during the testing. The data gathered included: time to complete each task and success/failure, user path through the site and comments, subjects' ratings of difficulty, how quickly the subject completed each task compared to his/her expectations, how intuitive the path to the information was, and overall subject-reported ratings of the website on various dimensions. The dimensions were: Easy to Use, Responsive, Interesting, Efficient, Well-Organized, Good Labels, Direct, Focused, Attractive, Easy to Read, Useful, Valuable Information, Comprehensive, and Good Instructions. These dimensions were rated using Likert scales.

RESULTS

All data analyses used two-tailed independent samples t-tests. Task success was measured as the percentage of successfully completed tasks out of total tasks attempted, measurements of task time were in seconds, and all other measures used a 1-7 Likert scale (lower values are better).

Task Success

On the new site subjects achieved an average task completion rate of 95% (SD = 7.07), compared to an average task completion rate of 72% (SD = 11.35) on the old site, indicating a large and significant improvement for the new portal ($t(18) = -5.44$, $p < .001$).

Task Time

Task time was examined at both aggregate (considering all tasks) and task-specific (comparing performance per task on both sites) levels. At the aggregate level, subjects took an average of thirteen and a half minutes less (a savings of 79 seconds per task) to complete all ten tasks on the new site (M = 504.41 sec., SD = 251.07) compared to subjects using the old website (M = 1319.91 sec., SD = 381.50). ($t(18) = 5.65$, $p < .001$)

At the task-specific level, the new site also showed significant improvement over the old site. As shown in Table 1, mean task time was significantly shorter on the new site for the tasks related to finding a map of the state, locating four year colleges, discovering the state bird, registering as a voter, starting a business, determining road conditions, and getting a hunting license. For two other tasks, renewing a driver's license and getting licensed as a nurse, mean task times were shorter on the new website, although not to the level of statistical significance. Subjects using the old website had a lower mean task time when finding the U.S. representatives and senators from the state; this difference was not statistically significant.

Table 1: Task Time by Site

Task	Site	M (sec)	SD	p
Driver's license	old	174.01	91.1	0.34
	new	**126.42**	112.28	
Nurse licensure	old	134.35	125.71	0.45
	new	**95.19**	88.32	
*State Map	old	205.99	116.25	0.002
	new	**66.57**	33.55	
*Four year colleges	old	76.04	29.08	0.001
	new	**30.32**	19.01	
*State bird	old	120.05	72.92	0.014
	new	**47.24**	42.14	
*Voter registration	old	194.52	132.19	0.003
	new	**47.5**	16.92	
U.S. legislators	**old**	**25.73**	7.72	0.7
	new	27.71	13.91	
*Starting a business	old	58.57	32.87	0.003
	new	**20.5**	12.03	
*Road conditions	old	172.37	82.79	0.001
	new	**31.56**	32.3	
*Hunting license	old	158.28	115.53	0.011
	new	**46.2**	48.34	

Note. All times are in seconds.

* indicates statistical significance

Task Difficulty

The new site also showed improvement relative to the old site in terms of how difficult subjects thought tasks were to complete. (The question to the user stated, "How hard was this task?" with answers ranging from "very easy" to "very difficult.") As shown in Table 2, mean difficulty ratings were significantly lower on the new site for tasks related to finding a map of the state, registering as a voter, and determining road conditions, compared to mean difficulty

ratings for the same tasks on the old website. For the remaining seven tasks, the mean difficulty ratings favored the new site, but not to the level of statistical significance for tasks related to the following: renewing a driver's license, getting licensed as a nurse, locating four year colleges, discovering the state bird, finding the U.S. representatives and senators from the state, starting a business, and getting a hunting license.

Table 2: Task Difficulty by Site

Task	Site	M	SD	P
Driver's license	**old**	**3.5**	2.42	0.39
	new	4.4	2.12	
Nurse licensure	old	3.1	2.18	0.41
	new	**2.4**	1.43	
*State Map	old	4.8	2.39	0.026
	new	**2.6**	1.58	
Four year colleges	old	1.8	1.03	0.22
	new	**1.3**	0.67	
State bird	old	2.7	1.34	0.06
	new	**1.7**	0.82	
*Voter registration	old	4.8	2.2	0.008
	new	**2.5**	1.08	
U.S. legislators	**old**	**1.2**	0.42	0.26
	new	1.5	0.71	
Starting a business	old	2.1	1.29	0.34
	new	**1.6**	0.97	
*Road conditions	old	5.2	2.44	0.001
	new	**2**	1.05	
Hunting license	old	3.4	1.96	0.11
	new	**2.2**	1.14	

Note. Lower values indicate less difficulty.

* indicates statistical significance

Expected Speed

Additional improvements were seen with the new site in expected speed, which is the subject-reported rating of how fast the task was to complete, compared to his/her expectation. (The question to the user states "How quick was this task, compared to what you expected?" with answers ranging from "much faster" to "much slower.") The mean expected speed was significantly lower for the new site on the task involving determining road conditions, as shown in Table 3. Expected speed ratings for eight other tasks (renewing a driver's license, getting licensed as a nurse, finding a map of the state, locating four year colleges, discovering the state bird, registering as a voter, finding the U.S. representatives and senators from the state, starting a business, and getting a hunting license) also favored the new site, although the differences were not statistically significant. The magnitude of those differences, however, suggests that the new site showed substantial improvement compared to the old site. The task involving driver's license renewal achieved a mean expected speed rating favoring the old site; this difference also failed to reach statistical significance.

Table 3: Expected Task Speed by Site

Task	Site	M	SD	p
Driver's license	**old**	**3.7**	2.41	0.29
	new	4.9	2.56	
Nurse licensure	old	3.4	2.07	0.42
	new	**2.7**	1.7	
State Map	old	4.8	2.49	0.12
	new	**3.1**	2.18	
Four year colleges	old	2	1.15	0.39
	new	**1.6**	0.84	
State bird	old	3	1.33	0.06
	new	**1.9**	1.1	
Voter registration	old	4.4	2.12	0.08
	new	**2.9**	1.37	
U.S. legislators	old	1.5	0.97	0.6
	new	**1.3**	0.67	
Starting a business	old	2	1.25	0.58
	new	**1.7**	1.16	
*Road conditions	old	5.2	2.2	0.008
	new	**2.7**	1.49	
Hunting license	old	3.3	2.26	0.49
	new	**2.7**	1.49	

Note. Lower values indicate less task time than users expected.

* indicates statistical significance

Task Intuitiveness

The intuitiveness of a task, how obvious the path to the information was for subjects, also showed improvement on the new site relative to the old site. (The question to the user stated, "How obvious were the steps in this task?" with answers ranging from "always obvious" to "always unclear.") Three tasks (finding a map of the state, registering as a voter, and determining road conditions) were rated as significantly more intuitive on the new site, as shown in Table 4. The tasks involving renewing a driver's license, locating four year colleges, discovering the state bird, starting a business, and getting a hunting license achieved better mean intuitiveness ratings on the new site. The tasks involving getting licensed as a nurse and finding

the U.S. representatives and senators from the state achieved better mean intuitiveness ratings on the old site. None of these differences achieved statistical significance.

Post-Test Measures

Subject ratings of the sites on various dimensions suggested that the new site was making progress over the old site in most domains, although none of the differences reached statistical significance (see Table 5). The mean ratings for six of the ten tasks (Easy to Use, Comprehensive, Efficient, Good Instructions, Good Labels, Well-Organized, Valuable Information, Responsive, and Easy to Read) favored the new site. Mean ratings for Attractive, Direct, Focused, and Useful favored the old site, with the two sites having the same mean rating for Interesting.

Table 4: Task Intuitiveness by Site

Task	Site	M	SD	p
Driver's license	old	4.3	2.11	0.91
	new	**4.2**	1.87	
Nurse licensure	**old**	**3.3**	1.95	0.73
	new	3.6	1.9	
*State Map	old	4.6	2.27	0.04
	new	**2.6**	1.65	
Four year colleges	old	2	1.05	0.36
	new	**1.6**	0.84	
State bird	old	3.1	1.6	0.09
	new	**2.1**	0.74	
*Voter registration	old	4.9	2.13	0.01
	new	**2.7**	1.25	
U.S. legislators	**old**	**1.2**	0.42	0.45
	new	1.4	0.7	
Starting a business	old	2.2	1.48	0.21
	new	**1.5**	0.85	
*Road conditions	old	5.2	1.87	0.001
	new	**2.2**	1.32	
Hunting license	old	3.7	2	0.12
	new	**2.4**	1.51	

Note. Lower values indicate greater intuitiveness.

* indicates statistical significance

DISCUSSION

The findings from the user testing indicate that significant improvements have been made in the usability of the state e-government portal. Users of the new portal take significantly less time to complete most tasks and are more successful in finding desired information than users of the old website. Furthermore, the trends in the data suggest that users of the new site experience less difficulty finding information, locate information more quickly than they expected, and perceive the site architecture to be more intuitive than users of the old site.

Table 5: Post-Test Measures by Site

Task	Site	M	SD	p
Easy to Use	old	3.3	1.42	0.09
	new	**2.3**	1.06	
Comprehensive	old	2.3	1.06	0.37
	new	**1.9**	0.88	
Attractive	**old**	**2.3**	1.16	0.21
	new	3.2	1.86	
Direct	**old**	**2.3**	0.95	0.44
	new	2.8	1.75	
Efficient	old	3.4	1.65	0.25
	new	**2.6**	1.35	
Focused	**old**	**2.3**	0.67	0.18
	new	3.2	1.93	
Good Instructions	old	3.6	1.43	0.16
	new	**2.6**	1.65	
Interesting	old	2.9	0.88	1
	new	2.9	1.45	
Good Labels	old	3.7	1.34	0.34
	new	**3**	1.83	
Well-Organized	old	3.1	1.2	0.87
	new	**3**	1.49	
Easy to Read	old	2.5	1.35	0.73
	new	**2.3**	1.16	
Responsive	old	2.4	1.51	0.88
	new	**2.3**	1.49	
Useful	old	1.7	0.95	0.85
	new	**1.8**	1.32	
Valuable Information	old	2	0.67	0.21
	new	**1.6**	0.7	

Note. Lower values are better.

Where the new system rates slightly worse (e.g. attractiveness), the difference has been small, not statistically significant, and generally limited to only a few tasks.

We believe certain features of the architecture are responsible for the increased usability. First, there is less ambiguity between link labels. Most of the labels below the top level of navigation include descriptive text, which helps explain what the user will find when following that link. Also, the categorization of the portal links was made to be more intuitive to the end user. Focusing on users' actions and comments, rather than categorizing items by

how they were related to the appropriate government agency, is how this was accomplished. Finally, the path to any given piece of information was actually longer in the new design. The expectation is that this would reduce the speed at which people could find information, but instead speed increased.

This enhanced usability will have a far-reaching positive impact for the state and its citizens, as the citizens can find information and services with a minimum of time, effort, and frustration. The effort of the team's user-centered design approach–keeping the focus on user needs, rather than the structure of the existing government information—produced this improvement.

Although the findings show that the usability of state e-government portal has substantially improved, the limitations of the present study should be kept in mind when interpreting the results. The laboratory tasks, while based on what real users have done, did not perfectly represent what a user would do in the on their own. For instance, the subjects tested in this study did not live in the target state, although they were instructed to approach the tasks from the perspective of a citizen from the target state. Also, subjects were not allowed to use the search engines on the websites and some subjects mentioned a preference for searching over browsing. Finally, the task times may be compressed due to testing on a T-1 line; most users in the target state are likely to be using dial-up connections with significantly less bandwidth. There is no evidence to suggest, however, that the pattern shown in the task times (i.e., quicker times on the new site) would differ on the slower connections.

CONCLUSION

This evidence of improvement in many areas of usability has made the redesign project worth the cost for many reasons. One is the goodwill of the people. The government of the state contracts with the e-government company to provide this web portal service. If users were dissatisfied, complaints to government officials would weaken the company's relationship with the state. A site that is difficult to navigate or overly complex is likely to cause user frustration. People are particularly sensitive about government systems, whose capabilities are often underestimated.

As another example, the savings from increased website usability are calculable for its users [2]. Of the 2.7 million residents [3], we might conservatively estimate a quarter of them use the website at least once per month. If each of them saved 79 seconds (as was the average task savings in this study) then about 53 million seconds (14,800 hours) are saved per year. Converting this to labor costs, we find 370 person-weeks (at 40 hours per week) or 7 person-years are saved per month. 84 person-years are saved each year. On average, a citizen in the target state had a salary of $14,700 [3]. This leads to a yearly benefit of $1.2 million.

However, the actual revenue generated by the website is limited to a fixed set of information applications. Users must be able to quickly find the applications that they desire, or they might abandon the website in favor of traditional brick-and-mortar ways to obtain their information. The failure rate of the old portal was found to be 28%, while the new site was 5%. We might assume that 100,000 users would pay a service fee on the order of $2 per transaction at least once a month. Then the 23% of them who are succeeding on the new site, whereas before they were failing, are generating an additional $552,000 in revenue per year.

Even people who are looking for free information could potentially be customers of the pay services in the future. A good impression about how easy it was to find the regulations for fishing in the state, for example, might convince a person that it is worth while to use the online fishing license system. Citizens outside of the state might be more inclined to explore for tourism possibilities in the state. Outside businesses might increase their consideration for working with the state. A number of similar benefits can be imagined.

Finally, the goal of the e-government company, is to make government more efficient for its citizens. By making an *easy to use* website, the company will increase the number of people who choose to do business with government electronically, rather than over the phone, or at the state capital. This allows the government to better capitalize on its investment in information technologies. It also extends the reach of government to places that were too expensive to provide services in the past.

The design team expects that even greater improvements can be expected in usability with additional user testing and design iterations. The redesign effort was focused on only one aspect of the web portal—the architecture. The portal also includes a number of interactive services that could be improved with future testing and design as one example. As these other aspects are improved, the benefits will increase, in the same manner.

ACKNOWLEDGMENTS

Our thanks to developers and managers at the aforementioned e-Government company, especially, Diana Persell, Debra Luling, and Melinda Morris-Black. Supporting personnel at Diamond Bullet Design included Rod Lowe, Juliane Morian, and Bryce Erwin.

REFERENCES

1. Bias, R., and Mayhew, D. *Cost Justifying Usability*. . Morgan Kaufmann, San Francisco, 1994.
2. Brinck, T., Gergle, D. and Wood, S. *Usability for the Web: Designing Web Sites that Work*. Morgan Kaufmann, San Francisco, 2001 (in press).
3. Wong, J. D, and Durkes, A. H. *The Governor's Economic and Demographic Report, 2000-2001*. Kansas Division of the Budget, Topeka, KS, 2001. Available at: http://da.state.ks.us/budget/.