

Museum website features, aesthetics and visitors' impressions: a case study of four museums.

Abstract

In an effort to extend previous work on developing general recommendations for design of art museum websites, we evaluated four museum websites (Tate, the Dallas Museum of Art [DMA], the Solomon R. Guggenheim Museum and the Rijksmuseum) and examined the relationships between website features, virtual visitors' impressions and estimates of repeat visits to the site. The study produced a comparative evaluation of the four museum websites and identified strengths and weaknesses in the sites' navigation, design and content features. The study found that website aesthetics was the strongest predictor of the visitor's overall impressions. Ability to manipulate content (e.g. download, print, zoom in and out) was also linked to the visitor's inclination to make repeat visits to the site. This paper describes the methods and results of the study and proposes a set of recommendations for the development and evaluation of art museum websites.

Keywords: art museums, digital art collections, websites, usability, evaluation, case study

Introduction

Much has been written about the importance of an online presence to further the mission of art museums (Honeysett, 2007; Jones, 2008) yet there are few clear guidelines to aid museums and similar institutions in developing engaging and informative websites. In an effort to further the understanding of visitors' perceptions of museum websites and to develop design recommendations, this article a) reviews several relevant studies aimed to develop a framework for museum website design and evaluation and b) describes a two-phase study that examined user perspectives of desirable museum website features.

Relevant literature

In recent years, museums have actively begun to develop their online presence in order to increase and improve visitor interaction with their virtual and physical collections (Barry, 2006; Ciolfi et al., 2008; Bartindale et al., 2011; Ciolfi & McLoughlin, 2012). Virtual visitors expect museum websites to enable and promote learning, to be functional, easily navigable, visually pleasing and enjoyable (Tractinsky et al., 2000; Zhang & Li 2005; Lin, Fernandez, & Gregor, 2012; Van Dijk et al., 2012). While museums develop and utilize their online presence in different ways, museums generally aim to encourage user interaction and engagement with their collections (De Angeli et al., 2006; Marty, 2007; Stewart & Marcketti, 2012).

Our review of the literature on the design of museum website interfaces identified two clusters of studies: a larger cluster that reports findings concerning the usability of specific museum websites and a smaller cluster that aims to develop more general frameworks for evaluation and design of museum websites.

A heightened interest in the usability of museum websites is not surprising, since in the past decade, a significant number of publications focused on the importance of usability in virtual visitors' interactions with museum websites. For example, within the context of image collections, Walz and Brenny (2005) conducted a heuristic self-inspection evaluation of the University of Minnesota Architecture and Landscape visual resources collection. By observing participants' interactions and collecting participants' assessments of the site, the authors identified specific design weaknesses such as navigation and labeling conventions, and proposed a baseline evaluation method for the website. In an effort to determine the effects of navigation designs on usability, De Angeli et al. (2006) compared two websites with the same content but different, interactive, menu-based, navigation designs. The study showed that perception of information quality was affected by the design style. Based on their findings, the authors outlined specific recommendations for the design of museum website navigation. Cunliffe, Kritou and Tudhope (2001) tested and critiqued several methods for evaluating small museum websites, including log analysis, direct observation and online questionnaires, and reviewed the strengths and weaknesses of each method. Haynes and Zambonini (2007) tested several usability tools, including Google Analytics, to examine virtual visitors' behavior on five museum websites. The data enabled the authors to construct user profiles of the museum's virtual visitors for the use in the museum website usability studies. In a study of the Indianapolis Museum of Art's website, Fantoni, Stein and Bowman (2012) examined how a visitor's intent to use a website affected the ways users interacted with the site. The authors concluded that users' reasons for visiting a website affected the time they spent on the website as well as the number of artworks searched and downloaded. Using the Quai Branly museum's and the Atlanta History Center's websites, Pallud and Straub (2014) examined variables that influence users' online behavior on museum websites. The authors concluded that website aesthetics was the most influential factor in motivating users either to return to the websites or to visit the physical museums.

A number of publications go further than the usability assessment of specific museum websites or their features and aim to develop more general evaluation and design frameworks. In one such study, Pallas and Economides (2008) developed a Museum's Sites Evaluative Framework (MUSEF) and tested it on 210 museum websites. The proposed MUSEF included six areas that museum websites are evaluated according to: content, presentation, usability, interactivity and feedback, e-services, and technical. In turn, each area included evaluation rubrics for specific website features. For example, the usability area included user interface, site structure and organization, navigation, orientation, and search-related features. A similar framework was proposed by Blas, Pia, Romagna and Paolini (2002). The authors developed the Milano – Lugano Evaluation Method (MiLE) framework for evaluating cultural heritage websites and explored the “universe of possible functions that a museum website should support.” MiLE emphasized the inspection method of testing, using various facets of abstract and concrete tasks for museum websites such as technology, navigation, content, illocutionary force, graphic and other evaluative areas similar to MUSEF's. The authors tested the method on various websites, including the Louvre and the Royal Ontario Museum websites, and proposed a model for evaluating information currency, information richness, quality of information organization, and multilinguisticity of museum websites. In an effort to develop a general website evaluation framework, Chiou, Lin and Peng (2010) analyzed the trends in 83 website evaluation articles published between 1995 and 2006. The authors found that in the wake of usability testing, museums often overlooked web strategy as the means of conveying museums' missions and

objectives through their web presence. To ensure consistency between web strategy and website presence, the authors proposed a five-stage process to support internal evaluation of websites, including identifying the website strategy, developing an evaluation scheme, running the evaluation, computing scores, and analyzing the data.

While evaluation frameworks like MUSEF and MiLE offer the means to test the quality of existing sites, few studies take a step further and offer specific recommendations for design of museum websites. One early example of such a study is the work of Dyson and Moran (2001) who analyzed seven websites with the aim of generating recommendations for the design of the Rural History Centre website at the University of Reading as well as developing a general framework for evaluating digital visual collections. The study developed a set of recommendations for design of museum websites, including the importance of guided browsing and clear and consistent website architecture. While the MiLE framework developed by the Bologna Group (Blas et al., 2002) did not include design recommendations, the authors hoped the framework could be used to prescribe functions for museum websites. More recently, Lin, Gregor and Ewing (2009) and Lin, Fernandez, and Gregor (2012) proposed a website evaluation framework that focused on assessing user enjoyment associated with the use of virtual museum collections. The authors attempted to evaluate enjoyment associated with multi-media content, interactive features (e.g., games), and proposed incorporating the principles of “stickiness,” storyline, mood building, fun and social interactions in museum website designs.

In an effort to extend previous work on developing general recommendations for the design of art museum websites, we developed a study to evaluate several museum website design features, examine the relationships between these features and describe virtual visitors’ initial and overall impressions and their estimated repeat visits to the site. This paper describes the methods and results of the study and provides recommendations for the development and evaluation of art museum websites.

Methods

This study is part of a multi-phase project that aims to develop design recommendations for art museum websites.

During the initial phase, we explored and evaluated the functionality of six digital art collections of various sizes and institutional affiliations using Nielsen’s (1993) evaluation framework criteria:

- Learnability: degree of difficulty in accomplishing basic tasks (e.g., search, browse, review and manipulation of results, etc.) during the first encounter with the interface.
- Efficiency: after gaining familiarity with a system, speed with which basic tasks can be performed.
- Aesthetics: perceived attractiveness of the interface.
- Errors: number and severity of errors encountered while using the system, and ease of recovery from the errors.

- Memorability: after a period of non-use, the degree of difficulty in re-establishing proficiency.
- Reflection: quality and polarity of memories associated with the use of the interface.
- Satisfaction: degree of satisfaction with the interface design based on the overall evaluation of task-specific, aesthetic and other features.

The seven collections that comprised the initial study sample included: 1) Google Art Project 2) Artsy 3) Brooklyn Museum 4) Los Angeles County Museum of Art (LACMA) 5) Museum of Fine Arts Boston (MFA) and 6) the Stedelijk Museum, Amsterdam.

The initial study (Lopatovska et al., 2013) established five categories of design features that users found valuable in their interactions with online collections, including:

1. Search/browse features
2. Image manipulation features
3. Interactive features
4. Website aesthetics
5. Usability

This paper reports on the second phase of the study, which examined relationships between website features within the five outlined above categories (search, image manipulation, interactivity, aesthetics and usability), initial and overall impressions of the website, and likelihood of return visits to the website. We were particularly interested in determining which categories of website features or visitors' initial impressions affected virtual visitors' overall impressions and their estimated likelihood of repeat visits. Figure 1 represents the main study constructs and the relationships investigated.

[Insert Figure 1]

The study was conducted with a group of ten participants: 4 males and 6 females, ages 24 to 53. Since the study was designed as a usability test, a small sample size of participants was considered acceptable (Nielsen, 2000). The study was designed as a course exercise for Information Science students enrolled in a master-level course at Pratt Institute. All participants had high information literacy skills. Most of the participants were art novices with only two students having formal education in art and/or design and architecture.

After piloting the method on several museum websites in 2013 and in early 2014, four websites representing a balance between traditional and innovative website features were selected for the study (i.e., the Tate, the Dallas Museum of Art, the Solomon R. Guggenheim Museum and the Rijksmuseum, Amsterdam). The spring 2014 versions of the four museum websites were analyzed in the study (Appendix B).

For each of the websites, participants were asked to indicate the types of features offered and rate their experiences using a 5-point semantic Likert scale where 1 corresponded to “extremely negative” and 5 corresponded to an “extremely positive” experience. The questionnaire was designed to collect data about the five categories of online features identified during the first study phase, including: 1) search/browse features 2) image manipulation features 3) interactive features 4) aesthetics and 5) usability (the list of specific features within the five broad categories

can be found in Table 1). Since we were interested in examining how these features affect the overall impression of a website and a visitor's estimate of his or her repeat visits, we added these questions to the study instrument. In addition to identifying the presence and rating of a certain feature, participants were also encouraged to provide comments explaining their scores (Table 1 illustrates the types and summary of data collected by the online questionnaire).

All evaluations were completed in a lab during approximately one hour.

The online questionnaire collected two types of data: numeric ratings of various website features and text of participants' comments explaining their ratings. The data were analyzed by applying appropriate statistical and content analysis techniques discussed in the next section.

Results

For the purpose of analysis, participants' numeric responses were grouped into eight broader variables representing the following website characteristics: visitors' initial impression, overall/final impression, likelihood of a repeat visit, aesthetics, search/browse features, usability, image manipulation and interactive features. The composite scores were created for each of the variables by summing up the answers to individual questionnaire items. For example, if a participant rated a particular museum website's Social Media feature as "2," Comment support as "3," Tagging as "3" and Gaming elements as "5," the total Interactive Features score for the website was calculated to be "13."

[Insert Table 1]

Based on participant ratings of museum features, the Rijksmuseum website received the highest average scores on Initial and Overall Impressions, Aesthetics, Search/Browse, and selected Usability features. The Tate received high scores on selected Interactive features and, despite lack of certain features, the DMA received the highest average score on the Likelihood of a repeat visit.

Stepwise, or statistical, regression analysis was used to test whether Overall impression and Likelihood of a repeat site visit scores were predicted by initial impressions, site's aesthetics, usability, search/browse, image manipulation and interactive feature scores.

The results of the stepwise regression¹ suggest that Aesthetics scores were good predictors of Overall Impression scores: the higher participants rated the aesthetic properties of the website, the better overall impressions they had of the website.

¹ Regression analysis is a statistical test used to examine relationships between two or more variables and estimate specific changes to the value of the dependent variable (or 'criterion variable') when one or more of the independent variables is changed (Freedman, 2005). The stepwise regression analysis was chosen to identify independent variables that had the strongest effect on a dependent variable (i.e., Overall impression). The regression test that aimed to explain the variance of the Overall Impression scores resulted in two statistically significant models. The first one explained 52% of the Overall Impression scores: $R^2 = .52$, $F(1, 32) = 32.861$, $p < 0.01$. In this first model, Aesthetics scores explained the largest proportion of variance in the Overall Impression scores: $\beta[\text{beta}] = .72$, $t(5.732) = 5.73$, $p < 0.01$. In this first model, aesthetics scores explained the biggest proportion of variance in the Overall Impression

The regression test that aimed to explain the variance of the estimated Repeat Visits scores resulted in a weak statistically significant model that explained 27% of the Repeat Visits scores: $R^2 = .27$, $F(1, 32) = 11.307$, $p < 0.01$. Scores on Image Manipulation features explained the biggest proportion of variance in the Repeat Visits scores: $\beta[\text{beta}] = .52$, $t(3.363) = 3.36$, $p < 0.01$. Analysis suggested that the Image Manipulation scores were the strongest predictors of estimated Repeat Visit to the website.

The online questionnaire used to collect participant ratings of various website features also collected comments on specific factors contributing to their ratings. The text of participants' comments was coded using a content analysis technique. The three major themes that emerged were related to the sites' navigation ($N=217$), design ($N=99$) and content ($N=40$). The website associated with the highest number of participants' comments was that of the Rijksmuseum ($N=107$), followed by the Guggenheim ($N=89$), Tate ($N=82$), and DMA ($N=78$).

[Insert Table 2]

Polarity of participants' comments was also analyzed by coding text of participants' responses into positive, negative and neutral categories. Appendix A illustrates the distribution and types of participant comments that were mentioned in more than 5% of participants' responses. In evaluating the Rijksmuseum website, the largest number of comments focused on the site's navigation. More specifically, participants were largely positive about the site's browsing and filtered search features, while a regular search received negative comments. Participants' comments also provided positive feedback on the Rijksmuseum's aesthetics (8 positive comments out of the total 10). In evaluating the Guggenheim Museum website, participants were critical of some of the navigation features, including advanced search, refined results and filtered search features. Judging from participants' comments, participants expressed positive or neutral sentiments about the site's aesthetics (clean, "low-risk" design) and mixed feelings about the site's content organization and presentation. In participant comments for the Tate, refined search features received both positive and negative comments from participants, while search result filters received mostly positive feedback. In evaluating the Tate website, participants provided positive feedback on the site's aesthetics and noted a good balance between traditional catalog features and modern stylistics. The DMA website received the lowest number of participants' comments, with participants giving mixed feedback on the site's aesthetics, advanced search features, content organization, and negative feedback on filtered search features, specifically pointing to the insufficient number of filtering options.

Limitations

The study had a number of limitations. The study was conducted with ten participants whose information literacy skills and knowledge of art history did not significantly vary. While small sample sizes are acceptable on multi-phase usability projects (Petrie & Nigel, 2009), a higher number of participants with diverse characteristics would have strengthened the statistical model

scores: $\beta[\text{beta}] = .38$, $t(2.625) = 2.625$, $p < 0.01$. The second model explained the 61% of the Overall Impression scores: $R^2 = .61$, $F(2, 32) = 22.998$, $p < 0.01$. In the second model, the biggest proportion of variance in the Overall Impression scores was explained by the aesthetics scores ($\beta[\text{beta}] = .50$, $t(3.454) = 3.46$, $p < 0.02$) and scores on the Image Manipulation features.

and improved the generalizability of findings. Future studies should aim at recruiting a more diverse pool of participants representing various demographic characteristics, information literacy skills, and knowledge and interest in art. Future research should also consider diversifying a sample of museum websites and including museums of different sizes and types of collections. The study was conducted under controlled conditions that could have influenced participants' interactions with museum websites and evaluations. Future research should expand to naturalistic settings and investigate behavior of museum visitors who have personal interests in the outcomes of their websites' interactions.

Discussion and Conclusion

We conducted a study to examine the relationships between the five groups of website features (search, image manipulation, interactivity, aesthetics and usability), visitors' initial impression of the website, visitors' overall impressions and likelihood of return visits to the website. Since we did not identify considerable support for the importance of specific website features in visitors' impression formation in the prior literature, we initially hypothesized that any of the listed above features and/or initial impressions might influence the visitors' overall impressions and their estimated likelihood of repeat visits (Figure 1). Through the analysis of visitors' feedback on various website features, we were able to determine that website aesthetics had the strongest effect on the overall impression, and image manipulation features were the strongest determinants of the estimated likelihood of repeat visits for the site. (Figure 2 helps illustrate the model of relationships between the website features, visitors' initial and overall impressions, and estimated likelihood of repeat visits).

[Insert Figure 2]

Our finding about the importance of website aesthetics in visitors' formation of overall impressions is consistent with previous studies that showed a strong correlation between website design aesthetics and a visitor's motivation to either return to a museum website or visit the physical museum (Pallud & Straub, 2014). This finding is also supported by studies that suggest users make impressions within the initial seconds of website use, before the user has a chance to explore navigation and content features (Lindgaard, et al., 2006; Tuch et al., 2012), and studies linking aesthetics to the user's perceptions of the website usability (Tractinsky, Katz, & Ikar, 2000). In commenting on the websites' aesthetics, participants mentioned website design colors, navigation and content presentation, and the overall site's "beauty" and pleasantness as factors attributing to their positive or negative judgments. While our findings illustrate the importance of creating aesthetically pleasing websites that can lead to repeat virtual visits or visits to physical collections, more research must be conducted in order to identify specific website features that generate positive aesthetics for target user groups.

We did not find support in the prior literature for relationships between image manipulation features (e.g., ability to save and re-size images) and visitors' estimated likelihood of repeat visits, suggesting that more research into these relationships is needed. It is worth noting that despite the high ratings of the Rijksmuseum website on the quality of its features and aesthetics, the highest estimates for repeat visits were associated with the DMA website. This finding suggests that while participants can be impressed by a website's aesthetics, additional factors, such as rich collection offerings and professional and leisure needs, might contribute to their decision to return to the site.

The site that received the highest scores on most of the participants' evaluations of web features, the Rijksmuseum, also received the highest number of participants' comments on design and usability. The high number of comments related to website design features can also be attributed to the characteristics of the study participants, who were Information Science students. It is possible that a study conducted with a different pool of participants would have produced different findings and would have generated more comments about website content than design. The analysis of the museum-specific comments yielded common patterns in user preferences across the four websites. For example, for all the analyzed websites, participants expressed preferences for simple, clean, and well-balanced sites, and emphasized the importance of advanced and filtered search features.

In the process of determining the relationships between user assessments of various website features, visitors' overall impressions and willingness to return to the site, the study evaluated four museum websites at a specific point in time. Considering that some of the institutions evaluated are in the process re-designing their websites, the study offers a valuable insight into possible improvements and creates a baseline for comparing the current websites designs to the future versions.

Based on the study findings, we offer the following recommendations to museum website developers

- **Aesthetics:** focus on website aesthetics as an important determinant of the user's overall impression. In our study, aesthetics was defined as the perceived attractiveness of the interface (Nielsen, 2000), a concept that is hard to measure and design for. Based on the "subjective" nature of aesthetic judgments, it would be difficult to recommend specific design features (e.g., colors and navigation) that should be incorporated into all museum website designs. We think a more productive approach would be to develop prototype interfaces and evaluate aesthetic responses to these interfaces by various user groups (e.g., novices/experts, younger/older visitors, etc). In our study, participants were asked to identify website features that contributed to their aesthetic judgments. Another way to measure aesthetics is to ask participants to rate the aesthetic dimensions of a website's design, including its perceived clarity, creativeness, sophistication, 'joyfulness' and other aspects (Lavie & Tractinsky, 2004). Regardless of the method chosen, designers should consider measuring aesthetics of specific interface designs to identify features and properties that positively or negatively affect the site's aesthetic appeal.
- **Search:** provide search, advanced search, and results filtering features. While the preferred availability of advanced search and sophisticated filtering by our participants might not indicate that all website visitors have similar preferences, we may assume that if other information retrieval systems (e.g., ArtStor, Artsy, Google Art Project) provide these features, users will expect to find them on museum websites as well.
- **Content manipulation:** enable image saving, printing, re-sizing and other image manipulation features. Our participants expressed a preference for being able to manipulate website content and disappointment when content manipulation was limited. Previous research indicates that providing access to digital content is one of the main functions of museum websites (Blas et al., 2002). Access to content is also one of the main reasons users visit museum websites (Marty, 2008). In light of our findings it may

also be important to enable user control of online content (copyright, technical and other constraints permitting).

- Website improvements and evaluations: continuously improve website design and conduct systematic evaluations to ensure that it meets ever changing user needs and preferences (Mayhew, 1999). For example, our study did not find compelling evidence for the importance of museum sites' interactive features, such as social media sharing and tagging options. This finding might be related to a low availability of these features on the websites analyzed or the profile of the study's participants. Web 2.0 features are gaining popularity and are increasingly being found on museum websites. These features create both new advantages and disadvantages (Nielsen, 2007), so the features best suited to the aims of a particular website (e.g., promotion, crowdsourcing, and user engagement), still need to be tested.

While additional work is needed to develop a more detailed set of museum website design recommendations, our study strengthened previous claims for the importance of aesthetics to museum websites and suggested the importance of content manipulation features to visitors' estimates of repeat visits. Future studies will continue to explore the importance of interactive Web 2.0 features to visitor interaction and their evaluation of museum websites and will expand research into factors contributing to the return visits to museum websites and decisions to visit physical collections.

References

- Barry, A. 2006. "Creating a Virtuous Circle between a Museum's Online and Physical Spaces." Available online at: <http://www.museumsandtheweb.com/mw2006/papers/barry/barry.html>
- Bartindale, T., Clarke, R., Shearer, J., Balaam, M., Wright, P., & Olivier, P. 2011. "Bridging the Gap: Implementing Interaction through Multi-User Design." *CHI EA '11 Extended Abstracts on Human Factors in Computing Systems*, 2071-2076.
- Blas, N.D., Pia, M., Romagna, E., & Paolini, P. 2002. "Evaluating the Features of Museum Websites: the Bologna Report." *Proceedings of Museums and the Web Conference*, Boston, MA.
- Chiou, W.-C., Lin, C.-C., & Perng, C. 2010. "A Strategic Framework for Website Evaluation Based on a Review of the Literature from 1995-2006." *Information and Management*, 47: 282-290.
- Ciolfi, L., Bannon, L. & Fernstrom, M. 2008. "Including Visitor Contributions in Cultural Heritage Installations: Designing for Participation." *Museum Management and Curatorship*, 23 (4) (2008): 353-365.
- Ciolfi, L., & McLoughlin, M. 2012. "Designing for Meaningful Visitor Engagement at a Living History Museum." *NordiCHI'12 Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense through Design*, 69-78.
- Cunliffe, D., Kritou, E., & Tudhope, D. 2001. "Usability Evaluation for Museum Web Sites." *Museum Management and Curatorship* 19 (3): 229-252.

- De Angeli, A., Sutcliffe, A., & Hartmann, J. 2006. *DIS 2006*, June 26–28, 2006, University Park, Pennsylvania, USA.
- Dyson, M. C., & Moran, K. 2001. "Informing the Design of Web Interfaces to Museum Collections." *Museum Management and Curatorship* 18 (4): 391-406.
- Fantoni, S. F., Stein, R., & Bowman, G. 2012, April. "Exploring the Relationship Between Visitor Motivation and Engagement in Online Museum Audiences." *Proceedings from Museums and the Web Conference*, San Diego, CA.
- Freedman, D. A. 2009. *Statistical models: theory and practice*. Cambridge University Press.
- Haynes, J., & Zambonini, D. 2007. "Why Are They Doing That!? How Users Interact With Museum Web Sites." In *Proceedings from Museums and the Web Conference*, edited by J. Trant and D. Bearman. Toronto: Archives & Museum Informatics.
- Honeysett, N. 2007. "Reach More and Earn More: Interacting with Audiences Online." In *The Digital museum: a Think Guide*, edited by H. Din and P. Hecht. Washington DC: American Association of Museums.
- Jones, K. B. 2008. "The Transformation of the Digital Museum." In *Museum informatics: People, Information, and Technology in Museums*, edited by P. Marty & K. Jones, 9-25. New York: Routledge.
- Lavie, T. and Tractinsky, N. 2004. Assessing Dimensions of Perceived Visual Aesthetics of Web Sites, *International Journal of Human-Computer Studies*, 60(3):269-298.
- Lin, A., Gregor, S. & Ewing, M. 2009. "Understanding the Nature of Online Emotional Experiences: a Study of Enjoyment as a Web Experience." *ICEC '09 Proceedings of the 11th International Conference on Electronic Commerce*, 259-268.
- Lin, C.H., Fernandez, W., & Gregor, S. 2012. "Understanding Web Enjoyment Experiences and Informal Learning: a Study in a Museum Context." *Decision Support System*, 53: 846-858.
- Lindgaard, G., Fernandes, G., Dudek, C., & Brown, J. 2006. "Attention Web designers: You Have 50 Milliseconds to Make a Good First Impression!" *Behavior & Information Technology*, 25(2): 115-126.
- Marty, P.F. 2007. "Museum Websites and Museum Visitors: Before and After the Museum Visit." *Museum Management and Curatorship* 22 (4): 337-360.
- Mayhew, D. J. 1999. "The usability engineering lifecycle." *CHI'99 Extended Abstracts on Human Factors in Computing Systems*. ACM: 147-148.
- Nielsen, J. 1993. *Usability Engineering*. Boston, MA: Academic Press.

- Nielsen, J. (2000, March 19). "Why You Only Need to Test with 5 Users." *Nielsen Norman Group*. Available online at: <http://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/>.
- Nielsen, J. 2007 "Web 2.0 can be dangerous." *Jakob Nielsen's Alertbox*.
- Pallas, J. & Economides, A.A. 2008. "Evaluation of Art Museums' Web Sites Worldwide." *Information Services & Use* 28, 45-57.
- Pallud, J. & Straub, D.W. 2014. "Effective Website Design for Experience-Influenced Environments: The Case of High Culture Museums." *Information Management*, 51: 359-373.
- Petrie, H. & Nigel, B. 2009. "The evaluation of accessibility, usability and user experience." *The universal access handbook*: 10-20.
- Stewart, T. S., & Marcketti, S. B. 2012. "Textiles, Dress, and Fashion Museum Website Development: Practices and Strategies." *Museum Management and Curatorship* 28(2): 523-538.
- Tractinsky, N., Katz, A. S., & Ikar, D. 2000. "What is Beautiful is Usable." *Interacting with Computers* 13: 127-145.
- Tuch, A., Roth, S., Hornbaek, K., Opwis, K., & Bargas-Avila, J. 2012. "Is Beautiful Really Usable? Toward Understanding the Relation Between Usability, Aesthetics, and Affect in HCI." *Computers in Human Behavior* 28(2012): 1596-1607.
- Tuch, A. N., Presslauer, E. E, Stocklin, M., Opwis, K., Bargas-Avila, J. A. 2012. "The Role of Visual Complexity and Prototypicality Regarding First Impressions of Websites: Working Towards Understanding Aesthetic Judgments." *International Journal of Human-Computer Studies* 70 (11): 794-811.
- Van Dijk, E., Lingnau, A., & Kockelkorn, H. 2012. "Measuring Enjoyment of an Interactive Museum Experience." *ICM '12 Proceedings of the 14th ACM international conference on Multimodal interaction*, 249-256.
- Walz, J., & Brenny, B. 2005. "Responding to Diverse User Groups: Usability Testing of the Web Interface of the Visual Resources Collection." *VRA Bulletin*, 31(2): 48-49.

Appendix A. List of the most frequently mentioned themes in participants' comments.

Broad Theme	Narrow Theme	Number of Comments	N of <u>positive</u> mentions	N of <u>negative</u> mentions	N of <u>neutral</u> mentions	Example Text
Rijksmuseum						
Navigation	Searching	8	1	7	0	"the search page is nothing special; in fact, the default endless scroll with pretty large thumb nails is a bit overwhelming."
Navigation	Filtered Search	12	12	0	0	"the explore pages are nice, i like the colored labels. also, the subject pages for each subject, material, or art style are a nice introduction."
Navigation	Browsing	6	6	0	0	"Browsing the collection was simple because it uses terms that I would use in my daily speech."
Design	Aesthetics	10	8	2	0	"aesthetically it doesn't seem to work for me, the images are decent but hard to find."
Guggenheim						
Navigation	Advanced Search	7	0	7	0	"There is no advanced search available"

Navigation	Filtered Search	7	1	6	0	“It seems as if you can't use multiple filters, which isn't that great.”
Navigation	Refined Search	9	0	8	1	“I didn't have an option to do this.”
Design	Aesthetics	5	1	0	4	“They have a clean website it just does not stand out. To gain an advantage in web traffic they need changes, to make risks to stand out.”
Design	Organization	8	4	4	0	“Organization/presentation of information seem to come across very clear.”
Tate						
Navigation	Advanced Search	11	1	10	0	“there doesn't seem to be a separate advanced search page”
Navigation	Filtered Search	9	5	3	1	“I found this feature easy to use and very helpful in sorting and narrowing results.”
Navigation	Search Filters	6	6	0	0	“the granularity of topics is fantastic. i like that i can search for all the art that features raised hands, or featuring certain flowers or themes”

Navigation	Refined Search	7	4	3	0	“i can save the search and then search within the search, which is good. don't see color filters though”
Design	Aesthetics	12	9	1	2	“this is a nice balance of modern tech styling and traditional catalog conventions”
DMA						
Navigation	Advanced Search	7	3	3	1	“no advanced search”
Navigation	Filtered Search	6	2	4	0	“all i can do is sort alphabetical by title and maybe by date? not even by related people.”
Design	Aesthetics	9	2	4	3	“It was very aesthetically pleasing”
Design	Organization	7	3	3	1	“It is very well organized and the collection seems to be organized in such a way that anyone can browse.”

Appendix B

Figure 3. Main page of Tate museum

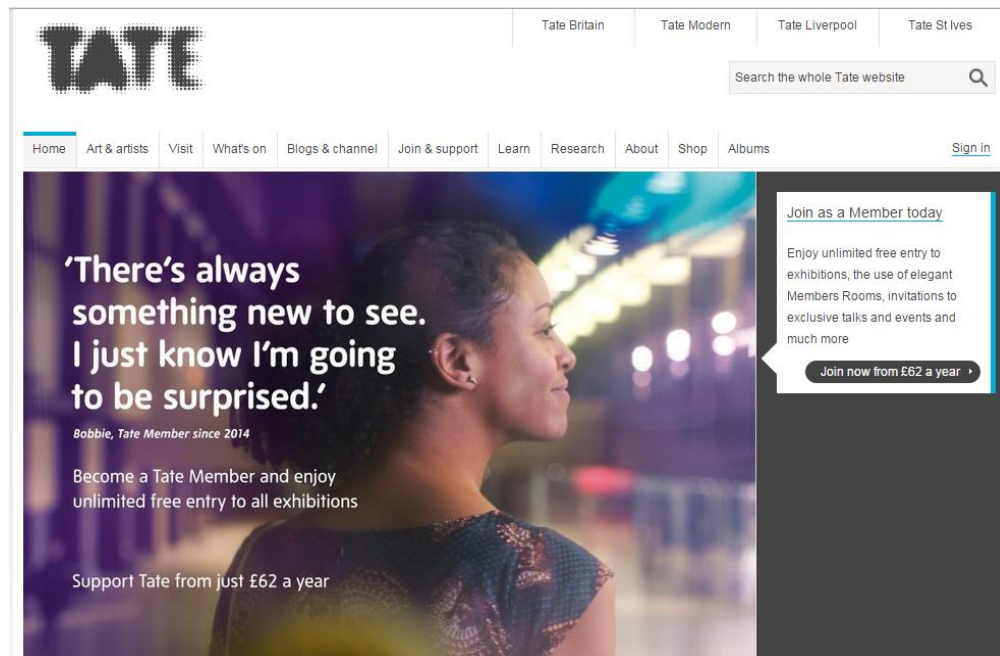


Figure 4. Main page of the Dallas Museum of Art museum

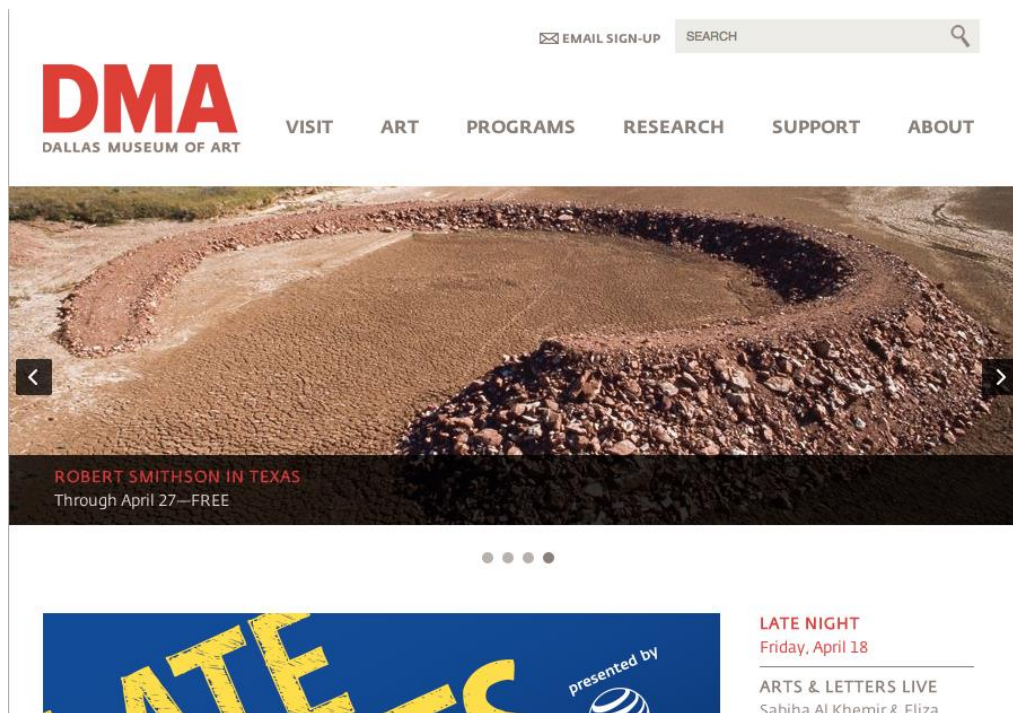


Figure 5. Main page of the Guggenheim museum



Figure 6. Main page of the Rijksmuseum

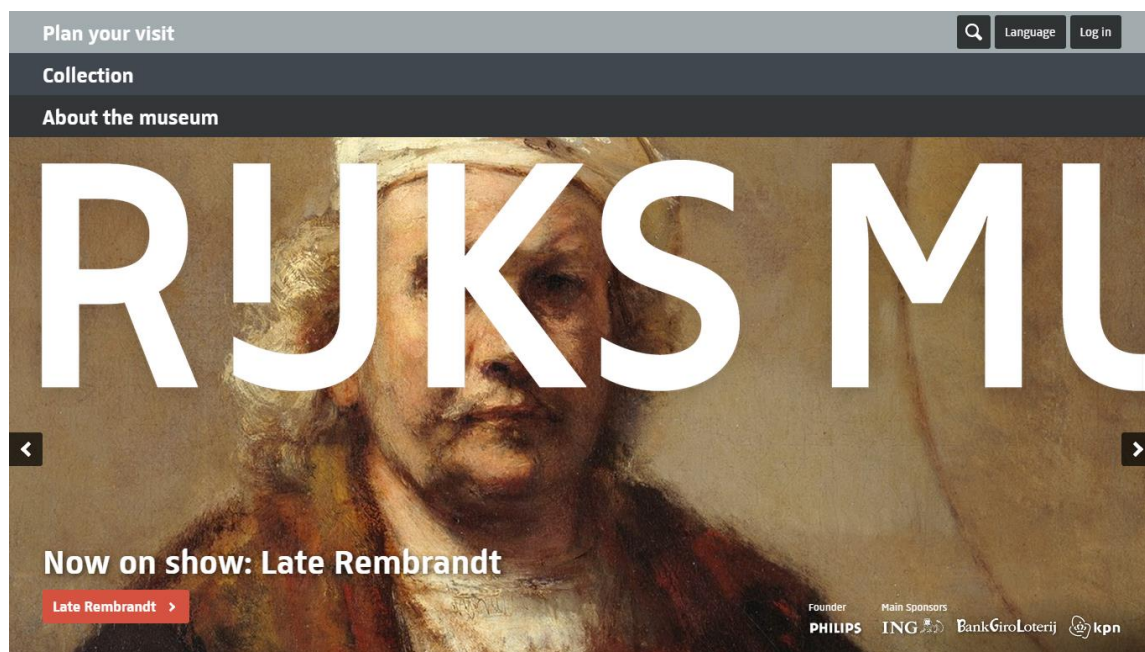


Table 1. Website features and their minimum, maximum and average ratings per museum.

Evaluation criteria\museum site	DMA min-max rating average rating	Rijksmuseum min-max rating average rating	Tate min-max rating average rating	Guggenheim min-max rating average rating
Initial impression (5 = extremely positive, 1 = extremely negative)	3-4 3.43	3-5 4.36	3-5 4.14	3-5 3.7
Aesthetics (5 = best, 1 = lowest)	2-5 2.71	3-5 4.36	4-5 4.13	3-5 3.4
Search/Browse features				
Browsing function	2-5 3.71	3-5 4.78	2-5 4.00	2-5 4.2
Keyword searching	2-5 3.29	2-5 4.22	3-5 4.50	3-5 3.9
Advanced searching	1-5 2.86	2-5 3.44	1-5 2.29	1 1
Filter searching	1-5 2.40	2-5 4.22	1-5 3.63	1-5 2.43
Refinement of search results	1-3 1.75	3-5 4.29	1-5 3.25	1-2 1.29
Usability				
Navigation Intuitiveness	4-5 4.28	4-5 4.36	3-4 3.70	1-4 3
Contact Info/Feedback Option	present	present	present	present
Text clarity	4-5 4.57	4-5 4.90	3-4 3.25	2-5 3.6
Handling of the spelling errors	n/a	0-1 0.22	0-1 0.38	0-1 0.3
Search history	0-1 0.14	0-0.5 0.28	0-1 0.44	n/a
Personalization	0-1 0.14	0-1 0.89	0-1 0.50	n/a
Image Manipulation				
Image zoom	n/a	5-5 5	n/a	n/a
Image size manipulation	n/a	3-5 4.75	2-4 2.75	n/a
Image download	n/a	3-5	3-4	2-4

		4.67	3.50	3
Image printing	n/a	3-5 4.60	3-4 3.50	1-5 3.38
Image saving	n/a	3-5 4.67	3-5 4.00	1-5 3.17
Interactive Features				
Social Media sharing	n/a	3-5 4.78	5-5 5.00	3 3
Comments support	n/a	0-1 0.22	n/a	n/a
Tagging	n/a	0-1 0.89	n/a	n/a
Gaming elements	n/a	0-1 0.11	0-1 0.25	n/a
Likelihood of a repeated site visit	1-5 2.71	0-1 0.83	0-1 0.75	1-2 0.45
Overall impression	1-4 2.86	2-5 4.50	3-5 4.00	2-4 2.8

Table 2. Themes in participants' comments.

	Categories of website feature discussed in comments and Most frequently mentioned features and polarity		
Museum Website	Navigation (N=217)	Design (N=99)	Content (N=40)
Rijksmuseum m (N=107)	67	33	7
Guggenheim (N=89)	54	22	13
Tate (N=82)	52	22	8
DMA (N=78)	44	22	12

Figure 1. Investigated study constructs and relationships between them.

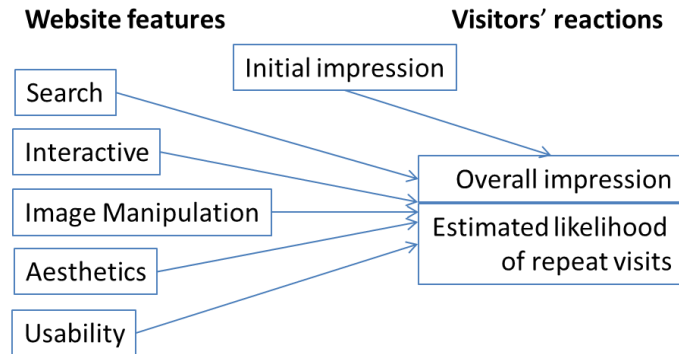


Figure 2. Relationships between study constructs.

