Scoring Example

WCAG-EM
Spreadsheet with calculations of scores
Old Scoring Example with testing a site with all WCAG 2.0 SC

Introduction

This example takes a person interested in making a conformance claim through the steps to make a conformance claim for a demo site through:

1. Declaring a scope
2. Taking a representative sample
3. Scoring against the new guidelines: Headings, Clear Language, Visual Contrast of Text
4. Total Score
5. Minimums
6. Level

To demonstrate these steps, we have used the W3C WAI Before and After Demo. We have mostly used the Before pages. This is too small for a demonstration of a representative sample, so we created a simplified representative sample of the WAI site.

Declaring a Scope

Organizations who want to make an optional conformance claim can select a logical sub-section of the site, application, or product. The scope has to be a logical sub-section (as is commonly done today) of the material. It cannot be used to make a claim outside the scope.

Since the Before and After Demo newsletter is part of the larger WAI website, which is in turn part of the larger W3C website, we are stating that the scope of this evaluation is the logical subset of w3.org, the Before and After Demo (w3.org/WAI/demos/bad/). Note that scope does not need to be expressed as a URL. Complex single page apps or mobile apps could declare a scope that could not be expressed as a URL.

Taking a Representative Sample of a Website

Because the Before newsletter is only 4 pages, we would have to test every page.

A better example of sampling would be to evaluate the WAI website at https://www.w3.org/WAI/. We would follow the rules of WCAG-EM. These rules are part of a thorough W3C spec with deep details. For the purpose of this example, we are giving a high level sketch of the

WCAG-EM rules and the sample. Caveat: This is not a thorough evaluation of the WAI site and has been grossly simplified to illustrate the process.

Step 1: Define the Evaluation Scope

WAI subsection of the W3C website
https://www.w3.org/WAI/

Step 2: Explore the Target Website

Step 2.a: Identify Common Web Pages of the Website

Home
Accessibility Fundamentals
Planning & Policies
Design & Develop
Test & Evaluate
Teach & Advocate
Standards/Guidelines

Step 2.b: Identify Essential Functionality of the Website

To provide information on a variety of topics

Step 2.c: Identify the Variety of Web Page Types

Landing pages (top navigation & common footer)
Detail pages on a topic (left navigation)
Video pages (captioning and translation options)
And many more, but we want to keep this example simple.

Step 2.d: Identify Web Technologies Relied Upon

HTML
CSS
JavaScript
WAI-ARIA
SMIL
SVG

Step 2.e: Identify Other Relevant Web Pages

Contact WAI
Site Map
News
Archive
Accessibility Statement
Translations
Resources for Roles

Step 3: Select a Representative Sample

Step 3.a: Include a Structured Sample

1. Home
2. Accessibility Fundamentals
3. Planning & Policies
4. Design & Develop
5. Test & Evaluate
6. Teach & Advocate
7. Standards/Guidelines
8. Introduction to Accessibility
9. Video Introduction to Accessibility
10. Contact WAI
11. Site Map
12. News
13. Archive
14. Accessibility Statement
15. Translations
16. Resources for Roles

Note: There would be more samples in a real evaluation, but we want to keep this simple.

Step 3.b: Include a Randomly Selected Sample

We did not count all the pages on the WAI site, but for this example, let us assume that there are 120 pages. Therefore, we would add 12 random pages, as WCAG-EM recommends.
[12 randomly selected pages]

Step 3.c: Include Complete Processes

[login, checkout, create new account]

Total Sample selected would be the:

16 structured pages
12 random pages
0 complete processes (not applicable)
28 pages in total

Taking a Sample of a Non-Website

Paraphrased from WCAG-EM as if WCAG-EM has been rewritten for non-websites. Note: These changes have not been discussed or reviewed by WCAG-EM editors to date, but will be coordinated with them before finalizing.

Step 1: Define the Evaluation Scope

New York Times App for iOS, the For You section.

Step 2: Explore the Target

Step 2.a: Identify common repeating templates or components
- Your daily bulletin
- Your columns and series
- Improve recommendations
- From your interests
- We think you might like
- Saved for later

Step 2.b: Identify Essential Functionality
- Provide customized information
- Modify selections

Step 2.c: Identify the Variety of Screen Types
- For you
- The Morning/Evening Briefing
- The Daily Podcast (audio)
- The Mini Crossword
- Customize Bulletin
- Article
- Multimedia
- Virtual Reality

Step 2.d: Identify Web Technologies Relied Upon
(Note: We should discuss how this would be adapted. For this example, we will pick standard screens, audio, and video.)

Step 2.e: Identify Other Relevant Screens
- Comments
- Bookmark
- Share
- Footer

Step 3: Select a Representative Sample

Step 3.a: Include a Structured Sample
1. For you
2. The Morning/Evening Briefing
3. The Daily Podcast (audio)
4. The Mini Crossword
5. Customize Bulletin
6. Article template
7. Multimedia template
8. Virtual Reality template

9. Comments
10. Bookmark
11. Share
12. Footer

Step 3.b: Include a Randomly Selected Sample

If we pretend that there are 15 different article templates, we would add 2 random samples

Step 3.c: Include Complete Processes

Create an account (3 screens)

Login (2 screens)

Total representative sample: 19 screens (12 structured, 2 random, 5 complete processes)

Scoring the New Guidelines

Headings

Before Demo (4 pages)
B-Home: 7 headings, no semantics, visual hierarchy, succinct, nested
B-News: 8 headings, no semantics, visual hierarchy, succinct, nested
B-Tickets: 3 headings: 2 with no semantics, 1 with table heading. Visual hierarchy, succinct, nested
B-Survey: 3 headings, no semantics, visual hierarchy, succinct, nested
21 heading total

No semantics	1 pass of 21 headings	=1/21 (total # of headings) = 4%
Visual hierarchy	21 pass of 21 headings	=21/21 100%
Succinct	21 pass of 21 headings	= 21/21 100%
Nested	20 pass of 21 headings	=20/21 95%
		= (sum)/4 75%

Total Score: 75%

Clear Language

This is an example of scoring the Clear Language Rubric from the W3C WAI Before and After Demo (BAD) This is an example of how that site could have been scored.

I took the rubric (First column from the Editor's draft of Clear Language. I added a column for the test results and the score. At the end, the scores are averaged (normalized).

Highlighted text following needs someone to look at the 4 Before pages and score them. This is a linked table. Make changes to the underlying spreadsheet.

Action	**Test**	**Score**
Use proper grammar for the language, as defined by a style guide or by standard language rules	Ran Hemingway grammar checker and made changes where they made sense (human override of automated test).	100%
Use correct spelling	Ran the browser spell checker and made corrections	100%
Use active voice.	Counted the use of passive voice (after editing) and saw it was used in about 30% of sentences. 70% of the sentences use active voice.	70%
Use simple verb tense.	Counted the use of simple verb tense (after editing) and saw it was used in 65% of sentences.	65%
Use literal and concrete language.	In 4 pages of text, there were 5 examples of figurative language. Estimating 75 paragraphs in 30-page web app, 5 of 75 used figurative language.	93%
Limit use of metaphor, similes, sarcasm, or irony.	All metaphors, sarcasm and irony removed. 15 similes remain to help clarity.	80%
Use common words or terms to represent concepts where it makes sense. When uncommon words are necessary, define them in context.	All uncommon words are included in a glossary that displays the definition in-line	100%
Remove unnecessary words	Reviewed by an editor looking for unnecessary words	100%

Avoid the following when possible. Define or spell out any you use in context: Technical terms Jargon Idioms Slang Acronyms Abbreviations	All technical words and acronyms are including in glossary. There are no jargon or slang phrases.	100%
Break text into sections and provide headings. See the guideline on headings and the guideline on white space (to be written).	Headings used that describe content. Evaluated by an editor.	100%
Keep sentences and paragraphs short.	Ran grammar checker. Some sentences are still long. 73% of sentences are short	73%
Number sequential steps.	N/A	
Provide brief summaries at the top of long documents.	Blog articles have summaries.	100%
Provide words and numeric symbols for numbers. Explain numeric concepts in simple words.	Inconsistent. About 30% of numbers have symbols and words descriptions.	30%
Provide alternatives or explanations for symbols and icons.	N/A	
	Total	**85%**

The total score of the Clear Language Rubric becomes one of the rows (the last row) of the Overall Guideline Scoring Example.

Visual Contrast

https://www.w3.org/WAI/demos/bad/before/home.html (4 pages)

Total Score

[Add the 3 scores, divide by 3 guidelines for a normalized score]

Minimums

WCAG 2.0 Level A	Silver Score	Usage without vision	Usage with limited vision	Usage without perception of colour	Usage with limited hearing	Usage without vocal capability	Usage with limited manipulation or strength	Usage with limited reach	Minimize photosensitive seizures	Usage with limited cognition
Headings										
Clear Language										
Visual Contrast of Text										

Level

OLD - To Be Deleted.

Keep in mind, that the Clear Language rubric:

- does not require 100% conformance to pass. Working with human languages is complex and needs to be flexible.
- is a draft designed to illustrate a concept
- is designed to be tested by an editor, not a quality assurance tester

The Clear language results are normalized for that guideline and the total score for the rubric is included as the last row in the table of the Overall Guideline Scoring Example

Where did this data come from?

Units of Scoring

- The unit of scoring varies by guideline, for whatever is appropriate. :
    - page or screen-based
    - site or project-based
    - instance-based (the number of instances of the element involved)
- The notes column was a quick evaluation of the unit that would be used to measure the guideline. It is certainly a point of greater discussion as we develop that content.
- Most scores are page-oriented, because that was the data Jeanne had. It would work the same way for instance scoring. 1.1.1 data was changed to be scored by the number of images, so you can see how scoring by instance works.

Overall Guideline Scoring Example

NOTE: Clear Language is the last row in the table before the totals. It is based on the rubric scoring which is in the Scoring a Rubric section.

WCAG 2.0 Level A	**Success Criteria Totals**	Total instances	Passes	Silver Score	Unit of Scoring
1.1.1 Non-Text Content (Level A)					
All non-text content must have a text alternative that serves the same purpose. For images, ensure the alt attribute is an equivalent if the image is not presented.	**13 Fails, 63 Passes, 0 N/A's**	76	63	83%	instance-based
1.2.1 For prerecorded audio-only and prerecorded video-only media, the following are true, except when the audio or video is a media alternative for text and is clearly labeled as such: (Level A)					
* Prerecorded Audio-only: An alternative for time-based media is provided that presents equivalent information for prerecorded audio-only content.					
	0 Fails, 0 Passes, 32 N/A's			N/A	instance-based

*Prerecorded Video-only: Either an alternative for time-based media or an audio track is provided that presents equivalent information for prerecorded video-only content.					
1.2.2 Captions are provided for all prerecorded audio content in synchronized media, except when the media is a media alternative for text and is clearly labeled as such. (Level A)	**0 Fails, 0 Passes, 32 N/A's**			N/A	instance-based
1.2.3 An alternative for time-based media or audio description of the prerecorded video content is provided for synchronized media, except when the media is a media alternative for text and is clearly labeled as such. (Level A)	**0 Fails, 0 Passes, 32 N/A's**			N/A	instance-based
1.3.1 Information and Relationships (Level A)					
Ensure that information, structure, and relationships are programmatically determined. This includes using the most semantically appropriate elements, such as headings, paragraphs, lists, and so on, to provide structure.	**30 Fails, 2 Passes, 0 N/A's**	32	2	6%	Needs restructuring
1.3.2 Meaningful Sequence (Level A)					
The correct reading sequence must be programmatically determined.	**16 Fails, 15 Passes, 0 N/A's**	31	15	48%	Needs more discussion
1.3.3 Sensory Characteristics: Instructions provided for understanding and operating content do not rely solely on sensory characteristics of components such as shape, size, visual location, orientation, or sound (Level A)	**0 Fails, 15 Passes, 17 N/A's**	15	15	100%	instance-based
1.4.1 Use of Color (Level A)					
Color is not used as the only visual means of conveying information, indicating an action, prompting a response, or distinguishing a visual element.	**7 Fails, 25 Passes, 0 N/A's**	32	25	78%	instance-based

1.4.2 Audio Control:	**0 Fails, 1 Pass, 31 N/A's**	1	1	100%	instance-based
2.1.1 Keyboard (Level A)					
All functionality of the content must be operable using the keyboard-alone.	**30 Fails, 1 Pass, 1 N/A**	31	1	3%	site-based
2.1.2 No Keyboard Trap (Level A)					
If focus can be moved to a component on the page, focus can be moved away from that component.	**0 Fails, 29 Passes, 3 N/A's**	29	29	100%	Needs more discussion - maybe a percentage of focusable elements not impacted by the keyboard trap.
2.2.1 Timing Adjustable (Level A)					
For content with a timeout, the user must be able to either turn off the timeout or extend the timeout.	**15 Fails, 17 Passes, 0 N/A's**	32	17	53%	page-based
2.2.2 Pause, Stop, Hide (Level A)					
Content that blinks, scrolls, or updates very regularly should be able to be stopped so as to avoid causing distraction.	**0 Fails, 0 Passes, 32 N/A's**				
2.3.1 Three flashes or below (Level A)					
Pages must not contain content that flashes more than three times in any one second period, or the flash is below the general red flash threshold, as this can cause light-sensitive seizures.	**0 Fails, 17 Passes, 15 N/A's**	17	17	100%	
2.4.1 Bypass Blocks (Level A)					
A mechanism must be available for users to bypass blocks of content.	**17 Fails, 13 Passes, 1 N/A**	30	13	43%	
2.4.2 Page Titled (Level A)					
Web pages must have titles that describe their topic or purpose.	**14 Fails, 14 Passes, 3 N/A's**	28	14	50%	
2.4.3 Focus Order (Level A)					
	29 Fails, 2 Passes, 1				

The order focusable components receive focus must be in an order that preserves meaning or operability.	**N/A**	31	2	6%	page-based
2.4.4 Link Purpose in context (Level A)					
The purpose of each link must be determinable from the link text alone, or from the link text along with the programmatically determinable context, such as a heading.	**8 Fails, 20 Passes, 1 N/A**	28	20	71%	
3.1.1 Language of the Page (Level A)					
The default human language of the page must be programmatically determinable.	**1 Fail, 27 Passes, 3 N/A's**	28	27	96%	page-based
3.2.1 On Focus (Level A)					
When a component receives focus, it should not initiate a change of context, such as opening a new page when an element receives focus.	**6 Fails, 22 Passes, 3 N/A's**	28	22	79%	
3.2.2 On Input (Level A)					
Changing the settings of a control should not cause a change of context. For example, it should be possible to scroll through a dropdown list without an action being invoked that causes the user to open another page.	**0 Fails, 29 Passes, 2 N/A's**	29	29	100%	
3.3.1 Error Identification (Level A)					
If an input error is automatically detected, the item that is in error is identified and the error is described to the user in text.	**3 Fails, 17 Passes, 11 N/A's**	20	17	85%	
3.3.2 Labels or Instructions (Level A)					
Labels or instructions are provided when content requires user input.	**28 Fails, 0 Passes, 3 N/A's**	28	0	0%	page-based
4.1.1 Parsing (Level A)					
Elements have complete start and end tags (except where it's allowable in HTML), elements are nested according to their specifications, elements do not contain duplicate attributes, and any IDs are unique.	**9 Fails, 7 Passes, 1 N/A**	16	7	44%	

4.1.2 Name, Role, Value (Level A)					
For all user interface components (including but not limited to: form elements, links and components generated by scripts), the name and role can be programmatically determined; states, properties, and values that can be set by the user can be programmatically set; and notification of changes to these items is available to user agents, including assistive technologies.	**31 Fails, 0 Passes, 1 N/A**	31	0	0%	
Clear Language Rubric	See Rubric Example			85%	
Total Score is the score per guideline divided by the number of applicable guidelines		**22**		**1330 %**	**60%**

Minimums for Category

Notes: We took the Silver score from the above table and then assigned it to the appropriate category. These may not be correct, and certainly can be debated. The point is to show how it could work—not whether the assigned categories were correct.

The required minimum score in the categories is still to be determined and will require more research.

WCAG 2.0 Level A	Silver Score	Usage without vision	Usage with limited vision	Usage without perception of colour	Usage with limited hearing	Usage without vocal capability	Usage with limited manipulation or strength	Usage with limited reach	Minimize photosensitive seizures	Usage with limited cognition
1.1.1 Non-Text Content (Level A)	83%	83%	83%							83%
1.2.1 For prerecorded audio-only and prerecorded video-only media, (Level A)										
1.2.2 Captions (Level A)										

1.2.3 An alternative for time-based media or audio description (Level A)										
1.3.1 Information and Relationships (Level A)	6%									
1.3.2 Meaningful Sequence (Level A)	48%	48%								
1.3.3 Sensory Characteristics:	100%	100%	100%	100%	100%					
1.4.1 Use of Color (Level A)	78%	78%		78%						
1.4.2 Audio Control:	100%	100%								
2.1.1 Keyboard (Level A)	3%	3%					3%	3%		
2.1.2 No Keyboard Trap (Level A)	100%	100%					100%	100%		
2.2.1 Timing Adjustable (Level A)	53%	53%	53%							53%
2.2.2 Pause, Stop, Hide (Level A)										
2.3.1 Three flashes or below (Level A)	100%								100%	
2.4.1 Bypass Blocks (Level A)	43%	43%					43%	43%		43%
2.4.2 Page Titled (Level A)	50%	50%	50%							50%
2.4.3 Focus Order (Level A)	6%	6%					6%	6%		
2.4.4 Link Purpose in context (Level A)	71%	71%	71%							71%
3.1.1 Language of the Page (Level A)	96%	96%								
3.2.1 On Focus (Level A)	79%	79%					79%	79%		79%
3.2.2 On Input (Level A)	100%	100%					100%	100%		100%
3.3.1 Error Identification (Level A)	85%	85%	85%							85%
3.3.2 Labels or Instructions (Level A)	0%	0%	0%							0%
4.1.1 Parsing (Level A)	44%									

4.1.2 Name, Role, Value (Level A)	0%									
Total of columns		1095 %	442%	178%	100%		331%	331%	100%	564%
Number of rows		17	7	2	1		6	6	1	9
Score		**64%**	**63%**	**89%**	**100%**		**55%**	**55%**	**100%**	**63%**

We have not determined what a passing score is for each category. We need more research.

Assigning a level

These numbers are a *placeholder* pending research testing existing web sites and products.

Bronze	85% - 89.9%
Silver	90%-96.9%
Gold	97% and up

Scoring a Rubric

This is an example of scoring the Clear Language Rubric from the web site analized. The authors of the site had put a lot of editorial work into the site and the language had been tested with users (although not with users with disabilities.) This is an example of how that site could have been scored.

I took the rubric (First 4 columns from the Editor's draft of Clear Languge. I added a column for the test results and the score. At the end, the scores are averaged (normalized).

Action	Substantially = 1	Partially = .5	Limited = 0	Test	Score
Use proper grammar for the language, as defined by a style guide or by standard language rules	Passes a grammar checker. Errors improve clarity.	Some errors	Little to no effort has been made.	Ran Hemingway grammar checker and made changes where they made sense	100%

Use correct spelling	Passes a spellchecker. Context errors (false positives and negatives) are corrected	Passes a spellchecker, but there are context errors	Uncorrected spellchecker errors	Ran the browser spellchecker and made corrections	100%
Use active voice.	No passive voice, or passive voice only where it improves clarity	Occasional use of passive voice.	Most sentences have passive voice.	Counted the use of passive voice (after editing) and saw it was used in about 30% of sentences. 70% of the sentences use active voice.	70%
Use simple verb tense.	Verb tense is consistent and appropriate for the content.	Verb tense has some inconsistencies or incorrect tense, but content is still understandable	Content is difficult to understand because mixed verb tenses or perfect or continuous tenses are used.	Counted the use of simple verb tense (after editing) and saw it was used in 65% of sentences.	65%
Use literal and concrete language.	No figurative or abstract language is used.	Uses some figurative or abstract language is used, but content is still understandable	Content is difficult to understand because figurative or abstract language is used.	In 30 pages of text, there were 5 examples of figurative language. Estimating 75 paragraphs in 30-page web app, 5 of 75 used figurative language.	93%
Limit use of metaphor, similes, sarcasm, or irony.	No metaphor or sarcasm used.	Uses some metaphor or sarcasm,	Uses abstract metaphor, like “getting	All metaphors, sarcasm and irony removed. 15 similes remain to help clarity.	80%

		but includes context to aid understanding	cold feet" without context to aid understanding		
Use common words or terms to represent concepts where it makes sense. When uncommon words are necessary, define them in context.	A common words list has been used where it makes sense.	Some uncommon words have been used but the content is still understandable.	Unnecessary uncommon words have been used or necessary uncommon words are not defined.	All uncommon words are included in a glossary that displays the definition in-line	100%
Remove unnecessary words	Text is concise and understandable. Detailed information has a summary or a heading structure to make it easy to skim.	Text has some extra words, but is still understandable.	Text contains extra words that make it difficult to understand the content.	Reviewed by an editor looking for unnecessary words	100%
Avoid the following when possible. Define or spell out any you use in context: Technical terms Jargon Idioms Slang Acronyms Abbreviations	One of the following is true: No technical terms, jargon, idioms, etc. are used. Definitions are provided in context or are one	One of the following is true: Uses some technical terms, jargon, idioms, etc. but don't always provide definitions	Uses technical terms, jargon, idioms, etc terms without definitions.	All technical words and acronyms are including in glossary. There are no jargon or slang phrases.	100%

	click away and one click to return.	and context to aid understanding. Definitions are provided but the user must search for them or locate them in a glossary.			
Break text into sections and provide headings. See the guideline on headings and the guideline on white space (to be written).	Text is consistently broken into sections with headings.	Some of the text is broken into sections with headings, but not consistently.	Text is not broken into sections or headings are not used.	Headings used that describe content. Evaluated by an editor.	100%
Keep sentences and paragraphs short.	Sentences and paragraphs are short.	Most sentences and paragraphs are short.	Content is difficult to understand because sentences and paragraphs are too long.	Ran grammar checker. Some sentences are still long. 73% of sentences are short	73%
Number sequential steps.	Sequential steps are organized logically and numbered appropriately	Sequential steps are organized logically but not numbered	Sequential steps are not organized logically nor numbered	N/A	

Provide brief summaries at the top of long documents.	A brief summary that describes the document content is provided.	A summary is not provided but another means of aiding understanding is available.	No summary is provided or alternative is available.	Blog articles have summaries.	100%
Provide words and numeric symbols for numbers. Explain numeric concepts in simple words.	Numbers and numeric concepts have text and symbol alternatives	Some numbers and numeric concepts do not have symbol alternatives, but content is understandable	No text or symbol alternatives are provided for numeric concepts	Inconsistent. About 30% of numbers have symbols and words descriptions.	30%
Provide alternatives or explanations for symbols and icons.	Text alternatives or explanations are available for symbols and icons	Some icons or symbols do not have text alternatives but content is understandable	Symbols and icons do not have text alternatives or explanations	N/A	
				Total	**85%**

The total score of the Clear Language Rubric becomes one of the rows (the last row) of the Overall Guideline Scoring Example.

Still to Do:

We need to get the critical path of the organization's scope. It doesn't impact the score, but it gives an organization the pass/fails where it counts. It isn't tolerable to have a lower score on a critical path. For example, on the critical path, you would need 90% to pass and 85% on the rest of the site.