Chapter 8

Applying the Web Interface Profiles: Example Web Site Assessment

8.1 Introduction

This chapter describes the use of the profiles developed in Chapter 6 to assess and improve the quality of an example Web site. The intent of this chapter is three-fold: 1. to demonstrate how the models can be systematically applied to this problem; 2. to illustrate the type of design changes informed by the models and how they vary across models; and 3. to highlight current limitations of the models. The example assessment closely follows the evaluation scenario depicted in Chapter 4, which is the overarching goal of the work in this dissertation. Currently, interpreting model predictions and determining appropriate design changes is a manual process; future work will focus on automating recommendations for improving designs as well as implementing these recommendations. Identifying comparable good designs to aide in site improvements will also be incorporated in future work.

8.2 The Example Site

Figures 8.1–8.3 show three pages taken from a small (nine page) site in the Yahoo Education/Health category. The site provides information about training programs offered to educators, parents, and children on numerous health issues, including leukemia and cerebral palsy. The site, which was not included in the profile development sample (see Chapter 6), was selected because on first glance it appears to have good features, such as clear and sharp images and a consistent page layout, but on further inspection it seemed to have problems. The site assessment focused on answering the following questions.

- Is this a high-quality site? Why or why not?
- Are these high-quality pages? Why or why not?
- What can be done to improve the quality of this site?

The first step was to download a representative set of pages from the site. For this particular site, only eight level-one pages were accessible, and no level-two pages were reachable, for a total of 9 downloaded pages. Although there is a page containing links (Figure 8.2), the links are to pages external to the site.

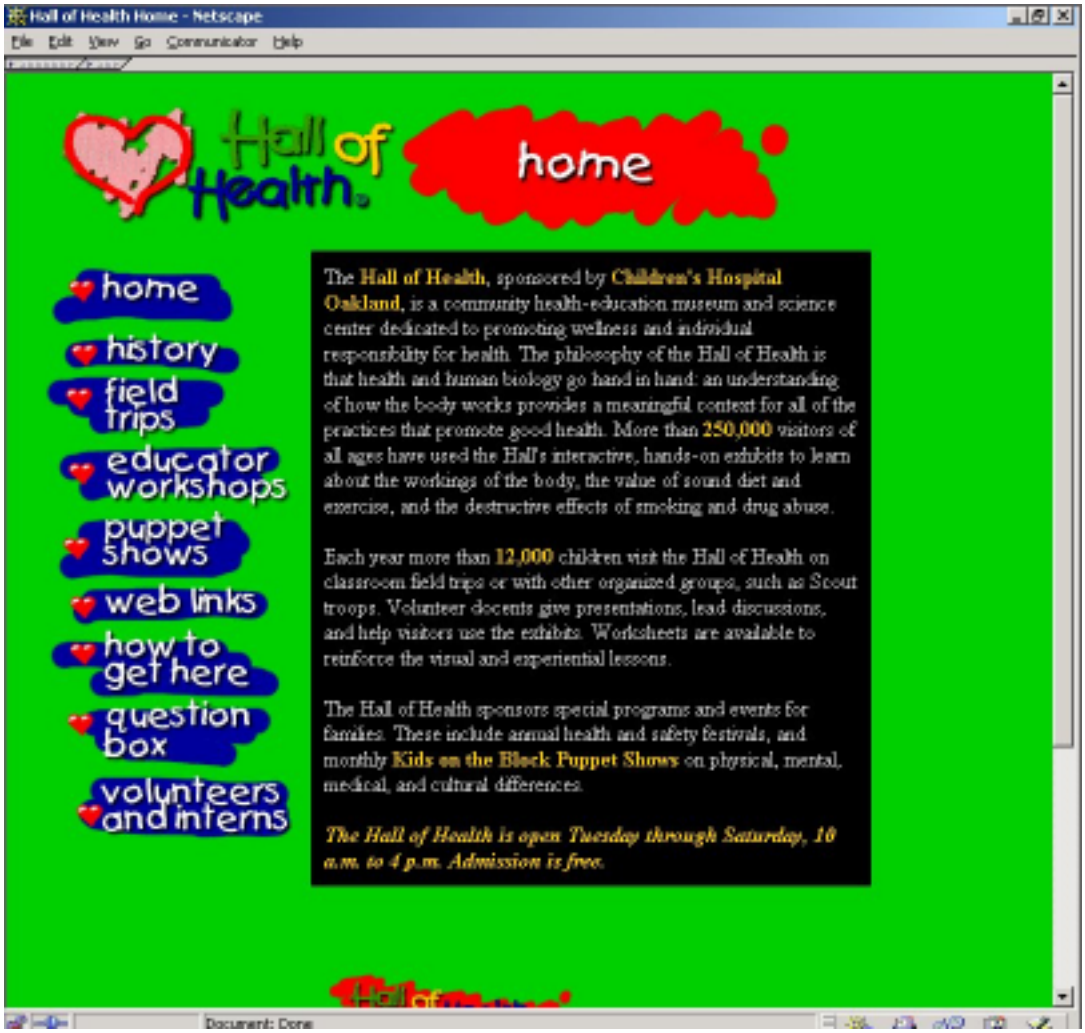

Figure 8.1: Home page taken from the example health education site (http://www.hallofhealth.org/home.html; September 14, 2001).

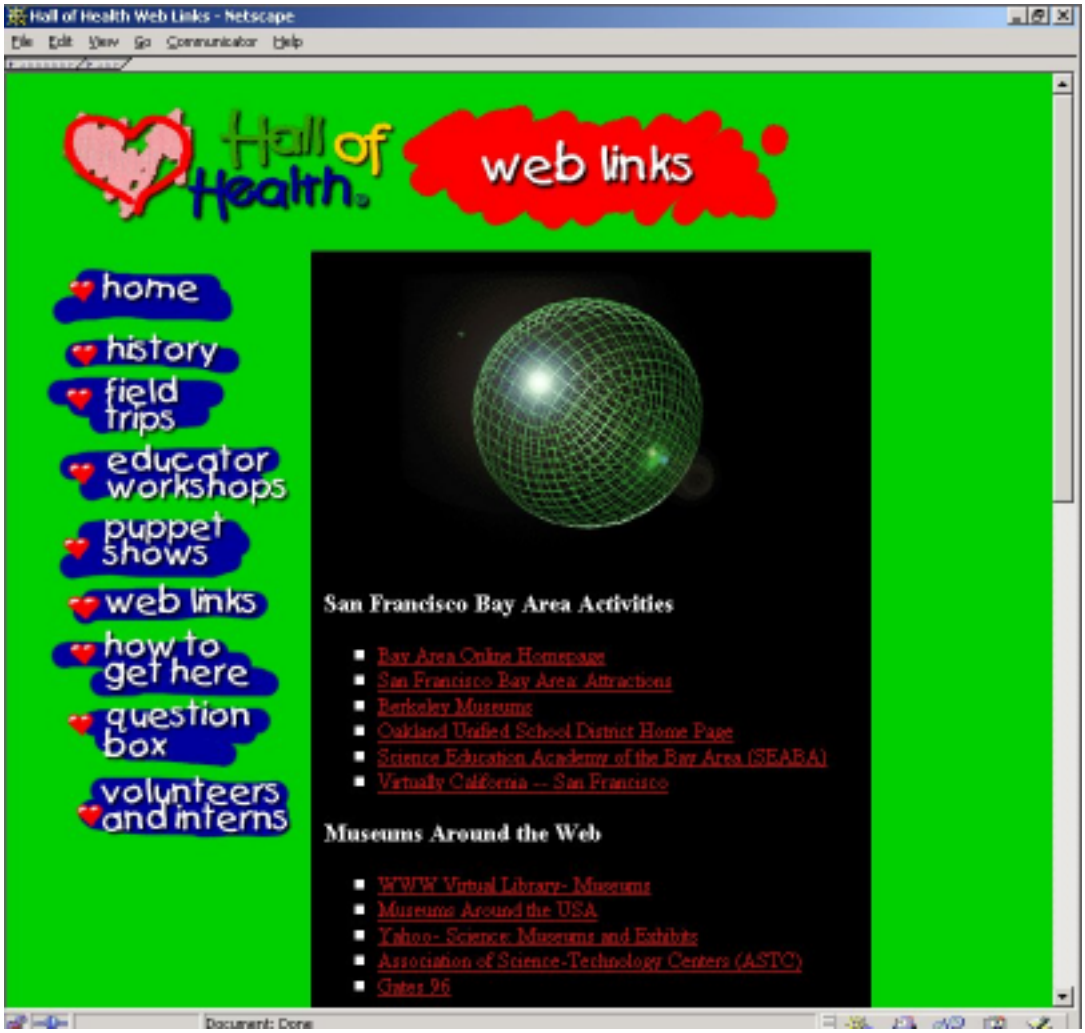

Figure 8.2: Link page taken from the example health education site (http://www.hallofhealth.org/weblinks.html; September 14, 2001).

Figure 8.3: Content page taken from the example health education site (http://www.hallofhealth.org/puppetshows.html; September 14, 2001).

The next step was to use the Analysis Tool (see Chapter 4) to compute site-level and page-level measures and to apply the models to individual pages and to the site as a whole. Each model encapsulates relationships between key predictor measures and can be used to (i) generate quality predictions and (ii) determine how pages and sites are consistent with or deviate from good pages and sites.

In the discussions below, when decision tree rules are used to generate predictions, the consequences are interpreted manually. When cluster models are applied, the score for each measure on an individual page is compared to that of the cluster centroid, and if the measure differs by more than one standard deviation unit from the centroid, the measure is reported as being inconsistent with the cluster. Cluster deviations are also interpreted manually.

8.3 Site-Level Assessment

The example site can be classified in both the health and education content categories, so the site-level decision tree model was run initially without differentiating by content category. The site-level model predicted that the site is similar to poor sites overall; the median page quality prediction (i.e., median computed over the overall page quality model's predictions for the nine pages; poor) is consistent with the overall site quality model's prediction. The corresponding decision tree rule (top of Figure 8.4) reveals that the site has an unacceptable amount of variation in link elements (31%), although variation for other site-level measures is acceptable. The combination of the link element variation and the lack of a comparable overall element variation violates patterns discovered on good sites.

The major source of link element variation is the text link count. Eight out of nine pages have from two to four text links; the remaining page has 27 text links, and acts as a links page (see Figure 8.2). The decision tree rule suggests that a link element variation level below 29% is typical on good sites. One interpretation of this finding is that good sites strive to keep the navigation structure consistent among pages and may even distribute links over multiple pages to maintain this consistency. Hence, the rule may indicate the need to similarly redistribute the links on this page.

Site quality was also assessed according to the two applicable content categories – health and education. The decision tree for health sites predicted that this is a poor health site (middle of Figure 8.4). In this case the problem is inadequate text element variation. Most of the pages on the site contain paragraphs of text without headings and use only one font face (serif); this may actually make it harder for users to scan the page to find the information they are looking for [Nielsen 2000; Spool *et al.* 1999]. The median health page quality prediction (poor) is consistent with the health site prediction.

The decision tree for education sites made a prediction contrary to that for sites overall and health sites; it found this site to be consistent with good education sites (bottom of Figure 8.4). Good health and good education sites are similar with respect to graphic formatting variation, but are quite different on the other measures, which is the cause for this disparity. However, as will be discussed below, the median education page quality is poor.

8.4 Page-Level Assessment

The decision tree model for predicting page quality reports that all nine of the pages are consistent with poor pages. The home page (Figure 8.1) contains seventeen italicized body words; pages with more than two italicized body words are considered poor pages in the model (see rule

Overall Site Quality

if ((Page Performance Variation is missing OR (Page Performance Variation $\leq$ 90.2)) AND (Overall Variation is not missing AND (Overall Variation $\leq$ 14.49)) AND (Link Element Variation is missing OR (Link Element Variation $>$ 29.195)) AND (Overall Element Variation is missing OR (Overall Element Variation $\leq$ 26.07)))

Class = Poor

This rule classifies the site as poor because the pages have acceptable page performance, overall, and overall element variation, but they have more than 29.2% variation in link elements (30.68%).

Health Site Quality

if ((Graphic Element Variation is not missing AND (Graphic Element Variation $\leq$ 32.695)) AND (Text Element Variation is missing OR (Text Element Variation $>$ 47.45)) AND (Text Element Variation is missing OR (Text Element Variation $\leq$ 92.25)))

Class = Poor

This rule classifies the site as poor because the pages have acceptable graphic element variation, but they have between 47.45% and 92.25% variation in text elements (53.18%).

Education Site Quality

if ((Median Page Breadth is missing OR (Median Page Breadth $\leq$ 11.25)) AND (Page Title Variation is missing OR (Page Title Variation $\leq$ 196.7)) AND (Page Formatting Variation is missing OR (Page Formatting Variation $\leq$ 27.785)) AND (Page Title Variation is missing OR (Page Title Variation $\leq$ 132.495)) AND (Graphic Formatting Variation is not missing AND (Graphic Formatting Variation $\leq$ 16.165)))

Class = Good

This rule classifies the site as good due to an acceptable combination of measures: the median page breadth (8) is less than twelve; and pages in the site have very little similarity in page titles (37.5%), page formatting variation (0%), and graphic formatting variation (3.19%).

Figure 8.4: Decision tree rules reported for the example health education site. The rules were reported by the overall (top), health (middle), and education (bottom) site quality models.

Home Page

if ((Italicized Body Word Count is not missing AND (Italicized Body Word Count > 2.5)))
    Class = Poor

This rule classifies the home page as poor because it contains more than two italicized words (17) in the body text.

Link and Content Page

if ((Italicized Body Word Count is missing OR (Italicized Body Word Count $\leq$ 2.5)) AND (Minimum Font Size is not missing AND (Minimum Font Size > 9.5)) AND (Minimum Graphic Height is missing OR (Minimum Graphic Height $\leq$ 36)) AND (Minimum Color Use is not missing AND (Minimum Color Use > 15.5)))
    Class = Poor

This rule classifies both the link and content pages as poor because they contain an acceptable number of italicized words in the body text and contain at least one image with a height less 37 pixels, but all of the text is formatted with a font greater than 9pt and all of the colors are used more than fifteen times. Recall that good pages tend to use a font smaller than 9pt typically for copyright text, and they use an accent color (see Section 6.6).

Figure 8.5: Decision tree rules reported for the three example pages. These rules were reported by the overall page quality model.

at the top of Figure 8.5). Schriver [1997] suggests that italicized text should be avoided because it is harder to read on computer screens than in printed documents.

Recall from Section 5.10.1 that a minimum color count metric was developed to track the number of times each color is used on a page and to report the minimum number of times a color is used; this measure detects the use of an accent or sparsely-used color. The content page (Figure 8.3) is classified as poor mainly because the minimum number of times a color is used is sixteen and all of the text, including the copyright text at the bottom of the page, is formatted with a font greater than 9pt (bottom of Figure 8.5). Good pages tend to have an accent color that they use sparingly, whereas poor pages seem to overuse accent colors (see Section 6.6). Good pages also tend to use a smaller font size for copyright or footer text unlike poor pages. Additionally, the example content page contains 34 colored body text words, which is twice the average number found on good pages; in the extreme case, a large number of colored words could result in the uncolored words standing out more so than the colored words. The same prediction and decision tree rule is reported for the link page.

To gain more insight about ways to improve page quality, each page was mapped into one of the three clusters of good pages – small-page, large-page, and formatted-page. All of the pages map into the small-page cluster and are far from the cluster centroid (median distance of 10.9 standard deviation units); the page closest to the center of this cluster has a distance of 4.0 standard deviation units (see Section 6.7). Pages in the example site deviate on key measures that distinguish pages in this cluster, including the graphic ad, text link, link text cluster, interactive object, and link word counts. Table 8.1 summarizes, for the sample content page, the ten key measures (i.e., measures that play a major role in distinguishing pages in this cluster) that deviate

Measure	Value	Cluster Range
Vertical Scrolls	2.0	(0.56–2.00)
Text Column Count	5.0	(0.62–4.36)
All Page Text Terms	129.0	(138.59–353.24)
Link Count	12.0	(12.40–41.24)
Text Link Count	2.0	(4.97–27.98)
Good Link Word Count	3.0	(7.43–49.67)
Bobby Browser Errors	6.0	(7.54–14.99)
Font Count	6.0	(3.64–5.80)
Sans Serif Word Count	0.0	(13.91–253.57)
Display Word Count	33.0	(1.13–18.67)

Table 8.1: Top ten measures that deviate from the small-page cluster for the example content page. The measures are presented in their order of importance. Each range reflects one standard deviation unit around the metric value at the cluster centroid. The page's measures are 8.33 standard deviation units from the cluster centroid.

from the cluster centroid; deviations are similar for other pages in the site. Most of these deviations, including two of the top ten measures (text link count and good link word count), can be attributed to the fact that the site provides predominately graphical links instead of text links for navigation. Table 8.1 also shows deviation on the page height (vertical scrolls), the use of words formatted with sans serif fonts (san serif word count), and the overall use of fonts (font count – combinations of a font face, size, bolding, and italics).

The quality of these pages was also evaluated using the more context-sensitive page quality models for health and education pages (as opposed to the overall model). All but two of the pages were predicted to be poor health pages, which mirrors the results of the site-level model. However, all of the pages were also predicted to be poor education pages, contrasting with the site-level model. In both cases, predictions were based on the features mentioned above. Table 8.2 summarizes the top ten measures that deviate from the two models and shows that there is some similarity between the two sets of measures, especially for measures related to text links (text link, link word, and good link word counts).

The contrast between site-level and page-level predictions demonstrate the need to incorporate page-level predictions into the site-level prediction. For example, a site can only be considered a good site if the site-level measures are consistent with good sites AND most of the pages are consistent with good pages. At the site level, the example site was highly consistent on page formatting, graphic formatting, and page performance; however, the page quality predictions show that several design aspects, such as text formatting and link elements, need to be improved. If the site-level model for education sites incorporated page-level measures, then this site would be considered a poor education site. Considering the median page quality predictions in conjunction with site quality predictions is one way to mitigate this limitation.

Finally, the quality of these pages was evaluated using the models for each page type – home, link, content, form, and other. The page type decision tree made accurate predictions for six of the nine pages, but inaccurately predicted that three pages were consistent with link pages; visual inspection suggested that these pages were actually content pages. As shown in Figure 8.6, the mispredictions were mainly due to an improper balance of link, body, and display text stemming from an overuse of image links. After correcting the page type predictions, all nine of the pages were classified as poor pages. Table 8.3 summarizes the top ten measures that deviate on

Health Page Quality			**Education Page Quality**		
Measure	Value	Model Range	Measure	Value	Model Range
Weblint Errors	0.0	(6.06–82.82)	Bobby Priority 2 Errors	3.0	(3.09–5.41)
Internal Link Count	10.0	(16.16–69.36)	Bobby Browser Errors	6.0	(9.08–20.28)
Bobby Browser Errors	6.0	(6.63–22.49)	Minimum Font Size	10.0	(8.88–9.10)
Link Count	12.0	(18.81–76.87)	Minimum Color Use	16.0	(0.20–6.22)
Redundant Link Count	1.0	(1.30–19.62)	Fixed Page Width Use	0.0	(0.42–1.20)
Graphic Pixels	224.1K	(80.6K–214.8K)	Minimum Graphic Height	32.0	(0.00–21.45)
Text Link Count	2.0	(3.41–53.91)	Text Positioning Count	14.0	(0.00–4.44)
Good Link Word Count	3.0	(6.23–98.25)	Text Link Count	2.0	(5.69–43.07)
Link Word Count	4.0	(6.58–136.78)	Link Word Count	4.0	(7.66–112.22)
Text Positioning Count	14.0	(1.03–7.11)	Good Link Word Count	3.0	(6.10–81.16)

Table 8.2: Top ten measures that deviate from the health and education page quality models for the example content page. The measures are presented in their order of importance. Each range reflects one standard deviation unit around the mean in the model.

if ((All Page Text Score is missing OR (All Page Text Score > 10.5)) AND (Good Body Word Count is not missing AND (Good Body Word Count ≤ 86.5)) AND (Link Word Count is missing OR (Link Word Count ≤ 73)) AND (Interactive Object Count is missing OR (Interactive Object Count ≤ 2.5)) AND (Good Body Word Count is missing OR (Good Body Word Count > 22.5)) AND (Good Text Color Combination is missing OR (Good Text Color Combination ≤ 5.5)) AND (Good Display Word Count is not missing AND (Good Display Word Count ≤ 0.5)) AND (Graphic Bytes is missing OR (Graphic Bytes > 38540)))
PageType = Link

This rule classifies a page as a link page because it has some similarity in content with the source page and contains few interactive objects and good text color combinations, but it also contains few link, good body, and good display (heading) words and more than 38 Kbytes for images. In other words, the page is dominated by images, image links in particular, and has inadequate text and text links.

Figure 8.6: Decision tree rule that mispredicts content pages to be link pages. This rule was returned for pages other than the three discussed in this section.

the sample content page. This model reports several deviations that were also reported by other models, including the minimum font size, minimum color use, sans serif word count, and text link count.

8.5 Summary of Assessment Findings

Tables 8.4 and 8.5 summarize the measures reported as being inconsistent for the individual pages and the site overall. Most of these measures were discussed above; however, some of the page-level measures were reported for pages other than the three example pages. Several page-level measures were reported as being inconsistent by over half of the models, including the link word and good link word counts, text link count, minimum font size, minimum color use, and Bobby browser errors. In addition, the text and link element variation measures were reported as being inconsistent at the site level.

The models provide some direct insight for resolving design issues associated with some of the measures. For example, decision tree rules reported by the overall page quality model indicate inconsistent measures with a > in the threshold (e.g., italicized body word count > 2.5); they also indicate consistent measures with a < in the threshold. In these cases, the designer could explore ways to reduce measure values below reported thresholds, such as removing italics or text coloring, changing font sizes, breaking text into multiple column, etc. The same guidance holds for the other decision tree models. Similar to decision tree rules, the cluster and discriminant classification models also provide ranges for acceptable metric values, and they report the top ten measures that deviate from the underlying models. Some of the model deviations are straightforward to correct, provided the designer understands the model output and relevant measures. Other model deviations are not as straightforward to correct, such as introducing additional links and content or reducing the reading complexity. Future work on automating design changes should make it easier to interpret and use the models to improve designs.

Based on the patterns reflected in Tables 8.4 and 8.5 and the observations generated by the analysis discussed above, a list of possible ways to improve the site was derived. The changes below are ordered based on their potential impact (how much they mitigate measures that were frequently

Measure	Value	Model Range
Minimum Font Size	10.0	(8.76–9.16)
Minimum Color Use	16.0	(0.26–6.42)
Spelling Error Count	0.0	(0.14–2.88)
Good Panel Color Combinations	1.0	(0.00–0.90)
Self Containment	2.0	(0.72–1.78)
Body Color Count	3.0	(0.80–2.68)
Average Graphic Width	207.0	(41.77–191.96)
Sans Serif Word Count	0.0	(20.18–614.98)
Minimum Graphic Height	32.0	(0.00–28.61)
Text Link Count	2.0	(2.75–37.15)

Table 8.3: Top ten measures that deviate from the content page quality model for the example content page. The measures are presented in their order of importance. Each range reflects one standard deviation unit around the mean in the model.

reported as being inconsistent). The recommendations only apply to the results generated during the initial application of the models; subsequent model applications revealed further changes that are not discussed here. No recommendations are made to address the accessibility and Weblint errors, since the roles of these measures in improving design quality are unclear. Specific changes made as well as the results of the changes are discussed in the next section.

1. Increase the number of text links and corresponding link text (text link, link word, and good link word counts). This will simultaneously increase the total number of links and internal links (link and internal link count) and decrease link element variation.

2. Use a smaller font size for some text, such as the footer text (minimum font size).

3. Decrease color overuse for page text and introduce an accent color (minimum color use).

4. Minimize or eliminate the use of italicized words in body text (italicized body word count).

5. Minimize text positioning (changes from flush left and columns where text starts; text positioning and column counts).

6. Minimize font combinations (font face, size, bolding, and italics combinations; font count).

7. Reduce the sizes of images (average graphic width, minimum graphic height, and graphic pixels).

8. Improve the page layout to reduce vertical scrolling (vertical scrolls).

9. Use tables with explicit widths to control the page layout (fixed page width use).

10. Vary the text elements and the formatting of text elements on the page (text element variation, good body and display word counts, sans serif word count).

11. Reduce the number of colors used for body text (body color count).

Measure	Page-Level								Site-Level		
			CCQ		PTQ					CCQ	
	OQ	Cls	H	E	H	L	C	Freq.	OQ	H	E
Text Element Measures											
Good Body Word Count							√	14.29%			
Display Word Count		√						14.29%			
Good Display Word Count							√	14.29%			
Link Word Count			√	√	√		√	57.14%			
Good Link Word Count		√	√	√	√			57.14%			
Spelling Error Count							√	14.29%			
Link Element Measures											
Text Link Count		√	√	√			√	57.14%			
Link Count		√	√		√			42.86%			
Internal Link Count			√		√	√		42.86%			
Redundant Link Count			√					14.29%			
Graphic Element Measures											
Graphic Ad Count					√			14.29%			
Text Formatting Measures											
Italicized Body Word Count	√				√			28.57%			
Sans Serif Word Count		√					√	28.57%			
Minimum Font Size	√			√		√	√	57.14%			
Body Color Count							√	14.29%			
Text Cluster Count						√		14.29%			
Text Column Count		√						14.29%			
Text Positioning Count			√	√				28.57%			
Link Formatting Measures											
No measures reported											
Graphic Formatting Measures											
Average Graphic Width						√	√	28.57%			
Minimum Graphic Height				√			√	28.57%			
Graphic Pixels			√					14.29%			

Table 8.4: Measures reported as being inconsistent with the page-level and site-level models for the example health education site (Table 1 of 2). A √ indicates that a measure was reported as being inconsistent on at least one of the nine pages by at least one of the models. The page-level models include the overall page quality (OQ), small-page cluster (Cls), content category quality (CCQ), and page type quality (PTQ) models. The health (H) and education (E) page models are used. The home (H), link (L), and content (C) page type models are used. The frequency column (Freq.) reflects the total number of times a measure is reported as being inconsistent divided by seven (number of page-level models). The site-level models include the overall site quality (OQ) and content category quality (CCQ) models; the health (H) and education (E) site models are used.

Measure	Page-Level								Site-Level		
			CCQ		PTQ					CCQ	
	OQ	Cls	H	E	H	L	C	Freq.	OQ	H	E
Page Formatting Measures											
Minimum Color Use	√			√		√	√	57.14%			
Good Panel Color Combinations							√	14.29%			
Bad Panel Color Combinations					√			14.29%			
Vertical Scrolls		√						14.29%			
Font Count		√				√		28.57%			
Fixed Page Width Use				√		√		28.57%			
Self Containment							√	14.29%			
Page Performance Measures											
Bobby Priority 2 Errors				√	√			28.57%			
Bobby Browser Errors		√	√	√		√		57.14%			
Weblint Errors			√		√			28.57%			
Graphic Bytes							√	14.29%			
Object Count					√			14.29%			
All Page Text Terms		√				√		28.57%			
All Page Text Score						√		14.29%			
Site Architecture Measures											
Link Element Variation									√		
Text Element Variation										√	

Table 8.5: Measures reported as being inconsistent with the page-level and site-level models for the example health education site (Table 2 of 2). A √ indicates that a measure was reported as being inconsistent on at least one of the nine pages by at least one of the models. The page-level models include the overall page quality (OQ), small-page cluster (Cls), content category quality (CCQ), and page type quality (PTQ) models. The health (H) and education (E) page models are used. The home (H), link (L), and content (C) page type models are used. The frequency column (Freq.) reflects the total number of times a measure is reported as being inconsistent divided by seven (number of page-level models). The site-level models include the overall site quality (OQ) and content category quality (CCQ) models; the health (H) and education (E) site models are used.

8.6 Improving the Site

Although the example site is somewhat aesthetically pleasing and highly consistent across pages within the site, the individual pages and the site as a whole are classified as being of poor quality. The pages were modified to incorporate a subset of the recommendations discussed above.

- To improve the color and text link counts and simultaneously reduce the link count variation, a link text cluster (i.e., an area of text links shaded with a different background color to make it stand out) was added as a footer at the bottom of each page; the text links in the cluster mirror the content of the graphical links. It was not necessary to split the link page into multiple pages, because adding the footer decreased the link element variation from 31% to 7%.

- To improve text formatting and the text element variation score: headings were added to break up paragraphs; additional font variations were used – Arial font (sans serif) for body text and Trebuchet (serif) for headings; and the font size of the copyright text was reduced to 9pt. The color of headings was also changed to gold for consistency with the models. All of these changes were implemented via an internal stylesheet; the stylesheet also improved the self-containment scores.

- To improve the emphasized (i.e., bolded, colored, italicized, etc.) body text scores, italics and colors within body text were converted to bold, uncolored body text on all pages. Colored, non-italicized body text was also converted to uncolored body text.

- To improve the minimum color usage scores, a color accent was added to the vertical bars between the text links in the footer of each page. A browser-safe color was selected as dictated by a subsequent prediction by the overall page quality model.

- To reduce vertical scrolling, the logo and copyright notice at the bottom of the pages was placed adjacent to each other in one table row. The sizes of images and borders around them were also reduced to improve space utilization. Furthermore, text was wrapped to the left of the images versus images not being inlined with text.

- To further improve the page layout, fixed widths (640 pixels) were used for the main layout table.

Figures 8.7–8.9 depict the revised pages corresponding to the pages in Figures 8.1–8.3; many of the changes are not visible since they appear at the bottom of the pages. Furthermore, only a subset of the potential changes were implemented. Appendix D provides side-by-side comparisons of the original and modified versions of the three pages.

After making these changes, all of the pages were classified correctly by functional type, and they were rated as good pages overall as well as good health pages. Figure 8.10 depicts the complex decision tree rule that classified all of the pages as good overall. The median distance to the small-page cluster was 4.7 as compared to 10.9 standard deviation units for the original pages. Eight pages were rated as average pages based on their functional type; one was rated as poor. In addition, five of the nine pages were rated as average education pages; the four remaining pages were rated as poor. These differences in predictions demonstrate the potential difficulty of satisfying all of the models simultaneously. Hence, a clear design objective needs to be chosen prior to making any changes, since the models could reveal a different set of changes to make.

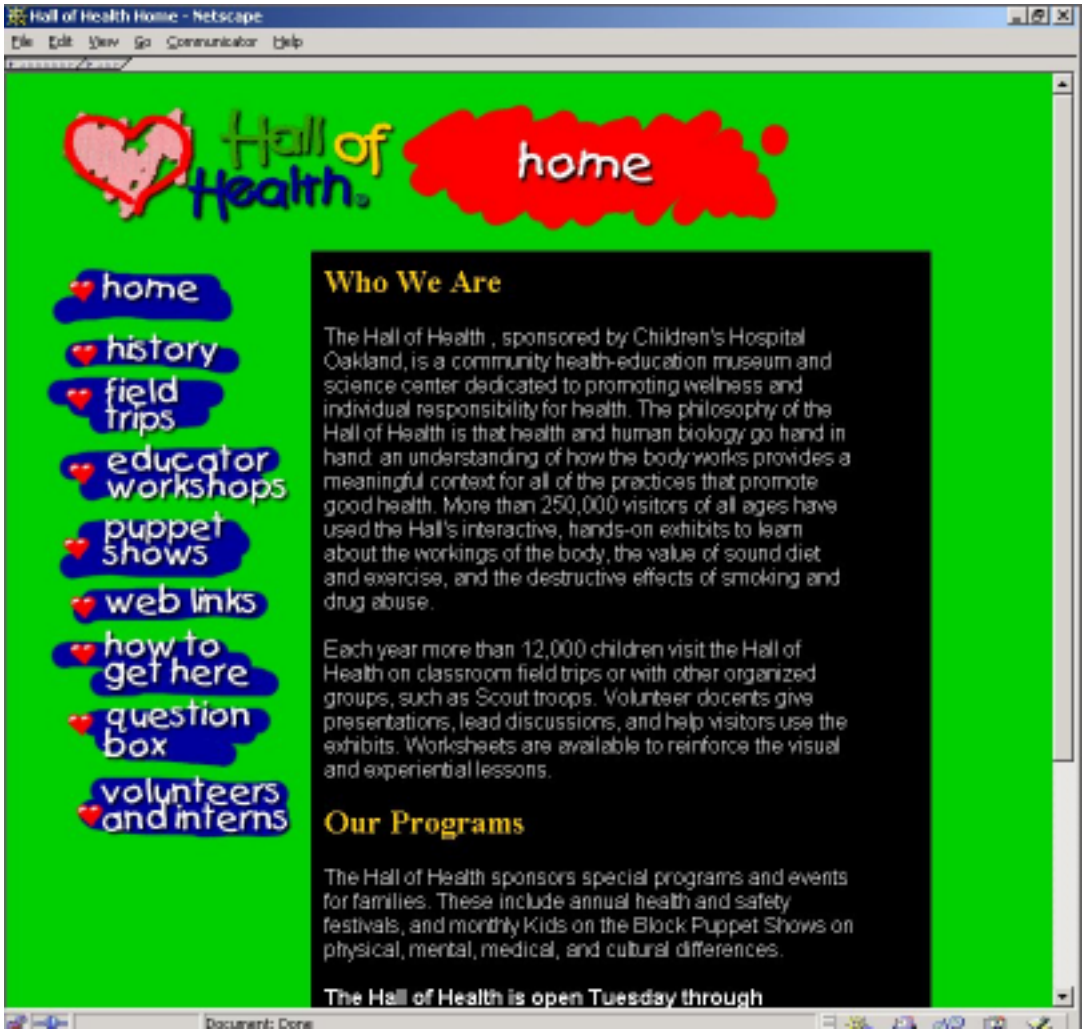

Figure 8.7: Modified home page for the example health education site. Gold headings were added, sans serif fonts were used for body text, colored and italicized body text was removed, and a fixed page width of 640 pixels was used. A footer navigation bar was added to the bottom of the page, an accent color was added to the footer navigation bar, footer elements were reorganized to reduce vertical scrolling, and the font size of footer text was reduced; none of these changes are visible in the screen shot. See Figure 8.9 for the footer navigation bar.

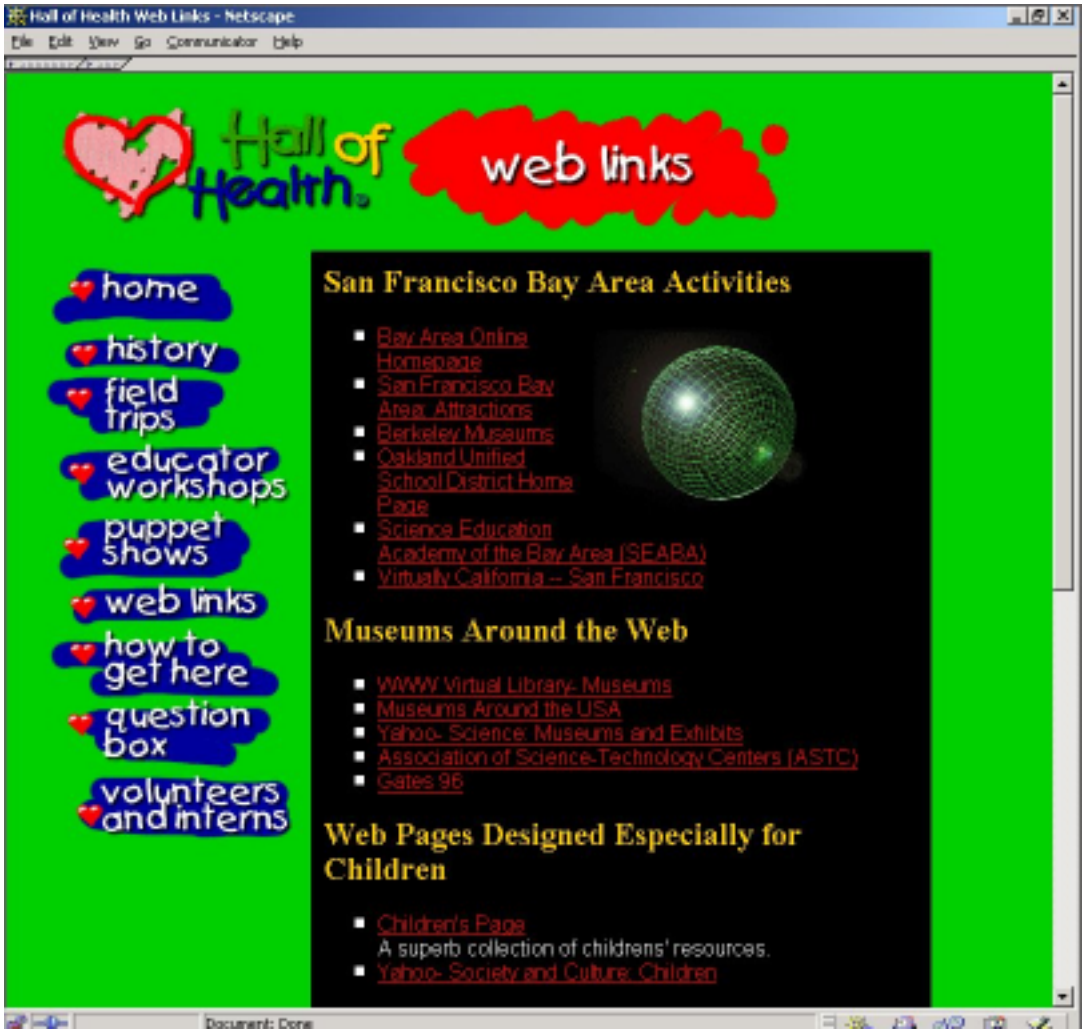

Figure 8.8: Revised link page for the example health education site. Gold headings were added, sans serif fonts were used for body text, colored and italicized body text was removed, the sizes of images were reduced, and a fixed page width of 640 pixels was used. A footer navigation bar was added to the bottom of the page, an accent color was added to the footer navigation bar, footer elements were reorganized to reduce vertical scrolling, and the font size of footer text was reduced; none of these changes are visible in the screen shot. See Figure 8.9 for the footer navigation bar.

Figure 8.9: Revised content page for the example health education site. Gold headings were added, sans serif fonts were used for body text, colored and italicized body text was removed, the sizes of images were reduced, and a fixed page width of 640 pixels was used. A footer navigation bar was added to the bottom of the page, an accent color was added to the footer navigation bar, footer elements were reorganized to reduce vertical scrolling, and the font size of footer text was reduced; not all of these changes are visible in the screen shot.

if ((Minimum Font Size is missing OR (Minimum Font Size $\leq$ 9.5)) AND (Graphic Ad Count is missing OR (Graphic Ad Count $\leq$ 2.5)) AND (Exclaimed Body Word Count is missing OR (Exclaimed Body Word Count $\leq$ 11.5)) AND (Minimum Graphic Height is missing OR (Minimum Graphic Height $\leq$ 38.5)) AND (Vertical Scrolls is missing OR (Vertical Scrolls $\leq$ 3.5)) AND (Bad Panel Color Combinations is missing OR (Bad Panel Color Combinations $\leq$ 2.5)) AND (Object Count is missing OR (Object Count $\leq$ 4.5)) AND (Good Meta Tag Word Count is missing OR (Good Meta Tag Word Count $\leq$ 42.5)) AND (Minimum Color Use is missing OR (Minimum Color Use $\leq$ 12.5)) AND (Horizontal Scrolls is missing OR (Horizontal Scrolls $\leq$ 0.5)) AND (Weblint Errors is missing OR (Weblint Errors $\leq$ 54.5)) AND (Colored Body Word Count is missing OR (Colored Body Word Count $>$ 0.5)) AND (Emphasized Body Word Count is missing OR (Emphasized Body Word Count $\leq$ 183)) AND (Bolded Body Word Count is missing OR (Bolded Body Word Count $\leq$ 43.5)) AND (Script Bytes is not missing AND (Script Bytes $\leq$ 173.5)) AND (Text Positioning Count is missing OR (Text Positioning Count $\leq$ 9)) AND (Serif Word Count is missing OR (Serif Word Count $\leq$ 325.5)) AND (Italicized Body Word Count is missing OR (Italicized Body Word Count $\leq$ 1.5)) AND (Graphic Count is not missing AND (Graphic Count $\leq$ 15.5)) AND (Minimum Graphic Width is missing OR (Minimum Graphic Width $\leq$ 97.5)) AND (Bobby Browser Errors is missing OR (Bobby Browser Errors $>$ 6.5)))

Class = Good

This rule classifies a page as a good page because it: uses a smaller font size for some text; has fewer than sixteen images and no graphical ads; uses at least one image with a height smaller than 39 pixels as well as at least one image with a width smaller than 98 pixels; has fewer than 183 total emphasized (i.e., italicized, bolded, colored, etc.) body words, but has fewer than 11.5 exclaimed body words (i.e., body words followed by exclamation points), fewer than 44 bolded body words, fewer than two italicized body words, and at least 1 colored body word; requires fewer than four vertical scrolls and no horizontal scrolls; starts text in nine or fewer vertical positions; uses fewer than 2.5 bad panel color combinations and uses an accent color; uses no scripts, applets, or other objects; uses fewer than 43 good meta tag words and has fewer than 325 words formatted with serif fonts; and has fewer than 55 Weblint errors and more than six Bobby browser errors.

Figure 8.10: Decision tree rule reported for all of the modified example pages. This rule was reported by the overall page quality model.

The site was still classified as a poor site overall, but for a different reason – too much text element variation. The original site had very little variation in text elements (body and display text in particular); adding headings to pages increased the text element variation (75.5%) above the acceptable threshold of 51.8%. Ensuring that all pages contain similar amounts of display text is probably the simplest way to resolve this issue. Some pages, such as the example link page, have long headings, while other pages have relatively short headings. The site was also classified as a poor health site and a good education site, consistent with classifications before the modifications; the same decision tree rules were reported (see Figure 8.4). The median overall page, education page, and health page quality predictions contradicted the site-level models.

8.7 Summary

This chapter demonstrated the ability to apply the profiles of highly-rated interfaces towards assessing and improving new sites. This capability signifies a major first step towards achieving the fundamental goal of this dissertation – enabling everyday users to produce high-quality Web sites. However, much work remains to be done to fully support this goal. In particular, an approach for suggesting interface improvements in an automated manner needs to be developed. Furthermore, an interactive evaluation tool needs to be developed to support iterative design. Future work will focus on expanding the capability demonstrated in this chapter.

The example assessment provided more insight into what the profiles actually represent and the type of design changes informed by them. The assessment suggests that the profiles provide some support for refining an implemented site, mainly improving the amount of text on the page, text formatting, color combinations, font usage, and other page layout considerations. The profiles do not support improving early site designs or the content; future work will focus on these issues.