PRODUCING INFORMATIVE TEXT ALTERNATIVES FOR IMAGES

A Thesis Submitted to the College of
Graduate Studies and Research
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in the Department of Computer Science
University of Saskatchewan
Saskatoon

By
Lisa Tang

PERMISSION TO USE

In presenting this thesis/dissertation in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis/dissertation in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis/dissertation work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis/dissertation or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis/dissertation.

ABSTRACT

A picture may be worth a thousand words but what might those words be? How do we go about finding those words? Images are often used to convey information, supplement textual content, and/or add visual appeal to documents. Unless the user can see the image and properly interpret it, the user may not receive the same information. While containers exist for providing text alternatives in various types of electronic documents (including Web pages), they are rarely used. When they are used, the text alternatives are not informative. While guidance currently exists regarding which containers to use in order to provide text alternatives, there is little guidance available regarding what information to include in these containers and how to compose text alternatives. The purpose of this work is to establish a procedure for identifying the information being communicated within an image and provide guidance on how to produce informative text alternatives.

Based on related information in the areas of Web accessibility, library cataloguing, captioning and audio description, image retrieval and indexing, art description, and tactile representation, important information communicated by an image were identified and a procedure for producing informative text alternatives using that information was developed. Studies were conducted to determine the effectiveness of the procedure to identify important information about an image.

Study 1 determined the information identified about an image when the procedure was not available. It also suggested reasons why people would opt to not provide a text alternative for an image. Study 2 determined the information that people would identify when they were given the procedure and a set of questions to help identify information about an image. Study 3 determined the information people identified when they were required to consider all of the different types of information that may be important in an image. The results from these three studies were compared to determine the effectiveness of the procedure to identify important information about the image. Study 4 presented the information identified in the previous three studies to sighted and visually impaired users to evaluate the importance of such information. This study determined the quality of the information identified in the first three studies and the ability of the procedure to identify important information for a wide set of images.

The results of these studies showed that the procedure was effective in identifying a greater amount of important information than without the procedure. Additional guidance was also identified to further help people create informative and useful text alternatives. The studies also showed that the procedure could be applied by different user groups to a wide range of images. The procedure was submitted to the International Standards Organization to become a technical specification, which will be available to people around the world.

ACKNOWLEDGEMENTS

The main reason I chose to study Usability Engineering, Human Computer Interaction, and accessibility is because I wanted to make software that will be useful to people and can help people. I believe this research and artefacts resulting from this research will be able to do that in a much grander scale than I ever imagined possible. I would like to thank the many people who helped me in making it happen.

First and foremost, I would like to thank my supervisor, Dr. Jim Carter. Your guidance, encouragement, and vision throughout this entire process are what brought me here today. Without your support, I know that I would not have accomplished as much as I have.

I would also like to thank Microsoft Canada for sponsoring this work. Thank you, Jim MacFie and John Weigelt, for your guidance and suggestions that started this research project.

I would like to thank the members of my advisory committee: Dr. Regan Mandryk, Dr. Chanchal Roy, Dr. Kevin A. Schneider, Dr. Allison Muri, and my external examiner: Dr. Gregg Vanderheiden for your suggestions and recommendations that helped improve my work.

Many thanks to all the members of the ISO/IEC JTC 1/SC 35 WG6, who made it possible for this research to become an international technical specification.

Thanks to all those who volunteered their time to participate in my research studies. I know it took a lot of time and I truly appreciate it. Your effort has not been wasted.

Thanks to all my friends who listened to me complain, who said encouraging words, who made me relax, who helped me whenever I asked.

Finally, I would like to especially thank my parents and my brother for their support and encouragement to pursue this venture. Thank you for giving me the freedom and trust to do what I needed to achieve this goal, no matter how long it took.

DEDICATION

This thesis is dedicated to

My parents (Johnny and Sanna) and my brother (Timmy)

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1 Introduction

Images are often inserted into documents to add visual appeal or supplement and/or complement the document content. [45, 51, 64] The main document content often makes references to those images. Sometimes, the main document content explains exactly what the images are communicating. More often than not, the main document content will comment on the images without explicitly explaining what the images are communicating. The main document content may refer to a graph, for example, that indicates the revenue for the past year but it might not state the actual revenue values, which may be important to the reader. A person who can see the graph may be able to interpret the values and other information not provided by the main document content, while those who cannot see the graph would not receive such information.

As more and more content and information is being presented on the Internet (in the form of Web pages), the issue of images being accessible becomes an even bigger problem. With Web pages, the authors of Web page content make fewer references to the images (and other graphical content), such as background images, navigation images, and animations. The images may simply exist without explanation, especially in the case of advertisements. People who use screen reader technology to read content aloud or people who have images turned off may be aware that an image exists but might not know what the image is of. Again, they may be missing important information.

Text alternatives (also referred to as “alternative text”) are a literary composition that represents or describes an image (and other graphical content). [67] Text alternatives allow everyone to understand the information being communicated by the image. [48] Unfortunately, text alternatives are rarely provided or done well.

1.1 Lack of Text Alternatives and Guidance for Text Alternatives

McEwan and Weert reported that “non-use, and incorrect use, of ALT (alternative) text emerges as the most frequent, basic error", with regards to Web pages. [38] Web pages and the

Internet are increasingly populated with graphical content, such as images, applets, videos, and animations. As such, there is an increasing amount of information that may be inaccessible to users.

Technology already exists to allow those using assistive technologies (such as screen readers) to "see" graphical content, mainly through the use of text alternatives. Microsoft Office documents, Flash content, and HTML all have tags or containers available for people to provide an alternate textual representation. However, either people are unaware of such containers or choose not to use them. [48] In 2006, it was reported that 40% of images on Web pages do not have text alternatives. [8] When people do make use of the containers, the provided text alternatives are often uninformative. Antonacopoulos et al. reported that 56% of the time, the text alternatives were false, incomplete, or non-existent. [2] According to screen reader users, it would be better to not have text alternatives than to have poor text alternatives. [7, 38] This does not mean that people should omit text alternatives.

When text alternatives are not properly provided, users may be aware that an image exists but do not know what the image is of. Screen readers generally report the file name of the image if no text alternatives are provided. [38, 60] Some screen reader users create custom scripts such that the screen reader avoids reading poorly written text alternatives. [7] However, not every screen reader user knows how to do this and the user remains frustrated with poorly written text alternatives.

Existing accessibility guidelines (such as the Web Content Accessibility Guidelines (WCAG) [66, 68]) state that it is necessary to provide text alternatives for images (and other graphical content) on a Web page. However, they also suggest situations when text alternatives are not necessary and provide a method for instructing screen readers to ignore the image. Although it is true that some images may not be needed in order for the reader to understand the main document content, "this decision must be left to the reader, although providing a means of assessing the diagram could save them significant effort." [9]

While existing accessibility guidelines suggest when text alternatives should and should not be provided, the guidelines do not indicate what information should be included in text alternatives when they are to be provided. If people do not know what to write for text alternatives, then the text alternatives may be badly written and the users are "often left completely unaware that there is additional visual information on the page." [62]

1.2 Why Use Text

Visual content could be translated in various ways into the auditory and tactile or haptics modalities. [5] Why then should the focus be on text and not other formats? As WCAG states, "Text alternatives are a primary way of making visual information accessible, because they can be rendered through any sensory modality (for example, visual, auditory or tactile) to match the needs of the user." [68] Text is the easiest method of recording the information that needs to be communicated. Once specified, it can be transformed into the modality and form that the user needs.

Text alternatives could be produced for all non-textual content, both audio and visual, and translated into other modalities. In this thesis, the focus will be on extracting information from static visual content, such as images, and presenting the information as textual content.

Text alternatives help people who cannot see images understand the content of an image by providing the same information in a textual form. [48] Text alternatives can be beneficial in the following scenarios: [6, 8, 68]

- Users disable images on their Web browser due to a slow Internet connection
- Users have a vision impairment
- Users use text-to-speech software
- Users use a text browser
- The image file no longer exists but the image placeholder still exists
- Users are multitasking and cannot look at the screen

Text alternatives benefit not only those who cannot see the images, but also those who can see but do not have the knowledge or ability to interpret the image. For example, Figure 1.1 below is an optical illusion. Some people see an old woman, some people see a young woman, and some people may not see anything at all. An explanation of the illusion would help those who are unable to see it to understand the significance of the image. Users with cognitive disabilities could also make use of text alternatives to comprehend and interpret an image.

Search engines can also make use of text alternatives in searching for information or for images. [42] Having text alternatives for images (and other graphical content) can increase the accessibility and availability of content to everyone.

Figure 1.1 Optical illusion of either an old or young woman [30]

1.3 Problems with Text Alternatives

Text alternatives are beneficial but there are some problems when trying to write them. The fundamental problem being that developers do not know what information to include in text alternatives. In the research study described in Chapter 4, the results indicate that developers would provide text alternatives if and when they knew what to write, but would opt to use the empty text (i.e., “”) if and when they did not. A starting point for writing text alternatives would be the reason or purpose that an image serves; however, it is not always known why an image was included or selected. [19]

Also, an image consists of and communicates a vast amount of information. [30] While it is desirable to keep as much information about the image as possible, there may be too much information for a person to process. [51] Some of the information may also be irrelevant given the context or purpose of the image.

The information (once identified) needs to be filtered to be more manageable and to include only the relevant information. The process of identifying and filtering information could take a lot of time. How does one determine which pieces of information to filter out? People need guidance on identifying and focusing on the important information.

Once the information has been filtered, it then needs to be organized and presented such that it is concise without missing important information. [53] Writing text alternatives can be very difficult and can be considered an art. [8, 53, 63] It is understandable that people avoid writing text alternatives or end up writing poor text alternatives. Therefore, people need guidance on writing informative text alternatives that will be useful to users.

While the term “alt-text” is the abbreviation for “alternative text”, it is often associated specifically with the text provided through the `alt` attribute of the HTML `img` tag. “Alt-text” has become restricted to HTML and is thought to be a short description due to the restrictions placed by Web browsers. However, text alternatives can be more than a short description and can exist in platforms and technologies other than HTML. Therefore, for the purpose of this thesis, the term “text alternatives” will be used to refer to any length of textual description or representation of an image on any platform or document type.

1.4 Research Questions and Thesis Structure

This thesis attempts to answer the following questions:

- What are the steps to produce informative text alternatives?
- What information are images communicating?
- What information is important to include in text alternatives?
- How can the procedure for producing text alternatives be made faster, simpler, and easier for people to use?

This thesis consists of:

- Chapter 2 describes and analyzes the existing work and guidance regarding text alternatives and the work in other research areas regarding providing visual information in other mediums. It also summarizes the requirements for producing informative text alternatives.
- Chapter 3 describes a procedure to produce text alternatives and identifies types of information that may be communicated by an image.
- Chapter 4 provides the details of a research study to determine how developers currently produce text alternatives, without the help of a procedure or guidance.
- Chapter 5 provides the details of a research study where a group of participants identified information about images and produced text alternatives with the help of a procedure presented in a document format.
- Chapter 6 provides the details of a research study where a group of participants identified information about images with the help of a procedure presented as a prototype tool.

- Chapter 7 provides the details of a research study where a group of participants evaluated the importance of information about images, which was identified in the previous research studies.
- Chapter 8 summarizes the results of the research studies, the contributions of this thesis, and the future work of this project.

Chapter 2 Background and Requirements

This chapter analyzes the existing literature available regarding text alternatives. It also looks into the areas of library cataloguing, captioning and audio description, image retrieval and indexing, art description, and tactile representation to identify information that may be important to describe and to suggest a procedure for producing informative descriptions using that information. Finally, a set of requirements for creating informative text alternatives is provided.

2.1 Existing Text Alternatives Guidelines and Research

World Wide Web Consortium's (W3C) WCAG 2.0 states "Provide text alternatives for any non-text content so that it can be changed into other forms people need, such as large print, Braille, speech, symbols or simpler language." [66] The text alternatives serve an equivalent purpose as the non-text content, that is, to provide the equivalent information in a textual format. Many Web accessibility organizations (such as XS4ALL [48] and Web Accessibility In Mind (WebAIM) [63]) supported the guidance of WCAG 2.0 and provided additional examples of text alternatives for sample images.

Much of the existing guidance focuses on the containers available for providing text alternatives on the Web, specifically within HTML. Each container serves a different purpose and some have specific length restrictions. For example, the main containers for text alternatives in HTML are the `alt` attribute in various HTML tags (meant for a short description of non-textual content) and the `longdesc` attribute of the `img` tag (meant for longer descriptions of the image). [37, 52, 60, 66] HTML5 introduces additional containers (such as `caption`) for providing text alternatives to figures. [68]

The containers for text alternatives are useful only if they are used by both developers and users. If the text alternatives are poor, users will ignore the text alternatives. [7] While it is important to have containers to support text alternatives and to have guidance that state the containers be used, it is equally (or more) important to provide guidance on what information to include in text alternatives. This aspect is lacking in the current guidance.

Rather than guidance regarding the information to include, most of the existing guidance focuses on when text alternatives are needed and when they are not. Some of the situations when text alternatives are not recommended include: [48, 63, 68]

- The image is purely decorative.
- The image is a spacer or for formatting.
- The text alternative repeats other Web page content.
- The image is a bullet in a list.
- The image is invisible to the sighted user.

In such situations, the guidance recommends an empty or null `alt` attribute in the HTML `img` tag (i.e., “”, henceforth called the empty text) so that screen reader technologies ignore the image. [48] However, for people who could see the image placeholder on a Web page, they may be aware of the image’s existence but not its contents. They might not know that the reason there is no text alternative is due to the reasons listed above.

Automatic evaluation tools exist to ensure that the `alt` attribute in an HTML `img` tag is set for all images on a Web page. Developers could (and do) provide the empty text instead of a text alternative. While WCAG 2.0 recommends that the empty text be used in some situations, it is often abused since it sufficiently satisfies automatic evaluation tools. Needless to say, an image with the empty text does not necessarily improve accessibility.

Many of the images deemed as decorative were chosen for a particular reason. For example, it may have been chosen to create a certain atmosphere or mood. The image might not provide content directly related to the document but it affects the users’ impression of the document and its content. Therefore, it may be important to describe decorative images so the people who could not see the decoration would receive the same impressions.

When writing text alternatives, the most important thing to keep in mind is the context and/or purpose of the image. [46] The context and/or purpose determine if text alternatives are necessary and influence the information provided through text alternatives. A single image could have different text alternatives according to its usage (context) and purpose. [38, 64, 68] For example, if the purpose is to demonstrate a photography technique, it may be important to describe the lighting and shadows of the objects in the image. Otherwise, that information may not be important. There is no “right” or “absolute” text alternatives since it changes based on the situation. [68]

As the American Foundation for the Blind says, “If you can't think of something to say about an image, that doesn't mean it's meaningless.” [1] The Web Accessibility For All project at the University of Wisconsin-Madison recommended conducting an assessment before writing the text alternatives by considering the following: [62]

1. If the information is provided elsewhere in an alternate format.
2. If the information in the image is too extensive to be described easily.
3. If some information could be eliminated.

Text alternatives intend to provide the same information as the image itself. [68] WebAIM identified two classes of information that should be conveyed: content and function. [63] Image content is with regards to what is present in the image. Image function is the purpose that the image serves. The resulting text alternatives (consisting of information regarding both image content and image function) need to flow with the main document content. [11, 46]

There are online text-mining tools that formulated text alternatives based on the surrounding Web page content as well as the linked Web page if the image was a link. These text alternatives would then present the same information as the rest of the Web page content. W3C states that redundant text would be inappropriate for text alternatives. [68] As such, the text alternatives generated by online text-mining tools would be inappropriate but successfully fulfills the W3C requirement that text alternatives be provided.

Text alternatives must be “succinct, descriptive, and accurate.” [38, 53, 60] Petrie et al.’s study discovered that users might not read the entire text alternative if it was too long and detailed. However, they would prefer to have more information than to have less. [46] This way, they can access the information if they wished or they can choose to ignore it. Therefore, the ordering of the words and content is important so that users could properly judge the usefulness of the description and to get the most important information quickly.

If the user wishes to access more detailed information regarding the image, it is possible to inform the user where to locate such information. Petrie et al. found that users were willing to access external Web pages to retrieve additional details. [46] This shows that longer explanations can also be useful to users.

W3C makes some recommendations of models to use when describing images, such as Resource Description Framework (RDF) and Exchangeable Image File format (EXIF). RDF is metadata that uses the subject-predicate-object triple to describe relationships within an image.

[65] It would be useful for tools to automatically generate text alternatives based on the RDF subject-predicate-object triples. While RDF can be used to specify or record the relationships that exist in the image, it does not necessarily help users to know what relationships or other information exist within the image. It is more important to know what information to record than to have the method for storing the information.

EXIF specifies properties and technical information about the image file rather than the image itself. [33] EXIF specifications differ depending on the image format (such as JPG and TIFF), which can change and be deprecated over time. It also may not have considered all image formats; therefore, there may be some formats without specifications. The procedure for producing text alternatives should not be specific to particular technologies so that it can be applied to image types and formats that may be added in the future.

While the existing guidance stresses the importance of text alternatives for visual content and provides some recommendations on how text alternatives can be structured, it still lacks guidance on the information that should be presented in text alternatives and a method of extracting the information. This thesis attempts to fill that void.

2.2 Library Cataloguing

Library cataloguing sorts and organizes books according to subject and content. It tries to accomplish the same for images so that it is easier to locate the appropriate images. Library cataloguing helps to identify information contained within images and how to organize that information. The subject or information used for cataloguing the image can also be used to write text alternatives. In order to categorize the image (or write text alternatives), one must first know the information contained within the image.

As with books, images can be catalogued based on four facets: [26, 50]

1. **Who:** Who is the picture of? What is their gender? What is their (approximate) age? Is the person a symbol for other beings, objects, or ideas?
2. **What:** What are the objects or creatures doing? What is their condition or state? What emotions are being conveyed? What abstract ideas are being symbolized?
3. **Where:** Where is the image spatially located (geographical, architectural, space)? Does the place or location matter? Does the location symbolize another place?
4. **When:** What is the specific date or time period? Is the time period important?

The library cataloguing field recognizes that not every image would contain information from each facet or aspect of facets. “The categories are established and the questions are asked, in order to prevent possible subjects and details from being overlooked. It is perfectly possible to leave blank facets or facet aspects.” [50]

The meaning and interpretations of an image are distinct from its physical description (which is the image content) and, therefore, also needs to be captured. There are three levels of meaning: [26]

1. **Description:** What are the objects, persons, and actions depicted in the image? This is factual information.
2. **Analysis:** What is the context of the image? This may require expert knowledge in the specific topic area.
3. **Interpretation:** What is the intrinsic meaning? What emotions and feelings does it evoke?

Interpretation is very difficult to catalogue since it is subjective and changes depending on the person. It would be difficult to be consistent in the interpretation of the image; therefore, it is not used to catalogue images. In the case of text alternatives, users might not have the knowledge or ability to comprehend or interpret the image. Therefore, the interpretation of the image should also be considered for text alternatives.

An image contains a variety of information and a person may wish to have different information at different times. “The delight and frustration of pictorial resources is that a picture can mean different things to different people.” [50] Library cataloguing generally classifies books by subject. Because it is difficult to know exactly what the person is looking for at all times, the books must be classified under both generic and specific subjects. [26, 50] “With our approach to images, we tend to provide terms for specific things we see, as well as for the categories that those terms are part of. In essence, we supply both specific terms and broader terms.” [26] For example, an image of a bird could be classified under “bird” as well as the name of the bird depicted.

When identifying information, it may be advantageous to be as specific as possible since it is easier to generalize the subject than to make it more specific. Librarians are “trained to supply the subject terms that are as specific as possible.” [50] However, there are limitations on the depth of indexing and it is up to the person to decide when subject indexing should stop. The

same applies to text alternatives. At some point, it would be necessary to stop identifying information in detail and to start writing the text alternative.

Library of Congress categorizes for images by subject or contents of the image, as well as the image format. [36] This categorization is inappropriate for guiding users through the writing of text alternatives because librarians train for a long time to categorize images properly. The procedure for writing text alternatives must be usable without extensive training. Also, the categories used by the Library of Congress are used to tag an image in order to improve the probability of locating the image during a search. The category terms themselves do not form a comprehendible sentence, which text alternatives need to be. The terms, however, could be used as suggestions to users while identifying specific details about the image content, which would then be used to write informative and descriptive text alternatives.

Library cataloguing identified many types of information that may be communicated by any image.

2.3 Captioning and Audio Description

Captioning and audio descriptions for video are meant to convert auditory and visual content (such as dialog and movement) into text and audio so that it is accessible to people who could not hear or see the video, respectively. Text alternatives are similar in its intent to make visual content accessible so that people can understand the image. Standards and guidance regarding the information presented in captioning and audio description already exist. They could aid in identifying information to describe about an image as well as how to describe it.

WCAG 2.0 notes that “captions are similar to dialogue-only subtitles except captions convey not only the content of spoken dialogue, but also equivalents for non-dialogue audio information needed to understand the program content, including sound effects, music, laughter, speaker identification and location.” [66] Similarly, text alternatives may describe the mood or effect of the image intended on the viewer, the identification of objects, and the locations of objects.

According to the Canadian Association of Broadcasters, there are four main principles to captioning: accuracy (of information), responsibility (to preserve meaning and intent), consistency (in terminology and language), and clarity. [10] Text alternatives should also follow these principles in order to be useful to users.

Creating captions for video involves making choices regarding what to include and what to omit due to space and time limitations. Canadian Association of Broadcasters recommends a hierarchy of relevancy to determine the information to include and create better descriptions. [10] The hierarchy of relevancy categorizes information into three categories based on relevance. The primary descriptions are absolutely crucial for understanding; the secondary descriptions are important but not crucial; and the tertiary descriptions are the least frequently used. Based on the time and space constraints, the primary descriptions are included first, followed by the secondary descriptions, and then the tertiary descriptions. Text alternatives can have similar constraints and could follow a similar hierarchy of relevance or importance to determine which pieces of information should be provided to users.

Audio description (or descriptive video) is used not only for videos, but also in plays and operas, to name a few. "Audio description uses the natural pauses in dialogue or narration to insert descriptions of the essential visual elements ..." [4] Experts in audio description consider it to be an art and people need training in order to be a good describer. [3] To be a good describer of audio description, one must consider the information that needs to be provided and the length of the pause.

The standards for audio descriptions identifies the following types of information to be described: actions, expressive gestures and movements, physical appearances, people, places, objects, clothing, colour, light, texture, time shifts, facial features, attitude, decor, spatial relationships between characters, the weather, setting, temporal indicators, and textual information. [3, 4, 21] These types of information may also be described in text alternatives. The audio description standards also note that geometric shapes may have importance in the visual sense but loses its meaning in the auditory sense. [39] It may be true that the symbolic meaning behind a shape may be more important than the shape itself; however, it may also be important to know that a certain shape has a specific meaning within a specific context. Future references or mentions of the same shape in the same context could let the user know the meaning immediately even if the meaning is not explicitly given.

Given the limited time available to provide descriptions, the art of audio description could be perceived as "an exercise in what *not* to describe." [3] The translation process, as Metatla et al. calls it, consists of two steps.

The first step is to "discard any information not carrying explicit meaning that might affect comprehension of the essential information encoded in a given diagram." [39] According to the American Council of the Blind, much more information is omitted than included in audio descriptions. [3] Similarly, it will be necessary to determine which pieces of information to omit from text alternatives since some forms of text alternatives have character limitations. Given the vast amount of information that could be made accessible to the user, some information should be omitted for comprehensibility and manageability.

The second step is to "define a way to organise the preserved information to allow for appropriate possibilities for accessing and navigating such information." [39] Regardless of any limitations on the various forms of text alternatives, the information communicated by the image still needs to be presented in a way that is useful to the user. The audio description field has several guidelines on writing useful descriptions:

1. Be consistent in naming convention. [13] Referring to an object, person, or place by the same name every time can make the descriptions easier to follow.
2. Use succinct, vivid, and imaginative words. [3] There are numerous ways of saying the same thing. Using fewer, but descriptive, words can give users the same information in less time and space.
3. Describe what is most essential for the user to know in order to appreciate the image. [3, 4] Less important information can be included if time and space is available.
4. Provide general information before specific detailed information. [3] For a better understanding of the content, people often need or prefer to have an overview or a general sense of things before knowing the specifics. Providing details without a foundation or context could lead to unnecessary confusion.
5. Less is more. [3, 4, 21] Gagnon et al. conducted a study with participants with vision loss who shared, "too much information is worse than not enough, because in the latter case they can sometimes manage to get some grasp of the action from the sound track or from what follows, whereas in the former case, they sometimes become exhausted and lose track of what is happening." [21] The same problem could exist with text alternatives; therefore, care needs to be taken to refrain from overwhelming users with information.

All of these principles, which apply to captioning and audio description, can also apply to text alternatives. These principles have been used by media production institutions for many years. Following them would possibly create more informative and descriptive text alternatives.

2.4 Image Captioning and Image Indexing

Image captions are the short text descriptions that appear directly below images and are considered as part of the main document content. Existing guidance on image captioning can be used to guide the writing of text alternatives. While image indexing is outside the scope of this thesis, the methods and keywords used to locate images can help identify information contained within an image. Text alternatives can in turn be used for image indexing and image retrieval.

The Defense Visual Information (DVI) Directorate of the United States military's style guide for captions uses a two-step process for writing image captions. [12] It involves gathering information to include in the caption, followed by the construction of the image caption. This shows the importance of knowing the information communicated and contained in the image prior to writing text alternatives. The DVI Directorate recommends using the W5 system (i.e., Who, What, When, Where, and Why) to describe an image or video clip, similar to the types of information suggested by library cataloguing. The DVI Directorate suggests that the subject in the image, the action depicted, the date or time of day, and the geographic location be included in the caption.

In their discussion of Where, the DVI Directorate presented the concept of specific locations if known and generic areas or regions if unknown. [12] This once again raises the idea that there are different levels of abstraction and details that can be provided. This is discussed further in the image indexing field (below) and art description field (Section 2.5).

One new recommendation DVI Directorate made was to indicate when information was purposely being undisclosed. This is significant since the act itself becomes information and informs users that it was a deliberate decision to not share information. Similarly, if text alternatives are not provided for a reason, users should be informed of such a decision.

Research on image indexing and/or image tagging look at how images could be tagged in order to improve the search success rate. This is similar to library cataloguing images by subject keywords to increase the search success rate. People often search for specific images using a different set of keywords to locate the images. Experts in a particular field may use terminology

from their field while non-experts would use less-technical terminology to search for the same images. [15] This means a differing level of specificity is necessary when tagging images in order to improve the success rate. Similarly, a differing level of specificity could be used in text alternatives. Eakins and Jaimes each proposed their own levels of specificity. [15, 32]

Eakins identified three levels of information complexity in images. [15] First, the objective and primitive features that are derivable from the image without requiring external knowledge. Second, the derived or logical features that involve inferring the identity of objects (e.g., the Eiffel Tower and a double-decker bus). Finally, abstract attributes that require external knowledge to interpret the picture (e.g., name of the activity, event, or emotional significance).

Jaimes proposed a ten-level pyramid framework for the levels of visual attributes that is more specific than the one proposed by Panofsky, as depicted in Figure 2.1. [32] Visual attributes are classified based on syntax (e.g., colour and texture), semantics (e.g. meaning and events), perception (what the senses perceive), and conception (e.g. representation, abstract, or general idea). The pyramid framework allows for different levels of detail (abstract to specific) as well as relationships between different objects identified at the different levels. It also recognizes that not every image would need to be described at every level. Each level may require different levels of expertise to provide the information. This model does not, however, specify the amount of detail necessary to describe each image.

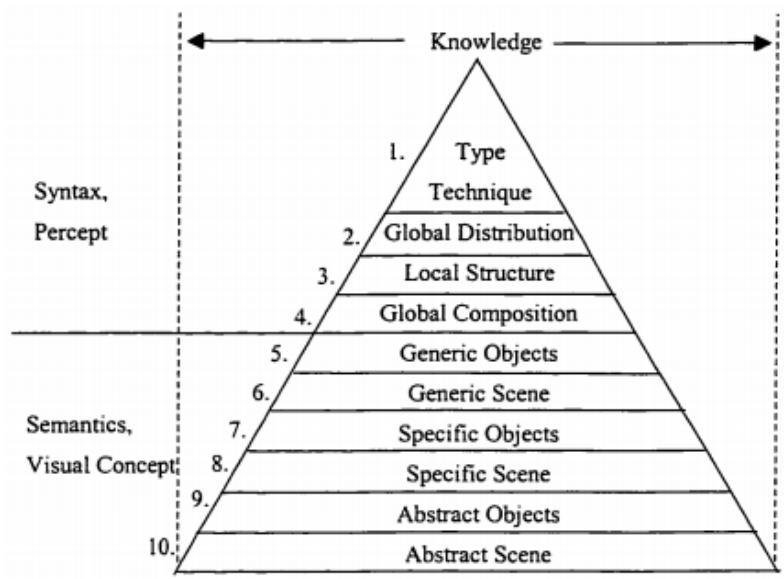


Figure 2.1 Pyramid for Indexing [32]

The biggest problems with image indexing are that it is labour and time intensive. It would take a long time and a lot of effort to identify and consider all of the information (and at different levels of detail). Eakins found that it took seven to forty minutes to index a single image. [15] These problems also apply to text alternatives composition since it is important to identify and consider the different types of information being communicated in order to determine what information to include or omit.

While image indexing and/or image tagging can be used to quickly identify keywords associated with an image, it has its limitations in its application to text alternatives. On their own, the keywords do not form actual sentences. Also, the keywords do not consider the context and purpose of the image; therefore, it may miss information needed for a particular use or it may contain information that does not apply.

Image indexing, combined with image captioning, can be a starting point for composing text alternatives. One missing element is a more comprehensive set of information presented in images (not only artistic images). This thesis attempts to present a set of information types that is as comprehensive as possible for the purpose of identifying information contained within an image.

2.5 Art Description

The field of art and art history focuses on analyzing and describing art. Over the centuries, several different models and methods of describing art have been developed. This section looks at some of those models and methods.

Roland from the University of Florida defined a five-step procedure for analyzing art pieces. [49]

1. Describe it, focusing on the appearance of the piece and the objects and people in the piece.
2. Relate it, focusing on connections or emotions the viewer has with the contents of the art piece.
3. Analyze it, focusing on establishing relationships between objects in the piece and meaning of elements of the piece.
4. Interpret it, focusing on events and actions depicted.
5. Evaluate it, focusing on the most important attributes about the piece.

While having the viewer relate to the piece (Step 2) may create a greater impact on the viewer, this relationship is highly subjective and may not be true for other viewers. It must be recognized that this information, although subjective, may be useful or helpful to those who are unable to interpret or comprehend the image. Even if presented with only one of many interpretations, the user would at least have an idea of what the image is about. It may be advantageous to present several views to illustrate the different possible interpretations.

Roland's procedure raises the significant concept of evaluating the piece of art or pieces of information in terms of importance. Although the procedures from other research areas mention that important information should be included in a description, they do not discuss the concept of evaluating the importance. This rating and evaluation of importance is understandably subjective and dependent on the person and situation. Roland did not discuss how to determine the importance of information. [49]

DiSimone described an old methodology in art history called formalism, which focuses on the form rather than the content of the work. [14] The methodology involves an outline of the form or subject matter for the piece (including the medium and techniques used) followed by a detailed description. This introduces the idea of providing an overview of the art prior to providing details when describing a piece of art. This idea could also be used when composing text alternatives. In formalism, directional words are used to guide the reader through the image, which may be used to quickly identify and focus on the section of the image being described. This shows that it may be important to indicate the location objects to direct the user's attention.

Doblin, as described by Hunter et al. [27], classified visual information into two categories: orthographic (which was alphanumeric or textual) and iconographic (which was visual). Both these categories of information may need to be conveyed through text alternatives. The iconographic category is further subdivided into "ideogrammatic (symbols that attempt to convey a single meaning, such as a road sign), diagrammatic (charts, graphs, or diagrams), and isogrammatic (… symbols that attempt to convey a visual representation of reality itself)" [27]. This model appears to focus on the manner in which the information is presented (through symbols or diagrams), but it does not consider the information itself being conveyed by the images.

Panofsky's theory on iconography of images also consists of three levels, focusing on the abstraction of information. [44] Level one is called the pre-iconographic level and describes

objects at a generic manner (such as “shoes”). Level two is called iconographical and describes specific objects (such as “glass slippers”). Level three is called iconological and describes the intrinsic meaning or content of the image (such as “Cinderella”). This method is used to analyze Renaissance art and would consider all three levels in the analysis. This method could be used to describe other images; however, such detailed analysis may not be necessary for all images.

Esaak believed that the content of an art piece was based on ideas. [18] It includes the artist’s intended ideas, the actual portrayed ideas, the viewers’ reaction to the intended and portrayed ideas, and the ways the art piece was influenced (such as by religion and politics). The concepts of intended versus portrayed ideas are something that the other models did not consider, but may also be important to describe to users in text alternatives. The users might not actually experience the intended ideas. Misinterpretation of the intent may affect the user’s impression. Therefore, it is important to consider both the intended and actual information.

The field of art and art history raised several concepts that were not brought up in other fields, that is, evaluating importance and considering the intended versus portrayed ideas. Both of these concepts are important to the process of composing text alternatives. It also further described different models of abstraction of information, which can be used when producing text alternatives.

2.6 Tactile Representation

Instead of text, images can also be represented in a tactile format called a tactile diagram. Current research has considered the types of information that could be presented in a tactile format. “Basic shapes and schematic renderings of a composition or object are translatable.” [34] It focuses strongly on information regarding the size and the location of objects within an image. [17, 25] This is understandable given that the user was exploring the image physically through touch. Most other types of information must be presented in a textual or auditory description.

Each tactile image uses the same methods to represent different types of information and a textual or auditory description accompanies the tactile image. [17] This textual or auditory description would act similar to a legend on a map. Similar to maps, the same symbol may have a different meaning on a different map; therefore, a legend is required to help users interpret the map properly. In the end, a textual description is still necessary to help communicate information about the image.

Depending on the complexity of the image, several tactile diagrams may be created to represent different aspects of the image. [34] An image is considered complex if there are too many objects or small details in the image. The resulting tactile diagrams would be a much simplified version of the original image and not contain the same information as the original image. By separating the image into several parts or components, the user can focus on each part individually.

Petrie et al. stated that the image analysis process to creating a tactile diagram consisted of classifying the type of image and then extracting the relevant information from the image. [47] For the accompanying textual description, Eriksson recommended separating the description into various stages or parts: "presentation of the relief image, general guidance so that the person who is to touch the picture will gain a general idea of it, before going into details." [17] The High Tech Center Training Unit in California went into further details and specified four stages to composing the textual description: give a general overview of the image, locate each element or component of the image, describe the image and elements in detail, and give a summary of the image. [25]

Tactile diagrams can be used to communicate an image's information but a textual description still must accompany the diagram. This shows the importance of a text alternative to images. In the process of creating tactile diagrams, the idea of segregating parts of the image to allow the users to focus on those parts was raised. This, along with possible ways of composing the textual description, can be used in the creation of text alternatives.

2.7 Image Type Specific Information

Many of the research areas discussed in this chapter mentioned various types of images, including photographs, advertisements, comics, and maps. [5, 8, 24, 27, 35, 37, 38, 45, 46, 61, 64] Some images were technical (such as UML diagrams and blueprints), some were statistical (such as charts and graphs), and some were artistic (such as paintings and sketches). The set of image types was vast and can be categorized in numerous different ways. This thesis does not attempt to compile a complete list of image types or to categorize them.

While there is little research on text alternatives for all types of images, there is research and suggestions regarding some types of images, such as paintings, photographs, charts, graphs,

technical diagrams, and maps. This section describes the existing guidance for these specific image types.

For paintings and photographs, the main focus is on the contents seen in the painting or photograph, the techniques used by the artist, and the intended message. It is suggested that the artist's name, the name of the painting or photograph, and the date it was created should be provided, along with a description of the place or event and other details that can be seen. [16, 22] The description can discuss angle, space, composition, patterns of objects depicted, as well as the gestures and expressions of the people depicted. [22] This indicates the importance of describing the physical appearance of objects and people in an image, along with the spatial relationships between those objects and people.

For charts and graphs, the focus is on the purpose and the types of data and relationships being communicated. The units of measurement need to be communicated along with specific (or approximate) values of significance and the trends that can be seen. The order in which the information is provided should be logical and consistent (such as left to right, top to bottom, and clockwise or counter clockwise). [41, 62]

For technical diagrams, such as UML and flow diagrams, Bennett and Edwards conducted a study to determine the types of questions non-sighted users asked when exploring technical diagrams. [6] They concluded that the focus needed to be on the elements (parts, nodes, or components) in the image and the relationships between the elements. An overview of the diagram was necessary to familiarize a person with the type of diagram being presented and the type of information that they will receive. The overview should be followed by a description of the elements or components and the connections between them. An increasing amount of information should be provided at the user's discretion as needed.

While Bennett and Edwards also found that information regarding the position and location of objects was not necessary for users to comprehend the diagrams, position and location information is important for maps (since the purpose of maps is to illustrate relative locations between points.) [6] The National Braille Association recommended selecting reference points and locating them in the image verbally. [41] These reference points should be described in a logical sequence. Other objects or locations can then be described in relation to the reference points.

To create a procedure that is applicable to a wide range of images, the information identified based on different types of images should be considered. The composition and presentation styles suggested could also be applied to writing text alternatives to other images. The procedure for creating text alternatives could also make such recommendations and guidance.

2.8 Requirements Based on the Literature

Many of the research areas described above identified the same types of information as important to describe and convey to users. They also had similar recommendations on how the information could be presented together. This thesis uses the knowledge from these research areas and applies it to producing informative text alternative.

This section summarizes the important concepts from the literature and extracts the requirements for producing text alternatives in order to help people to produce informative text alternatives. Each requirement is denoted with "REQ".

2.8.1 Essential Tasks to Produce Text Alternatives

In order to determine the information to convey or to not convey to users, it is important to first know the information being communicated. Therefore, an important step in producing text alternatives is to identify the information being communicated.

After the information is identified, another important step is to evaluate the importance of the information. As noted in the research areas of captioning and audio descriptions, the most important information needs to be presented first, followed by less important information if there was time and space available. To make this decision, the purpose and context of the image must be known and considered.

Several sources expressed that an image should be separated into parts or components. This would allow the person to focus on different parts of the image and provide additional details about each part. It may be important for identifying detailed information being communicated through and about different parts of the image.

After knowing the information being communicated, the information must be composed together into intelligible and usable text. The text should be written in a way such that it flows with the main document content. Due to time and space constraints, it is important for the text to

be succinct, concise, and vivid. This text should be tested by users to ensure its accuracy and usability.

REQ 1. Identify the purpose of the image.

REQ 2. Identify the context of the image.

REQ 3. Identify the information being communicated by the image.

REQ 4. Evaluate the importance of the information identified.

REQ 5. Separate an image into multiple parts in order for the person to focus on each part.

REQ 6. Organize the information identified in the image into intelligible and usable text.

REQ 7. The resulting text alternatives should flow with the main document content.

REQ 8. The resulting text alternatives should be succinct, concise, and make use of vivid words.

REQ 9. The resulting text alternatives should be tested by users.

2.8.2 Purpose and Context

The "purpose" of an image is the reason that the image exists in the document. It explains why the image is shown to the user. The purpose influences the importance level of the parts of the image, the importance level of each piece of information about the image, and the text alternatives produced.

The literature mentioned numerous possible purposes, including decorative, structure, navigation, advertisements, relating to the text, informational, logo, and linkage. [7, 9, 19, 38, 40, 43, 64] These purposes can be divided into (but not limited to) four categories:

1. **Decorative:** These images add visual appeal to the document. It may also be used to evoke emotion or create a mood or atmosphere to the document. For example, a bullet or a background image.
2. **Formatting:** These images provide structure to the document. For example, it could be used to separate the document into sections visually.
3. **Control:** These images allow the user to perform specific actions, such as navigate to another Web page or submit a form.
4. **Informative:** These images communicate specific information to the users. Most images will likely fall under this category.

An image can be used in multiple instances, for multiple and different purposes, and by different people. Depending on the context in which the image is being used, different versions of text alternatives should be produced to suit the specific context. The information that is important in one instance may not be important in another. For example, a photograph of a family portrait may focus on the people and the location where the portrait was taken. If the same image was used in a photography class, the focus may be on the camera settings and the lighting techniques used to create the photograph.

REQ 10. Emphasize the influences of purpose and context when producing text alternatives (i.e. different versions of text alternatives can be created for a single image).

2.8.3 Levels of Importance

The concept of evaluating the importance of information was raised within several research areas, including audio description and art description. It was recommended that the most important pieces of information (i.e., information that was critical to the understanding of the content) be provided first. However, the literature did not indicate what the levels of importance were nor how to rate the level of importance.

When writing text alternatives, people may focus on what they feel is important without considering the context or what the users may want to know about the image. They may need help and guidance on how to rate the levels of importance while considering the context.

REQ 11. Evaluate the importance of each piece of information about an image.

REQ 12. Identify the most important piece of information.

REQ 13. Communicate the most important piece of information first.

REQ 14. Identify the levels of importance.

REQ 15. Define the levels of importance based on the users' and content providers' points of view.

REQ 16. Provide guidance on how to rate the level of importance and what to do with information within a certain importance level.

2.8.4 Types of Information

Every research area described in Chapter 2 suggested many types of information that may exist within images. The suggestions were either vague (without clearly stating what information to include) or very specific to a particular type of image. There was no clear set of types of information that users should consider for all types of images.

Library cataloguing has the basic structure of “Who”, “What”, “Where”, and “When”. This structure is easily understood, therefore, it should be used to organize the different types of information contained in and communicated by an image. This structure can be further expanded to include “Why”, “How”, and “How Much”. All of the information identified in Chapter 2 can be organized into these seven categories.

REQ 17. Consider the seven categories of information when producing text alternatives.

REQ 18. Identify specific pieces of information within the seven categories of information that may be important for users to know.

REQ 19. Provide guidance regarding when certain categories of information may be of importance to the user.

2.8.5 Specific versus Generic Information

The information identified about an image can be generic or specific, and still accurately describe an image. For example, a general description could be “dog” and a specific description could be “poodle”. While both use the same number of words, the specific description is more descriptive. Within audio description (where there is limited time to provide as much information as possible), it is recommended that vivid, succinct, and imaginative words be used so that the user could receive a better mental image within the limited time. Similarly, vivid, succinct, and imaginative words can also improve the quality of text alternatives.

REQ 20. Acknowledge that both general and specific descriptions are accurate descriptions of an image. Discuss the benefits of being as specific and descriptive as possible (i.e. it can be descriptive while using the same number of words).

REQ 21. Use vivid, succinct, and imaginative words in text alternatives so that users can have a better mental image in a limited time and space.

2.8.6 Expertise Knowledge

Being knowledgeable about an image's usage and the content of the image can influence the information identified. The content provider of the document knows best as to why an image exists in a document. [20, 62] As previously mentioned, the purpose of an image highly influences the information identified and the resulting text alternatives. The content providers should be the one writing text alternatives for this reason; however, this is not always the case. Therefore, guidance is needed to support situations when the content provider is unavailable to create text alternatives.

Having the knowledge to properly interpret the image may also influence the information identified. If a person does not have the knowledge or ability to interpret the contents of the image, then they will not be able to provide as much information about the image. For example, a software engineer may be able to interpret a UML diagram while a biologist may not know the meaning of the various symbols used. Having the necessary knowledge to interpret the image would identify more information that may be useful to users.

Lacking the expertise to interpret the image is not an excuse to provide poor text alternatives. Even without expert knowledge, a person could describe what is seen within the image without interpretation. This is especially true if the purpose or context of the image does not require expert knowledge information to be identified. The goal is to be as descriptive as possible so that better text alternatives could be created.

REQ 22. Provide guidance such that people will be able to produce useful text alternatives even when they did not select the image.

REQ 23. Acknowledge the role of expert knowledge when identifying information about the image.

REQ 24. Emphasize the influence of the purpose and context of the image on the necessity to have expert knowledge.

2.8.7 Intended Versus Actual Information

While an image is chosen or created to convey certain pieces of information, users may receive different pieces of information, which may lead to miscommunication. This raises the concept of information that was intended to be conveyed versus information that the user

actually received. This is an important concept because knowing the intent can help users to interpret the image better and knowing the actual information received can help the content provider to re-evaluate the image and select a more appropriate one if necessary.

REQ 25. Acknowledge the importance of and difference between what is intended to be communicated and what is actually communicated.

2.8.8 Technology and Image Type Independent

The current set of guidance on text alternatives focuses on HTML and does not provide guidance for other types of electronic documents or technologies. There are other forms of graphical content (such as Flash videos) that also possess containers for providing text alternatives. Also, images and other graphical content are used in documents other than Web pages. Therefore, the guidance should support the creation of text alternatives regardless of the container and technology being used.

Much of the guidance on the types of information being communicated by an image is specific to the type of image. While it is possible to attempt to provide guidance on specific types of images, different people may call an image by a different name (for example, a photo versus a picture, or a diagram versus a blue print). There is no agreed-upon set of images or terminology of images such that it would be possible to organize the questions based on image types. Also, it would not be possible to guarantee that all image types are covered by the guidance. Providing a general approach does not constrain its application to the terminology. It would then be possible to apply the approach even in situations that may not have been previously considered.

REQ 26. Provide guidance that is not specific to a technology or platform.

REQ 27. Provide guidance that could be applied to a wide range of images.

Chapter 3 Procedure for Producing Text Alternatives

This chapter describes the procedure that was developed for producing text alternatives for images and summarized the contents of the ISO/IEC 20071-11 technical specification. [31] The procedure is designed to be applied to all static image types within all types of documents. Therefore, it does not specify or provide guidance with regards to specific technologies or documents that make use of images.

The procedure refrains from providing guidance regarding movies, as they are not static images and are outside the scope of this thesis. Guidance regarding movies can be found in the fields of audio description and captioning.

The procedure is a general approach that is not constrained by image types or terminologies of image types. While it is also possible to attempt to provide guidance on specific types of images, different people may call an image type by a different name (for example, a photograph versus a picture, or a diagram versus a blue print). There is no agreed-upon set of images or terminology of images to organize the guidance according to image type. It is possible that certain types would be missed if organized in this manner. In an attempt to be as inclusive as possible, the procedure does not specify image types but does include information that may exist within all image types. This satisfies REQ 27.

The procedure is not limited for use by content providers and/or Web page developers. Other persons, such as the average Internet user, can use this procedure to create text alternatives for their personal Web sites, for example. The procedure could also be incorporated into document authoring tools for all user groups to create text alternatives. The results of this procedure may vary due to the amount of time spent on applying the procedure and the amount and quality of information identified.

The intent of the procedure is to identify as much raw material about the image as possible before creating text alternatives. This way, the important information can be identified and chosen to be provided in text alternatives.

3.1 Details of the Procedure

As mentioned in Chapter 2, text alternatives are dependent on the purpose and context of the image. Therefore, text alternatives may change when the purpose of the image changes and with each usage of the image. [5, 38]

Based on the existing processes and procedures for describing art, captioning, audio or video description, and tactile representation, the following procedure was developed for producing text alternatives for images. Sections 3.1.1 to 3.1.5 elaborates on each step in the procedure.

Step 1. Identify the purpose that the image served within the document.

Step 2. Identify the image components within the image.

Step 3. Identify the image (or image component) content.

Step 4. Elaborate on the image (or image component) content.

Step 5. Organize the identified information into text alternatives.

Step 6. Evaluate the resulting text alternatives.

To illustrate some of the steps in the procedure, Figure 3.1 will be used.

Figure 3.1 Example Image

3.1.1 Step 1: Identify the Purpose

This first step identifies the reason the image exists in the document and explains the meaning of the image. [11] It influences the importance of the image components, the importance of each piece of information about the image, and the text alternatives produced. It answers the question "Why does this image exist in this document?" It is influenced by and interlinked with the context of the image. This satisfies REQ 1 and 2.

There are four categories of purposes: decorative, formatting, control, and informative. An image is not limited or restricted to one purpose category. An image may have several purposes for each usage. For example, an image can be decorative as well as informative. It is more useful to describe and explain the purpose rather than to classify the image into the four categories.

For the example image, the reason Figure 3.1 exists in this thesis is to illustrate certain steps of the procedure.

3.1.2 Step 2: Identify the Image Components

Several sources, including Ina [29], stated that an image can and should be broken down into several parts or image components. The process of breaking down the image into components can help the person focus on specific parts of the image and identify information being communicated by the individual components. Depending on the purpose of the image, different image components may be considered important to identify and describe. This satisfies REQ 5.

Identifying image components is an iterative process. It may be necessary to further break down an image (or image component) while identifying the image (or image component) content in order to provide more specific or detailed information. For example, one component may be called "group of stuffed animals" and it may become necessary to separate the group into individual stuffed animals.

Images can consist of one or more image components. Along with the image as a whole, each person, shape, object, text, landmark, or step in a sequence can be considered as an image component. For example, Figure 3.1 consists of three major components: the Whole Image, the plate of desserts, and the cup of coffee. Given the context of the image (which is to illustrate

certain steps of the procedure), each dessert on the plate of desserts could also be identified as major components while the spoons may be of less importance.

3.1.3 Step 3 & 4: Identify and Elaborate on Image Content

Both steps 3 and 4 involve identifying information about the image. This satisfies REQ 3 and 11. Once again, the intent of the procedure is to identify as much raw material for text alternatives as possible. It is recommended that the image information be identified through two iterations. During the first iteration, high level information is identified for each image (or image component), as specified in Section 3.2.1. During the second iteration, the high level information is elaborated for additional details about the image (or image component), as specified in Section 3.2.2. This satisfies REQ 17 and 18.

It should be noted that people could use a breadth-first approach (as recommended) to identify information or a depth-first approach, which would elaborate on the high level information as soon as the high level information is identified. They may also use a mixture of the two approaches. The purpose of these two steps is to identify as much information about the image as possible.

Depending on the context and purpose of the image, identifying image content may require expert knowledge in the field that the image is of. For example, an art historian would have more knowledge regarding a painting from the 18th century and can interpret the image. However, it may not be necessary to have that knowledge in order to comprehend the image, depending on the purpose and context of the image. This satisfies REQ 22 and 23.

During these two steps, it is important to rate the importance of each piece of information identified. As previously mentioned, the importance level of an image component or piece of image information depends on the purpose and context of the image. Therefore, the importance level can change with each usage of the image. Something that is considered unimportant in one instance may be considered important in another. This satisfies REQ 4.

3.1.4 Step 5: Organize the Text Alternative

Text alternatives can exist in several forms, including a short description, a long description, a caption, a tool tip, and the main document content itself. The form that it takes depends on the amount of information to communicate and the importance of the information, which is influenced by the purpose and context of the image. Different versions of text

alternatives may be written for the same image as a result. This satisfies REQ 8. Because each technology uses different forms or containers for providing text alternatives, this thesis will use the following terms in order to be non-technology specific:

- "primary alternative text" represents "the main method of providing text alternatives to users of screen readers." [31]
- "secondary alternative text" represents "other method(s) of providing text alternatives to users of screen readers." [31]

The information gathered from Steps 1 through 4 is organized to improve its readability and to determine the container in which the information would be presented. This satisfies REQ 6.

It is possible (and recommended) that several containers be used to provide different levels of detail about the image. Some readers may want a general idea of what the image is while others may need more specific details to fully understand the image.

This step involves:

1. Removing redundant information.
2. Allocating each piece of information to the appropriate container based on the importance of the information. In cases where the resulting text alternative is too long or detailed, it may be necessary to move less important information from the primary alternative text to the secondary alternative text.
3. Organizing the information in a logical, readable order.
4. Ensuring the compatibility of the text alternatives with the surrounding content and context of the image such that there is not redundant or conflicting information being presented. This satisfies REQ 7.

At this time, there is little guidance on how exactly to formulate text alternatives. It is based on the judgement of the person to determine which pieces of information are important to communicate. Below is a summary of recommendations for formulating text alternatives:

- State the purpose of the image.
- Provide general information followed by additional details.
- Describe objects in the image in a logical and consistent manner. For example, left to right or top to bottom.
- Be succinct and concise. Make use of vivid words. This satisfies REQ 8 and 21.

- Identify the most importance piece(s) of information and communicate it first. This satisfies REQ 12 and 13.
- Despite the recommendation that some images should not have text alternatives since the images could be ignored, users will not know that the image can be ignored unless it is explicitly stated. They will know only that text alternatives are unavailable. A person may see the image placeholder and wonder what the image is of. If the description is provided elsewhere, the text alternative should state where it can be found.

3.1.5 Step 6: Evaluate the Resulting Text Alternative

The resulting potential text alternatives are evaluated to ensure that it suitably describes the image given the context of the Web page. This satisfies REQ 9. The text alternative should be evaluated by both sighted and visually impaired users. Three different evaluations can be done. It is recommended that more than one evaluation be done.

One evaluation could be done by visually impaired users. Since the text alternatives will more frequently be used by visually impaired users, they are the most appropriate user group to perform the evaluation. They can determine whether or not the text alternatives flow with the surrounding content and if the text alternatives make sense. However, they may not always be available to evaluate each image's text alternatives.

While some images may be hidden from screen readers (and hence the screen reader users), a sighted person can still see image placeholders when an image is not loaded. Therefore, a second evaluation method would be for a sighted user to view the document with the images being not visible. For example, for a Web page, using a text-only Web browser or a Web browser with images not presented to view the Web page can help a sighted individual experience a Web page where text alternatives are needed.

One final evaluation method can be done by people who can see the image. They can help determine if the text alternatives truly represents the content of the image and communicates the important information (or identify important information that is missing).

Although it may be difficult to perform all three evaluations, performing at least one of these evaluations is better than none because it is a crucial part of the procedure. Based on the results of the evaluation, the procedure may be repeated to identify additional information or to re-write the text alternatives.

After successfully producing numerous text alternatives that are informative and useful to users, there may seem to be less need to evaluate every text alternative every time. However, it is good ergonomic practice to test all elements of an interface, including text alternatives. Therefore, text alternatives should still be evaluated occasionally to ensure its validity and usefulness. As previously noted, even if it is not possible or feasible to conduct all three evaluations, conducting even one evaluation is better than none at all.

3.1.6 Levels of Importance

In Steps 3 and 4, the individual pieces of information identified are rated in terms of importance. Depending on the importance level, the information may appear in a different container of text alternative or not presented at all. The importance level rating depends on the context and purpose of the image. Therefore, the importance level may change with each usage of the image. Images can present unique information, information that complements the main document text, visual appeal, and conflicting information.

In situations where the image presents unique information, this information adds to what is presented in the main document text. Additional objective and subjective information presented in an image is important to the understanding of the document.

In situations where the image complements (restates, modifies, elaborates, supplements) the main document text, the image might also contain information that is not provided in the main document text. Complimentary objective and subjective information might be important to the understanding of the document. A full understanding of the image can provide a better understanding of the document.

In situations where the image adds visual appeal, the image creates or modifies the mood of the document. The information might be focused on the subjective rather than the objective understanding of the document. Images used for visual appeal might have little or no objective information relevant to understanding the document. Subjective information might be important to or influences the understanding of the document.

Text alternatives for an image are intended to inform users of information that the image is communicating. There might be times when the image presents information that conflict with what is presented in the main document text. The conflicting information could be intentionally or unintentionally presented. If the conflicting information is intentionally presented visually, it should also be presented in text alternatives or the main document text. If the conflicting

information is unintentionally presented and in recognizing this, the decision is made to retain the image despite the conflicting information, then the conflicting information can be ignored in the text alternatives. This is not information that the content provider intended to provide and, therefore, should not be highlighted. It can also be ignored in hopes that sighted users would not notice the conflicting information.

There are three levels of importance (Essential, Significant, and Helpful) and one level of being unimportance (Not Important). This section satisfies REQ 14, 15, and 16.

3.1.6.1 Essential information

Essential information is necessary for the understanding of the image within the document in which it appears. Essential information should be presented either in the main document text (when referring to the image) or in the primary alternative text. Placing the information directly into the main document text ensures that all users had access to this information.

Essential information may have some or all of the following properties. As more of the properties apply, it is likely that the information is Essential:

- It is aimed at the target audience.
- It must be known in order to comprehend the document.
- Most people want or need it most of the time.
- The user would be confused as to what the document is talking about without this information.
- Without it, the user has no idea why the image is there or what the image is for.
- It provides a good first impression of the image.
- Based on this information, the user will determine if they need or want to know more about it.
- For the content provider, this is the information that the content provider absolutely wants to tell people about.
- It provides the essence, purpose, function, or intent of the image.
- It identifies that the image conflicts with the main document text and that this conflict is intentional.

3.1.6.2 Significant information

Significant information is useful for getting a comprehensive understanding of the image within the document in which it appears, when such information is desired by the user (based on the user's understanding of the Essential information). Significant information satisfies the more detailed interests of most users most of the time.

Significant information should be presented either in the primary alternative text or in secondary alternative text. The container that the Significant information is placed depends on the amount of Essential information already in the primary alternative text. It may be better to place some or all of the Significant information in the secondary alternative text so as to not overload the primary alternative text.

Significant information may have some or all of the following properties. As more of the properties apply, it is likely that the information is Significant:

- It is aimed at the target audience.
- It gives a more detailed and thorough understanding of the image and/or document.
- It is information that could be obtained by more than a quick glance.
- The user should know about it as they are reading the document in order to understand the document.
- The user decided to know more based on the Essential information. This information goes into more details about the Essential information.
- Without this information, the user has an idea of what the image is about and the reason the image is there, but does not have a detailed understanding about it.
- For the content provider, this is information that further explains and gives more details on what the content provider wants to tell the users.

3.1.6.3 Helpful information

Helpful information provides a more thorough understanding of the image within the document for users who wish for a more detailed description of the image. Because Helpful information is only of interest to some of the users some of the time, it should not be placed in the primary alternative text. It may be placed in secondary alternative text or in a separate document that is linked from the main document text or the primary alternative text.

Helpful information may have some or all of the following properties. As more of the properties apply, it is likely that the information is Helpful:

- It is specific details that might be of interest to some who are the target audience of the document.
- It is targeted towards very specific audiences (other than the target audience) or a subset of the target audience.
- It provides the user with a better understanding of the image when the user is not an expert in the topic area or not the target audience of the document.
- It might reassure the user that they have not missed something of greater importance.
- Without this information, the users have a fairly complete understanding of what the document is about but have some things that the users still want to know.
- It includes different or other possible interpretations of the information being expressed by the image.
- For the content provider, this is information that could clarify some things for some people.
- It includes optional extra information that is seldom wanted or needed, but elaborates on what is already there.

3.1.6.4 Not important information

Information is Not Important if it does not provide much additional understanding of the image or the document for any users. This could include information that is inappropriate given the context of the image. For most images, there are likely to be a number of questions that do not provide important information given the context of the image, resulting in Not Important information. Information that is Not Important should not be presented to users in either the main document text or text alternatives.

Information that is Not Important may have some or all of the following properties. As more of the properties apply, it is likely that the information is Not Important:

- Very few to no users want to know this information.
- It is rarely helpful.
- It is not important enough to mention.

- Without this information, the user knows everything they want or need to know in order to understand the document and/or image.
- This is information that might result in unintended confusion or boredom and does not help users understand the content provider's message.

3.2 Types of Information Within Images

It is important to know what information may be conveyed (and may be important to convey) to the user. [6] This information is then transformed into a textual medium so that those unable to comprehend or see images may receive and understand the content.

ISO 14915-3 identifies twelve categories of information that can be conveyed through all types of media: causal, conceptual, continuous action, descriptive, discrete action, event, physical, procedural, relationship, spatial, state, and value. [30] Depending on the type of information being conveyed, the standard recommends specific media types to communicate the information. However, this standard assumes that its information types are mutually exclusive, which results in focusing on only some of the information being conveyed by the image. ISO 14915-3 lacks guidance on what type of information is important to be transferred between media. Tang et al. described the importance of transferring information between media types, especially the importance of capturing secondary encoding information in this transfer. [59]

Based on the types of image information identified and described in Chapter 2, the following categories of image information were developed. It is a modified version of the W5H structure currently being used by library cataloguing systems. It focuses on What, Who, Where, When, How Much, and How. It includes the concept of having generic information as well as specific information. The structure also specifies questions that may lead people to identify the image information.

As with library cataloguing, it is recognized that each image may not contain information from all categories. The intent of this structure and set of questions is to be as inclusive as possible. It is recognized that many categories of information and questions may be ignored when describing any specific image.

3.2.1 Identify Basic Image Content

Basic image content consists of What and Who are in the image. This image content can be broken down into image classification, textual content, objective content, perceptual content, and subjective content.

The **image classification** can quickly tell the reader the kind of information to expect from the image description and the image itself. For example, knowing that the image is of a chart or graph would let people know to expect statistical information; or knowing it is a comic would let people anticipate something funny.

Text-only information is presented as images in some instances. It is highly recommended that text be presented as text in the document rather than as images. However, when it is presented as an image, it is important for people to know what the text states. It is crucial that the text be recorded exactly as written in the image.

Objective content involves information that can be seen in the image at first glance, including objects, people, and actions. It answers the question "What objects, persons, and actions are in the image (or image component)?" Elaborations on this content can be found in Section 3.2.3.1 and 3.2.3.4. Examples of objective content are "Tony playing the piano" and "tiger sleeping".

Objective content can range in detail, between highly generic and highly specific. The level of specificity depends on the person's knowledge of the image content and the importance of the information. Some situations required less detail than others. It is recommended that the information be as specific as possible in identification. The information can be generalized during the organization step (Section 3.1.4). An example of generic content is "car", while specific content would be "Toyota Yaris". While both descriptions are accurate, the use of specific content can be more descriptive while using the same number of words. This satisfies REQ 20.

Perceptual content specifies the low-level perceivable information in the image without specifying meanings, which is subjective content. Perceivable content includes (but not limited to) shape, colour, texture, size, and position. Elaborations on this content can be found in Section 3.2.3.2. Examples of perceptual content are "rectangle", "yellow", "bumpy", and "plaid".

Subjective content involves more subjective information than objective content, such as symbolic meaning, perceived and intended emotions, themes, concepts, theories, opinions,

judgments and other explanations that go beyond the obvious. It answers the question "What is the meaning of this image (or image component)?" Elaborations on this content can be found in Section 3.2.3.1. Examples of subjective content are "love", "happiness", and "courage". Subjective information influenced our perception and interpretation of what is presented. Since people who can see the image receive subjective information, this information should also be presented to those who cannot see the image. This information may differ between users.

3.2.2 Relationships

In addition to identifying image content, it is important to identify the relationships between the image and surrounding content, as well as the relationships between image components. This procedure focuses on three types of relationships: logical, temporal, and spatial.

Logical relationships explain which entities are interacting and how the interaction takes place. **Temporal relationships** explain when certain interactions take place relative to each other. This includes the time that an event occurs or the sequence of events that occurs. **Spatial relationships** explain where elements were physically located relative to each other. This can be with respect to the image to the surrounding document content, image components to the image, and image component to another image component.

When an image consists of multiple image components, knowledge of the relationships between the components can improve the understanding of the image. Elaboration on these relationships can be found in Sections 3.2.3.4, 3.2.5, and 3.2.5.2.

3.2.3 Elaborating on What

Basic information regarding "what" has been identified during Step 3 of the procedure. Step 4 of the procedure elaborates on this information and consists of elaborating on the physical object, perceptual content, and subjective content. It also identifies logical relationships and/or actions within the image.

3.2.3.1 Elaborating on Physical Object

Basic information regarding objective content has been identified in Step 3 of the procedure. The following questions regarding objective content should be considered for elaboration while writing text alternatives:

- What is the object?
- What is the brand / model / part name (number) of the object?

3.2.3.2 Elaborating on Perceptual Content

Basic information regarding perceptual content has been identified in Step 3 of the procedure. The following questions regarding perceptual content should be considered for elaboration while writing text alternatives:

- What is (are) the color(s) of the image (or image component)?
- What is the shape the image (or image component)?
- What is the size / dimensions of the image (or image component)?
- What is the texture of the image (or image component)?
- How is the image (or image component) positioned? (e.g. sideways, at an angle, facing left)
- What other perceptual information is important for users to know?

3.2.3.3 Elaborating on Subjective Content

Basic information regarding subjective content has been identified in Step 3 of the procedure. The following questions regarding subjective content should be considered for elaboration while writing text alternatives:

- What concepts are associated with the image or the image component?
- What is the image or image component representing or symbolizing?
- If the colour(s) of the image or image component is important, what is the colour(s) representing?
- What themes are represented?
- What emotions are being expressed?
- How is the user expected to respond emotionally (with feelings, judgments, and opinions) to the image?
- What other subjective information is important for users to know?

  This satisfies REQ 25.

3.2.3.4 Elaborating on Logical Relationships and Actions

Logical relationships describe which elements are interacting and how they interact with each other. This includes relationships expressed in technical diagrams. The interaction often involves a subject (performing the action) and an object (the action being performed onto). For example, in “Dad is cutting the turkey”, “Dad” is the subject, “the turkey” is the object, and “cutting” is the action.

It is not always the case that there is always both a subject and an object in the image. Sometimes, there is only a subject with the object implied or vice versa. For example, an image of a girl running with a number pinned to her shirt has the subject only. If the context of the image is that it is a marathon, then it implies that the girl is running towards the goal line, an object that is not visible in the image. The following questions regarding logical relationships and actions should be considered while writing text alternatives:

- What interaction or action is taking place?
- What or who is the subject of the interaction or action?
- What or who is the object of the interaction or action?
- How is the interaction or action being performed?
- What is the intended result of the interaction or action?
- What other information about the interaction or action are important for users to know?

3.2.4 Elaborating on Who

Basic information regarding “who” has been identified in Step 3 of the procedure. Elaboration on “who” identified particular details about the person. The following questions regarding people should be considered while writing text alternatives:

- Who is the image or image component of?
- What does the person look like? (i.e. age, gender, nationality, hair colour, eye colour, hair style, etc.)
- What is the facial expression of the person?
- What is the person doing?
- What position is the person in? (ex. Standing with hands across the chest)
- What other information about the person is important for users to know?

3.2.5 Elaborating on Where

Information regarding "where" can be categorized into three types: location information, spatial relationship information between image components, and spatial relationship information between the image and the surrounding document content.

3.2.5.1 Elaborating on Locations or Places

Location information specifies the place being depicted in the image (e.g. the beach, Paris, and inside the school gymnasium). The following questions regarding locations or places should be considered while writing text alternatives:

- What is the setting, place, or location depicted in the image (or image component)?
- What specific landmarks are visible in the image (or image component)?
- What other information about the location is important for users to know?

3.2.5.2 Elaborating on Spatial Relationships Within the Image

This information is with regards to the physical location of the image components relative to each other. The following questions regarding spatial relationships within the image should be considered while writing text alternatives:

- Where is the image component spatially located within the image?
- Where is the image component relative to other image components?

3.2.5.3 Elaborating on Spatial Relationships Within the Document

This information is with regards to the physical location of the image relative to its surrounding content. The following questions regarding spatial relationships within the document should be considered while writing text alternatives:

- Where is the image spatially located within the document?
- Where is the image located with regards to the content that describes it?

3.2.6 Elaborating on When (Temporal Relationships)

There are two types of information relating to time: information being presented within the image and information regarding the changes to the image over time. Information presented within the image includes time period, event, and sequential relationships. Information relating to changes over time includes state information and the use of slide shows.

3.2.6.1 Elaborating on Time Periods

Time period refers to when an image is taken or the time period that the image depicts. There are different ways of defining or describing the time period. It can be described by an actual date and time (the image was created), the time of day or year being depicted, or the historic period.

3.2.6.2 Elaborating on Events

It is a frequent occurrence that an image is used to commemorate an event, or occasion. An image captures a specific instance in time. The question this type of information answers is "What event is the image depicting?"

This information may appear to be similar to those regarding actions or logical relationships (Section 3.2.3.4). However, this information is specifically focused on the event or occasion where the action is being performed. Furthermore, it is better to identify redundant information to ensure it has been captured than to miss potentially important information. The redundant information can be removed or combined when organizing the text alternatives.

3.2.6.3 Elaborating on Sequential Relationships

There are some images that illustrate or demonstrate a process or sequence of events. Such images include assembly instructions, flow diagrams, and structure charts. The specific sequence illustrated is important to convey to users as it often tells the user how something works or the order that something needs to be done.

There are several types of sequential relationships: linear (an element leads to another), branching or hierarchical (one element lead to multiple elements), and cyclical and/or networked (any element can lead to every other element in the set, including itself). The relationships can be either unidirectional or bidirectional.

The following questions regarding sequential relationships should be considered while writing text alternatives:

- What is the basis of the sequential relationship?
- What types of sequential relationships are involved (linear, branching, cyclical, one directional, bi directional)?
- If there a start and/or end point to the set of relationships, what are the start and/or end point(s)?

- What are the individual steps (components) of the relationship?
- What is a suitable basis for logically ordering the individual steps / components?
- How is each step / component related temporally to other steps / components? This can include:
    - connections to previous and following steps / components
    - time involved in individual steps / components or for the transition between steps / components
    - logic (decision or event) involved in moving from one step / component to another
- What other information about the sequential relationships is important for users to know?

3.2.6.4 Elaborating on States

Individual parts of the image may change in order to communicate specific information regarding the current conditions (e.g., an icon to indicate the current weather conditions or an indicator on a device for battery charging status). The following questions regarding states should be considered while writing text alternatives:

- What different states (values) can occur for the image component?
- How are the different states (values) visually represented?

3.2.6.5 Elaborating on Slide Shows

Some images are used as part of a set of images. Sometimes, those images together can portray an action or animation. At other times, the set of images are completely separate images with a common theme.

The following questions regarding slide shows should be considered while writing text alternatives:

- Questions relating to the entire slide show
    - How many separate images does the slide show consist of?
    - How can the user control (e.g. interrupt, go back) the playing of the slide show?
- Questions relating to an individual image
    - What is the position of each image within the slide show?
    - What time interval will the image be presented for?
    - What action or motion is being portrayed in transitioning to the next image?

3.2.7 Elaborating on How Much

This type of information contains quantitative values. Charts and graphs are common images that portray such information. The following questions regarding quantitative information should be considered while writing text alternatives:

- What is the quantity?
- What is the quantity associated with? / What does the quantity represent?
- What is the unit of the quantity?
- What is the precision (or statistical significance) of the quantity?
- Is the quantity fixed or dynamic?

3.2.8 Elaborating on How

Some documents allow the user to make changes or open other documents. Buttons (in the form of images) are sometimes used to provide users with the controls to perform such actions. Images with linkages are one popular manner of doing this. The following questions regarding controls should be considered while writing text alternatives:

- How is the user supposed to use the image (or image component)?
- What action does the user perform to interact with the image (or image component)?
- What is intended to result from interacting with the image (or image component)?
- What can go wrong in the interaction?

3.3 Status of this Procedure

The procedure described above led to a Canadian contribution to the International Standards Organization (ISO) committee on User Interfaces Accessibility (ISO/IEC JTC1/SC35). This work was presented to SC35 and was internationally accepted by other usability and accessibility experts. This work is the basis for the technical specification ISO/IEC 20071-11 entitled “Guidance on text alternatives for images”. [31]

The technical specification is currently under development. It will be updated to incorporate the results of this thesis, including general guidance regarding when certain categories of information may be important to the user. This would satisfy REQ 19.

All of the requirements identified in Section 2.8 have been included and satisfied within the procedure described in this chapter.

Chapter 4 Research Study of Text Alternatives Without the Procedure

Before evaluating the effectiveness of the procedure in identifying the important information about an image and in creating text alternatives, it was important to know the kind and quality of alternative text being generated without the procedure. The purpose of this research study was to determine a baseline for the amount and quality of information people normally write for text alternatives when they were not given a procedure to follow.

The results of this study were presented and published at the HCI International conference 2011 [58], and the Annual Meeting of the Human Factors and Ergonomics Society (HFES) 2011 [57].

This chapter describes the research questions, participants, methodologies, results, and the analysis of results for this research study.

4.1 Research Questions

This research study was designed to answer the following questions:

1. What information do people normally include in text alternatives (when they are not given a procedure)?
2. What information do people normally miss in text alternatives (when they are not given a procedure)?
3. When given the option to use the empty text (i.e., “”), how often do people make use of the option?
4. What are the reasons for using the empty text? What are the reasons for not using the empty text?
5. How much time do people normally spend on writing text alternatives (when they are not given a procedure)?

6. What is the quantity of information that people identify for images (when they are not given a procedure)?
7. What is the quality of the text alternatives (when people are not given a procedure)?

In addition to creating a baseline for the amount and quality of information in text alternatives, this research study attempted to identify some of the reasons for the current quality of text alternatives and where people needed help in order to generate more informative text alternatives.

4.2 Methodology

4.2.1 Participants

For Web content, it is often the person creating the Web page who was responsible for writing the text alternatives. This usually means the Web site developers. Since this research study intended to determine the current status or standard of text alternatives being created and it is often the Web site developers creating text alternatives, this research study was targeted towards developers. Given the small sample, the participants may not be true representatives of their population. However, the results can identify possible suggestions and recommendations that may be true of the population.

Ten developers volunteered for this research study, ranging from Information Technology staff to database administrators to software developers. Similar to the Web site developers responsible for writing text alternatives, these volunteers were busy people with many other important tasks to do for their jobs. While they volunteered to write text alternatives for this research study, they were told to create text alternatives for images that they did not choose themselves. They may have had more incentive to write text alternatives but they also did not have any help to create text alternatives.

4.2.2 Materials and Execution

The group of ten developers were provided with a single Word document that consisted of a brief introduction to text alternatives and their purpose, a set of five images (along with a link to the original Web site where the image resides, providing them with context), and a feedback section. For each image in the set, the participants were asked to write the text

alternatives and record the amount of time spent doing so. The participants were not given examples or instructions on how to produce the text alternatives. The participants were then presented with the option to use the empty text (i.e., “”) instead of what they had already written. For each image in the set, they were asked if they would use the empty text and to explain their choice in each case.

The participants were randomly separated into two groups. Each group completed the research study on a different set of five images (Set A or Set B). The images in each set can be found in Appendix A. The images will henceforth be referred to as A1 to A5 and B1 to B5 to represent the five images within each set. The participants who completed the research study on Set A images will be referred to as “Group A participants”, similarly for “Group B participants”. A sample of the materials given to the participants can be found in Section B.1 in Appendix B.

4.2.3 Method to Answer Research Questions

Research Questions 1 and 2

In this research study, the participants did not explicitly specify the type of information and question within the set of questions that their information answers. Therefore, it was done at the discretion of the researcher. The researcher identified the specific question(s) within the set of questions that would or could have resulted in the information written in the text alternatives.

For each text alternative written, the researcher considered if the information identified Why, What, Who, Where, When, How Much, and How. If so, the researcher determined which specific question within the set of questions the information could derive from. If the information could be derived from more than one question, then all possible questions from the set of questions would be marked as answered. If the information could not be derived from any question in the set of questions, then it was noted to possibly add a question for it.

Components were also not explicitly identified by the participants and were done at the discretion of the researcher. A component was considered as identified if there was a significant amount of detail about that component written in the text alternative. A mere mention of a component’s existence was not considered as an identification of the component.

Based on this process, the types of information that people normally include or miss in text alternatives (research questions 1 and 2) were identified. The results are shown in Section 4.3.1.

Research Questions 3 and 4

The participants' choice to use the empty text (yes, no, or maybe) were tallied to determine how often the empty text was used (research question 3). Their explanations as to their choices were then analyzed for commonalities or trends between the participants (research question 4). The results are shown in Section 4.3.2.

Research Question 5 & 6

The averages, minimums, and maximums of the number of questions answered, the time spent, and the number of words written were then calculated (research question 5 and 6) so that they can be used for comparison against the future research studies. This was done to determine the effectiveness and efficiency of the procedure to produce informative text alternatives compared to without the use of the procedure. The results are shown in Sections 4.3.1, 4.3.3, and 4.3.4.

Since the amount of time spent on writing text alternatives for each image was reported by the participants themselves, it was possible that the times were not accurately reported and may be estimates of how long it actually took. However, the reported times were good enough for use to determine the major differences between the research studies.

While the number of words written was one measurement for the quantity of information (research question 6), it was not expected to be as an accurate measurement of the quality of the text alternatives that were written as other measures. It is possible to improve the quality of the words while the number of words remained the same. The number of questions answered was also used to measure the quantity of information. These calculations were focused on the quantity rather than the quality of information, which is research question 7.

Research Question 7

Based on the results of this research study alone, it was not possible to evaluate the quality of the results. To determine the quality of the text alternatives written by the participants (research question 6), the importance of the information provided in the text alternatives need to

be evaluated. The evaluation was done in the Research Study of the Identified Information (Chapter 70). The percentages of information rated as Essential, Significant, Helpful, and Not Important that was identified by the participants in this research study were calculated. An average, maximum, and minimum percentage was calculated for each importance level. The total number of pieces of information identified by the participants was also determined.

The specific types of information that was frequently missed and frequently identified within each importance level were calculated by looking at the percentage of participants who identified a piece of information. If more than 50% of the participants identified a piece of information, it was considered as more often identified. If less than 50% of the participants identified a piece of information, it was considered as more often missed. The total number of instances was tallied for each type of information.

The results are presented in Section 4.3.4.

4.3 Research Study Results

4.3.1 Types of Information Included and Missed

It was identified that the participants answered a combined total of 165 questions. Out of all the questions answered, the question identifying the object(s) in the image (Physical Object) was answered most often (24.8% of the total questions answered). The next most often question answered was identifying the type of image (Classification), which was 13.3% of the total questions answered. Some of the other more popular types of information identified included the image's purpose (Why) and the physical appearance of objects or persons, focusing mainly on colour (Perceptual and Who). This information is summarized in Table 4.1.

Table 4.1 Types of Information Most Often Identified Without Using a Procedure

Question	Category	Percentage
What is the object in the image/component?	Physical Object	24.8%
What kind of image is it?	Classification	13.3%
What is the image being used for?	Why	7.3%
What is (are) the colour(s) of the image/component?	Perceptual	5.5%
Who is the image/component of?	Who	4.8%
What does the person look like?	Who	4.8%

Because there were many individual questions from the question set which were not answered by any participants or were answered only once, the types of information missed will

be described by category or sub-category of information for simplicity. According to Table 4.2, although the information existed within the image, the following categories and sub-categories of information were more often missed when writing text alternatives without the help of a procedure: Textual, Logical Relationship, How, Sequential Relationship, Where, Spatial Relationship, How Much, Subjective, and When.

Table 4.2 Usage of Categories and Sub-categories of Information Identified Without Using a Procedure

Category or Sub-category	Percentage
What – Textual	0.0%
What – Logical Relationships	0.6%
How	0.6%
When – Sequential Relationships	1.2%
Where – Location/Place	2.4%
Where – Spatial Relationships	4.2%
How Much	4.2%
What – Subjective	6.1%
When	6.1%
Why	10.9%
What – Perceptual	10.9%
What – Classification	13.3%
Who	14.5%
What – Physical Object	24.8%

After mapping the text alternatives to the set of questions, it appeared that the participants answered an average of 3.4 questions, with a standard deviation of 2.7 questions, a minimum of one question, and a maximum of 15 questions. Figure 4.3 shows a distribution of the results. The actual values can be found in Table C.2 in Appendix C.

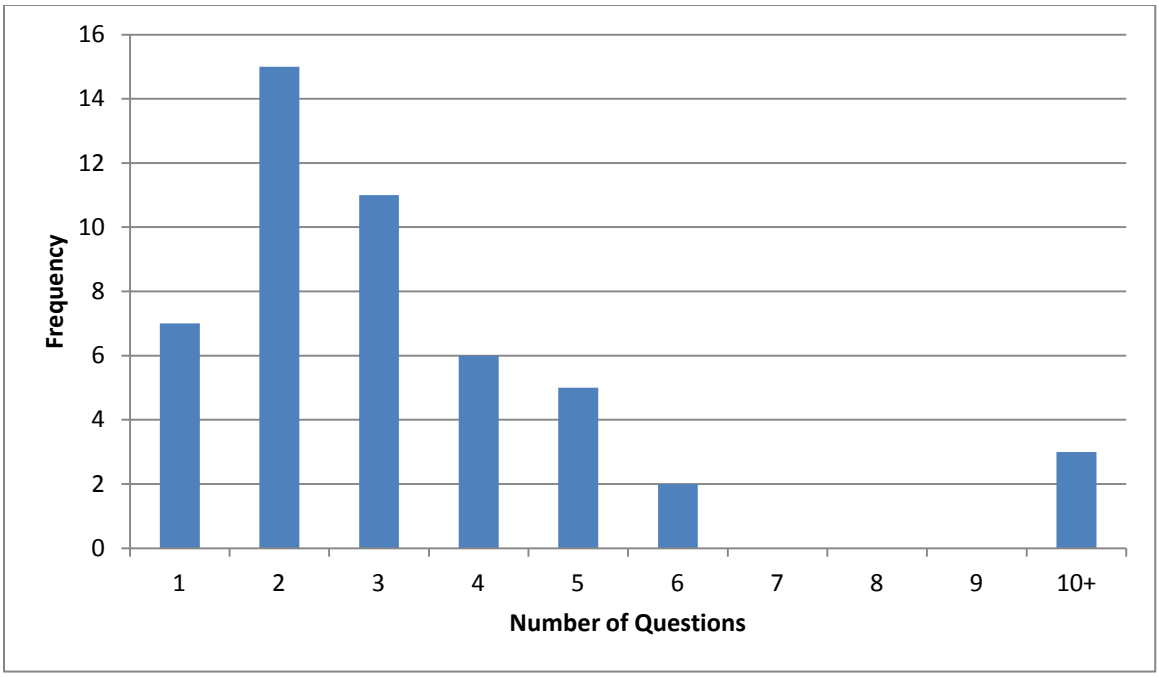


Figure 4.3 Number of Questions Answered Without the Use of the Procedure

4.3.2 Empty Text Usage and Reasons

The participants' answers regarding the usage of the empty text consisted of definitive yes, definitive no, and undecided or maybe. Thirty percent of the time, participants said that they would use the empty text. They also did not know or debated using the empty text 6% of the time. Sixty-four percent of the time, the participants said that they would not use the empty text. Table 4.4 summarizes the responses made by the participants.

Table 4.4 Empty Text Usages

	Use ""	Not Use ""	Maybe Use ""	Total
Make use of empty text	15	32	3	50
Percentage	30%	64%	6%	100%

The most common reason for using the empty text was the difficulty in coming up with the text alternative (73.3% of the time). They did not feel that the text alternative they had written was good enough but since they could not develop a better one, the empty text was preferable. Similarly, the participants said that they would not use the empty text if it was easy for them to come up with the text alternatives (43.8% of the time). If they knew what to write for

text alternatives or felt that they could come up with a good text alternative, then they would not use the empty text (9.4% of the time).

Three participants said that if the information was presented elsewhere, then it would be okay to use the empty text. However, one participant said that it was possible to at least include the title of the image as the text alternative. That is, it was easy to duplicate information that already existed in the document.

Two participants said that the empty text would be used if the image was not relevant or when no useful information was being conveyed by the image. They believed that if there is some information (whether it was important or useful) being conveyed by the image, then the empty text was not appropriate (46.9% of the time) and text alternatives should be provided.

Three participants noted that other methods of providing information may be more suitable than text alternatives, such as the use of tables to present numerical data and the use of audio.

Table 4.5 provided a summary of the reasons provided for choosing to use or not use the empty text in place of the text alternatives the participants had originally written.

Table 4.5 Summary of Reasons for Using or Not Using Empty Text

	Use “”	Not Use “”	Maybe Use “”
Difficult to write	73.3%		
Easy to write		43.8%	
What was written is good enough		9.4%	
Need text alternatives		21.9%	
Information irrelevant or text alternatives not needed	20.0%		33.3%
Possible to convey certain information		25.0%	
Use other methods instead	6.7%		66.7%

4.3.3 Amount of Time Spent

The participants were asked to record the amount of time it took for them to write each text alternative. Out of the fifty instances, only one instance was not reported. The participants took an average of 2.9 minutes, with a standard deviation of 2.4 minutes, a minimum of 0.5

minutes, and a maximum of 10 minutes to write text alternative for an image. Figure 4.6 shows a distribution of the results. The actual values can be found in Table C.10 in Appendix C.

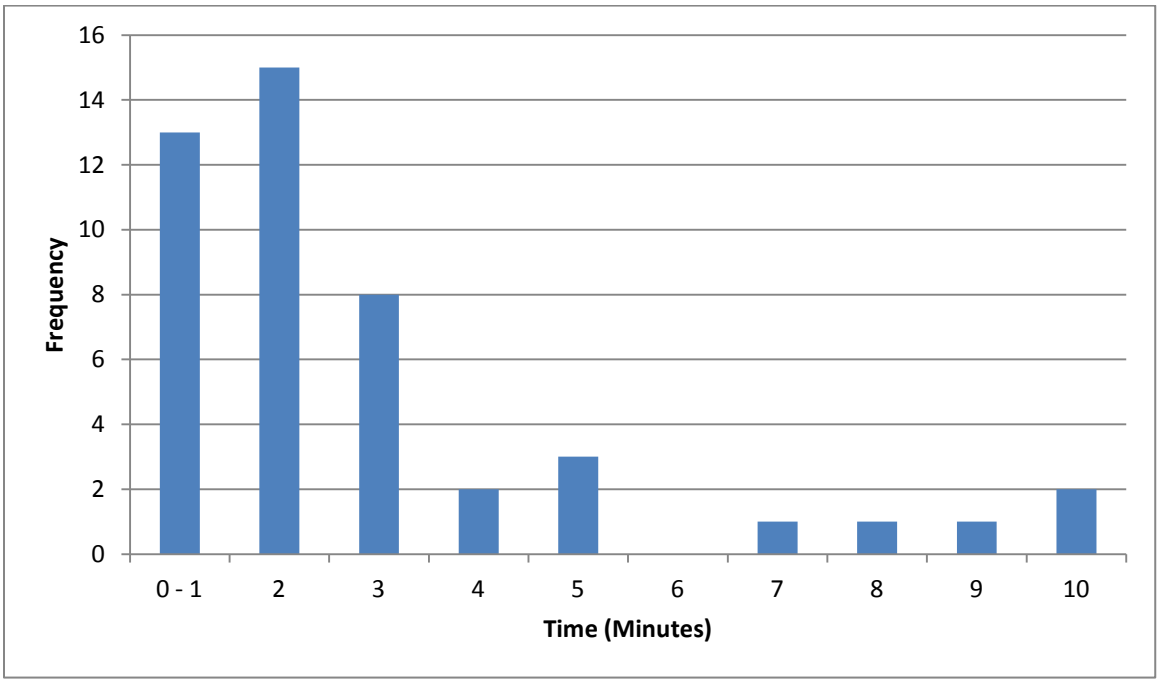


Figure 4.6 Amount of Time Spent Without the Use of the Procedure

Only one participant spent close to 10 minutes on average writing text alternatives, while the remaining participants spent less than 5 minutes on average to write a text alternative. However, four participants had spent a maximum of 5 to 10 minutes to write text alternatives for a given image during this research study.

4.3.4 Number of Words Written and Quality of Responses

A word processor program was used to count the number of words written by a participant for a given image. Based on these results, the participants wrote an average of 23.2 words, a minimum of 3 words, and a maximum of 116 words for a given image. Figure 4.7 shows a distribution of the results. The actual values can be found in Table C.18 in Appendix C. While a standard deviation is inappropriate since the data is not normal, 90% of the data were in the range of 3 to 53 words.

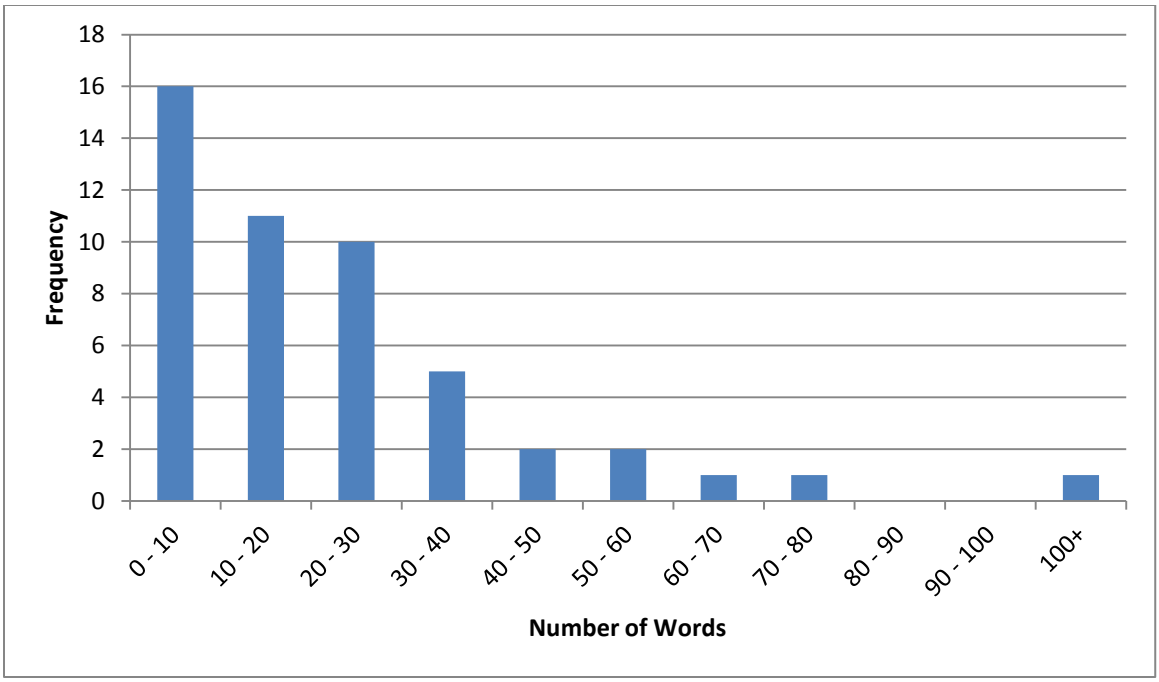


Figure 4.7 Number of Words Written Without the Use of the Procedure

Two participants wrote fewer than 10 words on average, six participants wrote between 10 and 25 words on average, and two participants wrote more than 45 words on average. In terms of the least number of words a participant wrote for any given image, eight of the participants had written fewer than 10 words for at least one image.

When the participants did not have the help of a procedure, they identified an average of 14.6% of the information that was rated as important to some extent based on the results of the Research Study of the Identified Information. This involved an average of 31.8% of the Essential information, 14.9% of the Significant information, and 4.8% of the Helpful information. Participants also identified an average of 2.2% of the information which was later considered Not Important. Table 4.8 illustrates more detailed results.

Table 4.8 Summary of Quality of Information Identified Without the Procedure

	Essential	**Significant**	**Helpful**	**Not Important**
Average (percent)	31.8%	14.9%	4.7%	2.2%
Average (#identified/total)	29.8 / 94	12.5 / 84	7.8 / 167	1.1 / 50
Maximum (percent)	49.2%	26.9%	19.6%	10.0%
Minimum (percent)	5.0%	3.4%	0.0%	0.0%

Table 4.9 lists the top five types of information that the participants often missed within each level of importance when they were not provided with a procedure. Table 4.10 lists the top five types of information that the participants often included within each level of importance when they were not provided with a procedure. With respect to the types of information participants included that was rated as Significant and Helpful, no other types of information were included in the text alternatives within these importance levels when the participants were not given a procedure.

Table 4.9 Types of Information Often Missed Without a Procedure by Importance Level

Essential	Significant	Helpful
1. What – Perceptual 2. What – Physical Objects 3. What – Subjective 4. When – Sequential Relationship 5. What – Textual	1. How Much 2. What – Physical Objects 3. What – Perceptual 4. What – Subjective 5. Where – Spatial Relationships	1. What – Perceptual 2. Where – Spatial Relationships 3. What – Subjective 4. When – Sequential Relationships 5. What – Physical Objects

Table 4.10 Types of Information Often Identified Without a Procedure by Importance Level

Essential	Significant	Helpful
1. What – Classification 2. What – Physical Objects 3. Why 4. Who 5. What – Subjective	1. What – Physical Objects 2. What – Perceptual 3. When – General 4. What – Subjective 5. Who	1. What – Physical Objects 2. Why 3. What – Subjective 4. Where – Spatial Relationships

4.4 Analysis of Research Study Results

4.4.1 Analysis of Information Included and Missed within Text Alternatives

The participants identified mainly Physical Objects and Classification information, with some mention of Who, Perceptual, and Why information. These types of information can be used to provide a general overview of the contents and purpose of the image. The images also contained and communicated Textual, Subjective, Logical Relationships, Where, Spatial Relationships, When, Sequential Relationships, How Much, and How information, but the participants rarely identified such information. It indicated that participants rarely provided a text alternative that communicated all of the same relevant information as the image (which is the

intent and purpose of text alternatives). While a summary may help people to determine whether the image contained important information, it might not give the important information itself.

When explaining the reasons for using the empty text, the participants expressed that more descriptive text alternatives (than what they had written) were necessary, which would then include more different types of information. However, they did not know what or how to write better text alternatives. Therefore, a procedure and guidance could help people identify more details and expand text alternatives to include more types of information.

4.4.2 Analysis of Reasons for Using Empty Text and Frequency of Use

The results showed that the participants would not use the empty text the majority of the time. The participants' decisions not to use the empty text were made after they had already written something for as the text alternative. It is understandable that in such cases, people would likely choose to make use of what they had already spent time creating. However, it may be a different story if the participants were given the choice to use the empty text prior to creating their own version of text alternative.

Despite the finding that the participants would not use the empty text 64% of the time, research studies indicated majority of images on the Internet do not have text alternatives. [2, 8] The participants explained some of the reasons why they would choose to use the empty text or not, including the ability of the person to create text alternatives quickly and the possession of knowledge to write what they feel is good text alternative. These reasons may explain why so many images on the Internet lack text alternatives.

The ability of the person to quickly create text alternatives highly influenced whether or not text alternatives were written. It identified two aspects of the problem, that is, ability and knowledge to write good text alternatives and time to write text alternatives. It must be easy (and quick) to know what information to include in text alternatives and then to generate the text alternatives. Otherwise, the person would use the empty text. First, the procedure can help identify the information that could be included in text alternatives and to formulate the text alternatives, which ensured that the person will be able to come up with good text alternatives. Second, additional research was needed to improve the amount of time to generate text alternatives using the procedure.

While repeating existing information (such as a caption or title) is easy to do and fulfills the requirement of providing text alternatives, the text alternatives would not serve its purpose,

which is to provide equivalent information in a textual format. This practice should be avoided. The issue remains that the person did not know what information to include in text alternatives other than what was already written. A procedure then can help identify information that could be provided in text alternatives.

While the participants recognized that text alternatives should be written if there was relevant information being communicated, the empty text could still be used if the person did not know what to put as text alternatives. The procedure can help to identify information to be provided through text alternatives, avoiding the use of the empty text.

4.4.3 Analysis of Time, Words, Questions, and Quality of Text Alternatives

Without providing guidance or a procedure for writing text alternatives, most of the participants spent an average of 2.9 minutes and wrote an average of 23.2 words. This was an average of 8.0 words per minute and 1.2 questions per minute. This acts as a baseline for comparing the efficiency and effectiveness of the procedure to identify information in the remaining research studies. The comparison was done in Section 6.5.1

The participants in this research study identified a low percentage of the information that was considered as Essential, Significant, and Helpful when they were not given a procedure. Given that the participants identified only 14.6% of the information that was rated as important to users, it is understandable that the amount of Essential and Significant information identified was limited.

Table 4.2 showed that Perceptual and Physical Object information were identified most often by the participants (10.9% and 24.8% respectively). However, these were also the types of information most often missed that were rated as Essential, Significant, and Helpful. Even for types of information that people often identified, a lot of information was still missed. This shows that improvements can be made to identify all types of information.

Ideally, most of the Essential and Significant information would be identified, while limiting the amount of Not Important information. The participants in this research study succeeded in minimizing the amount of Not Important information included in text alternatives. Even though they did not communicate all of the same information as the image, they only communicated information that users considered important.

Since the participants seldom identified the Significant and Helpful types of information, and the average percentages of Essential and Significant information identified were low, there is clearly room for improvement that supports the development of a procedure to help in this task.

4.5 Conclusion

This research study showed that some possible reasons for using the empty text instead of creating text alternatives include the difficulty and time required to create text alternatives. Since people could not easily (and quickly) determine the information to include in text alternatives, the empty text was used. Helping people determine the information being communicated by an image could help decrease the usage of the empty text.

Left to their own devices, the text alternatives created by the participants missed a large amount of information that was considered as important by users, especially information that was considered as Significant or Helpful. The procedure described in Chapter 3 can help people identify information that may be important and useful to users.

Chapter 5 Research Study of the Procedure in Document Format

The purpose of this research study was to determine the effectiveness of the procedure developed by the researcher when presented as in an ISO document format. Since the document would be read and interpreted by different types of users, it was important to evaluate how they would interpret and apply the procedure. This chapter discusses the research questions, methodology, and results of this research study.

The results of this research study were presented and published at the following meetings and conferences:

- International Standards Organization's (ISO) User Interface group meetings, in Venice Italy 2010 [54] and in Belleview, WA, USA in 2011 [55],
- International Technology & Persons with Disabilities Conference (CSUN) 2011 [56],
- HCI International conference 2011 [58], and
- The Annual Meeting of the Human Factors and Ergonomics Society (HFES) 2011 [57].

5.1 Research Questions

It is important to validate the procedure to make sure that it is usable and understandable by not only the technical experts, but also the general public. The objectives of this research study were to answer the following research questions:

General

1. What, if any, are the issues with the current document reported by the participants and potential improvements to be made?

Types of information and specific question related to the types of information

2. What types of information do people include after reading the document?
3. How many questions do the people answer after reading the document?

4. How do the user groups compare with each other with regards to the number of questions answered?
5. How do the images compare with each other with regards to the number of questions answered?

Amount of time spent

6. How much time is spent on identifying information about an image using the procedure?
7. How do the user groups compare with each other with regards to the amount of time spent?
8. How do the images compare with each other with regards to the amount of time spent?

Quantity and quality of information

9. What is the quantity of information that people identify for images by using the document?
10. How do the user groups compare with each other with regards to the quantity of information?
11. How do the images compare with each other with regards to the quantity of information?
12. What is the quality of the information that people identify for images by using the document?

Across Research Studies

13. How do the results of this research study compare with the results with the procedure was not used (in terms of the time spent, quantity of information, number of questions answered, and types of information identified)?

5.2 Methodology

5.2.1 Participants

Specific user groups of participants were sought out for this research study to ensure the validity and completeness of the results. These user groups were selected because they possess different levels of technical expertise and may describe images differently. Four user groups were identified:

1. Usability and Accessibility Specialists (henceforth referred to as “Specialists”): They are the leaders in this research area and their input is of high importance. Their knowledge and input can help improve on what already exists. They include persons from the SC35 international standards group, other ISO experts involved with usability and accessibility, persons from the Royal National Institute of Blind People (RNIB), persons from the World Wide Web Consortium (W3C), and assistive technology developers.
2. Web Site Developers (henceforth referred to as “Developers”): Their expertise is in the technical aspects of Web page programming and development. Although the content for the Web page should be provided by the Content Providers, Developers are often required to also develop the content. Therefore, Developers both program the Web pages and develop the content for the Web pages. They will interact with other types of users, such as content providers, when such users are involved.
3. Content Providers / Specialists (henceforth referred to as “Content Providers”): They provide the content for pre-programmed Web pages. Typically, they are the people who are knowledgeable in the field that the image is of since they are the person who chose the image. Since they are the experts in the image content itself, they are an important user group to include. They may be the people commissioning the Web page to be developed (hence, the clients who decides and approves of the final product) or specialists working for those commissioning the Web page whose main goal is content management. Within this research study, the Content Providers provided responses for images that were selected by some other content provider / specialist.
4. Average Internet Users (henceforth referred to as “Internet Users”): They are not involved with creating Web pages. They are the average person making use of the Internet on a daily basis. They do not have any technical background. This user group was included to ensure that anyone can make use of the procedure.

Due to the specific nature of the individual user groups, the participants in this research study were identified through personal contacts and approached personally to request their participation. The purpose of the study was not to generate formal statistical information to compare with an alternate approach but rather to validate the procedure. A random sample of participants was not considered necessary or feasible within a preliminary study. Given the small sample, the participants may not be true representatives of their population. Likewise, varying

demographics was not possible. However, the results can identify possible suggestions and recommendations that may be true of the population. Further research is needed to identify any dependencies on particular demographics.

Six participants in each user group were found. Each user group of participants were then separated into two groups (A and B), that is, three out of six participants from each user group were in Group A and three were in Group B. They were separated randomly. Each group applied the procedure to a different set of five images.

At the conclusion of the research study, five Specialists, six Developers, five Content Providers, and six Internet Users participated in the research study, for a total of twenty-two participants.

5.2.2 Materials and Execution

The research study was done without the presence of the researcher. It was recommended that the researcher be absent in order to not influence and bias how the participants interpreted and applied the procedure for producing text alternatives. This was done to avoid contaminating the answers of the participant. The usability of the materials provided to the participant was solely evaluated based on the participants' responses and feedback.

Each participant was provided with the following materials:

- A Word document of the sections from ISO/IEC 20071-11 containing the procedure for creating text alternatives (as outlined in Chapter 3 and provided in Appendix D) where the examples of the procedure were removed.
    - The Word document of the procedure provided to the participants was deliberately devoid of examples so as to not influence the participants' interpretation of how to apply the procedure. This would determine the understandability of the procedure document and help identify issues that affect the application of the procedure.
- A Word document for the participant to describe the images and provide feedback, which also included the instructions to the research study. This document was returned to the researcher. A sample of this document can be found in Section B.2 in Appendix B.

The research study consisted of three parts. First, the participants familiarized themselves with the Word document of the procedure and types of information they could consider when describing an image. Second, the participants were given a set of five images (along with links to

the original Web pages) and asked to describe each image by applying the procedure. The participants were asked to place their responses in a tabular form, identifying the component the information was about, the actual information about or communicated by the image, the question(s) leading to the information, and the importance of the information. This is illustrated by Table 5.1. Finally, the participants were asked to provide feedback regarding the procedure and their experience, including any problems they had with applying the procedure, any difficulties they experienced, and any suggestions they may have for improvement.

Table 5.1 Tabular Form for Specifying Information Identified About an Image

Whole Image or component name	Information about the Image	Question(s) Answered	Importance Level
Aaa	Bbb	Ccc	

Because the participants were not under the researcher's supervision during the research study, the participants were given one week to complete the research study. Since it was known that the research study may take a long period of time, it was recommended that the participant take breaks and complete the research study over several days.

5.2.3 Method to Answering Research Questions

Research Question 1

To identify problems and improvements to be made to the procedure and procedure document (research question 1), the comments and feedback from the participants were analyzed for commonalities. Quite a few issues were raised by multiple participants, which are summarized in Section 5.3.1.

Research Question 2 to 5

To determine the types of information that the participants identified, the information was mapped to the set of questions that could arrive at the information. Since the set of questions was presented to the participants, the question(s) could be explicitly stated by the participant. However, in most cases, the participants did not explicitly identify the specific question that

would arrive at the identified information. If they specified the category or sub-category of questions that the information came from, the specific question within that category or sub-category was deduced based on the information provided. If they did not specify a category or sub-category, then the specific question was deduced based on the information that was provided across all categories (the same process as discussed in Section 4.2.3). In cases where the participant specified a particular question or category of questions, there were times when the participant classified the information under the wrong category or question. In this situation, the information was marked as wrongly classified and re-classified into the more appropriate category and/or specific question(s). The marked information was later analyzed with regards to question misinterpretation along with the results of Research Study of the Procedure as a Prototype Tool in Section 6.3.2.

After mapping the information to questions, the results were analyzed to determine the types of information that were considered after reading the procedure and the various types of information that an image may contain (research question 2). The number of questions that were answered (research question 3) was calculated, along with the average, minimum, and maximum. The averages were also calculated for each user group within Group A and Group B participants for each image. These values were then used to compare between user groups (research question 4) and between images (research question 5).

Within each user group, the images were ordered from the most questions answered to the least questions answered with respect to the individual sets of images and overall. The ordering was then analyzed to determine if there was a consistent trend regarding the images that had more questions answered and regarding images that had fewer questions answered.

The results to these research questions are summarized in Section 5.3.2. These values were used for comparison with other research studies.

Research Questions 6 to 8

Similar to the number of questions answered, the amounts of time spent by participants was used to determine the average, minimum, and maximum values (research questions 6). Averages were also calculated for each user group within Group A and Group B participants and for each image. These values were then compared between user groups (research questions 7)

and between the images (research questions 8). The images that took the most and the least amounts of time were then identified by ordering the images within each user group in decreasing time in order to identify trends. The images were ordered with respect to each set of images and overall.

Once again, the amount of time spent on writing text alternatives for each image was reported by the participants themselves. It was possible that the times were not accurately reported and may be estimates of how long it actually took. However, the reported times were good enough for use to determine the major differences between the research studies.

The results to these research questions are summarized in Section 5.3.3 and used for comparison with the other research studies.

Research Questions 9 to 11

Similarly, number of words written by all participants was determined with the use of a word processor and the average, minimum, and maximum values were calculated (research question 9). Once again, the focus here was on the quantity of information rather than the quality of the text alternatives that were written. The intent of the procedure is to identify as much raw material as possible prior to formulating text alternatives. The number of words was one measurement of the raw information. There may be correlations between the number of words, the number of questions answered, and the types of information identified.

Averages were also calculated for each user group within Group A and Group B participants and for each image. These values were then compared with respect to user groups (research question 10), images (research question 11), and research studies. The images that contained the most and least words were identified by ordering the images within each user group in decreasing order with respect to each set of images and overall. These results were analyzed to identify trends.

The results to these research questions are summarized in Section 5.3.4. The results were used to compare with other research studies.

Research Question 12

The quality of the information identified and words written while using the procedure (research question 12) was calculated and analyzed by the same process as discussed in Section 4.2.3. The results are summarized in Section 5.3.5.

Research Question 13

The results of this research study were compared to the results of research question 6 from the Research Study of Text Alternatives Without the Procedure (research question 13). The results from the Research Study of Text Alternatives Without the Procedure were presented in Chapter 4 and the results from this research study are presented in Section 5.3. Both of these results are summarized and analyzed in Section 5.4.4. The analysis compares the amount of time, number of words, number of questions, types of information, and the quality of the information identified between the two research studies. It considers the types of information where the procedure was effective by comparing the frequency that the participants identified a certain piece of information. If participants identified a certain type of information more frequently with the procedure, then the procedure was helpful for that particular type of information.

5.3 Research Study Results

This section refers to the images given to the Group A participants as A1 to A5 (from Set A images) and those given to Group B participants as B1 to B5 (from Set B images). The images used in the research study can be found in Appendix A.

5.3.1 Participant Feedback

This section summarizes the feedback provided by the participants regarding the application of the procedure, the structure of the document, and the document content.

Applying the Procedure

Many participants (31.8%) indicated that they were uncertain about how the procedure could be applied in order to describe images. They stated that the document was logical and

made sense. However, they did not understand how to apply it. They requested that examples be provided for a better understanding of how to apply the procedure and to reassure them that they had applied the procedure properly.

One participant stated that there was a large disconnection between the document provided to them and how to apply the procedure as described. After reading such a long and technically detailed document, they had difficulty understanding what they were supposed to do with all of the information. Amidst the confusion, they also lost the purpose and intent of the procedure, even though they knew that it was important to provide useful text alternatives. Again, they suggested that the steps of the procedure be clearly stated and to providing examples of how the procedure could be applied in different contexts and for different types of images.

Document Structure

Some participants (18.1%) felt that the document could be rearranged to support the application of the procedure rather than the understanding of the theory. Again, although the information was understandable while reading about the procedure, the structure of the document made it difficult to follow while attempting to apply the procedure. One participant suggested a step-by-step guide on how to describe images.

Some of the participants (13.6%) expressed that an executive summary or an expansion to the current introduction was necessary to provide a broader sense of what the procedure is about. Some participants felt that after reading the long document, they lost the significance and importance of such a procedure. An executive summary or a more detailed introduction could help readers better understand the purpose and importance of the document.

Contexts

Some participants commented that images in different contexts would result in different information being extracted or used in text alternatives. Although this point was discussed in the document, the participants did not seem to be aware of it.

Because the participants were not the actual developer of the particular Web page nor an expert in the domain of the images' content, some participants (13.6%) felt that they were not

appropriate or suited for describing the images for that given context. Also, they were not certain of the intended context of the image. Therefore, they felt they could not provide information about the image.

Terminology

A few of the participants were confused as to whether or not it was necessary to use the terminology specified in the document when applying the procedure. Because the document was highly technical in nature, the participants may have felt their responses should be similarly technical. Some participants attempted to answer questions using the same terminology, but their responses were short and often non-descriptive as a result.

Some participants also felt that much of the terminology in the glossary were not necessary for the sake of describing the images and should be removed. They felt that this would eliminate some of the confusion and shorten the document length.

Length of Document

Most of the participants indicated that the document was too long for a research study. Even though the document made sense and was logical, they had forgotten what was said by the time they were to apply the procedure. Some participants have requested that the document be shortened and to remove the technical details. One participant stated that if they were not participating in this research study, they would not have completed reading the document.

Due to the length of the document and the amount of time it took to describe the images, some participants indicated the necessity of a tool to speed up the process. They also felt that a tool would clarify how the procedure was meant to be applied.

Another Approach

One participant suggested that the procedure be changed and instead, ask the person to describe the image in a certain number of sentences. For example, describe the image in one

sentence, then in three sentences, then in five sentences, and so on. The intent would be to provide more information each time.

Another participant provided results based on their own personal procedure. The procedure this participant followed was:

1. Decide on the importance of the image within the document.
2. Collect the general information (as described in Section 3.2.1).
3. Collect the specific information (as described from Sections 3.2.2 to 3.2.8).
4. Decide on the information to put into the primary alternative text and the information to put into the secondary alternative text.
5. Write both primary and secondary alternative texts.

The participant believed this procedure was clearer than what was described in the procedure document provided through the research study.

Levels of Importance

Several participants (22.7%) indicated that the levels of importance (described in Section 3.1.3) were subjective and wanted the levels to be more objective. They would like more concrete methods of determining importance and if certain information should be provided in text alternatives. They wanted to be certain that they were applying the procedure properly and that their rating was correct or accurate.

General Application

One Specialist expressed that the procedure and guidance could be applied for all types of media and electronic documents, not only HTML. They felt that the document was primarily focused on HTML and did not consider other types of documents. They thought the document should be phrased more generically to support a wider range of electronic documents and platforms.

5.3.2 Number of Questions Answered and Types of Information Identified

It was determined that the participants answered a total of 802 questions. Table 5.2 illustrates the frequency that each type of information was identified by the participants. The

participants of this research study mainly identified the object shown in the image, Perceptual information, and Subjective information, all of which are part of the What category of questions. Information regarding How, Location/Place, When, and Logical Relationships were more frequently missed.

Table 5.2 Usage of Categories and Sub-categories of Information with the Use of the Procedure in Document Format

Category or Sub-category	Percentage
How	1.6%
What – Logical Relationships	2.0%
Where – Location/Place	2.2%
When	3.4%
What – Textual	3.7%
How Much	4.4%
When – Sequential Relationships	5.6%
Why	7.4%
What – Classification	7.4%
Where – Spatial Relationships	7.6%
Who	8.0%
What – Subjective	11.8%
What – Perceptual	17.3%
What – Physical Object	17.6%

The participants answered an average of 7.5 questions, with a standard deviation of 1.1 questions, a minimum of 1 question, and a maximum of 44 questions while applying the procedure based on the document. Figure 5.3 shows a distribution of the results. The actual values can be found in Table C.3 in Appendix C.

Figure 5.4 shows the average number of questions each user group answered while applying the procedure to each image. The exact values can be found in Table C.5 in Appendix C. Table 5.5 shows the average number of questions that each user group answered for all images overall, and for Set A images and Set B images. Based on these results, it can be seen that some user groups answered more questions than others. For Set A images (the left side of Figure 5.4), it appeared the Developers answered the most questions while the Specialists answered the least questions. For Set B images (the right side of Figure 5.4), it appeared the Specialists answered more questions while the Developers answered the least questions.

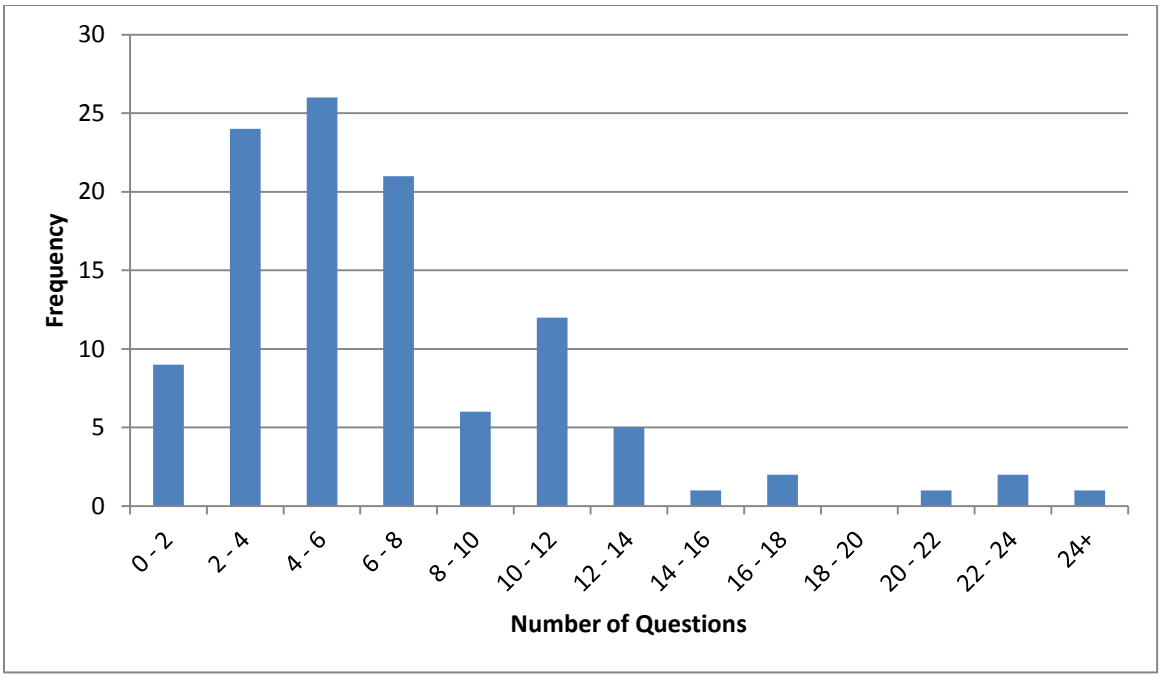


Figure 5.3 Number of Questions Answered with the Use of the Procedure in Document Format

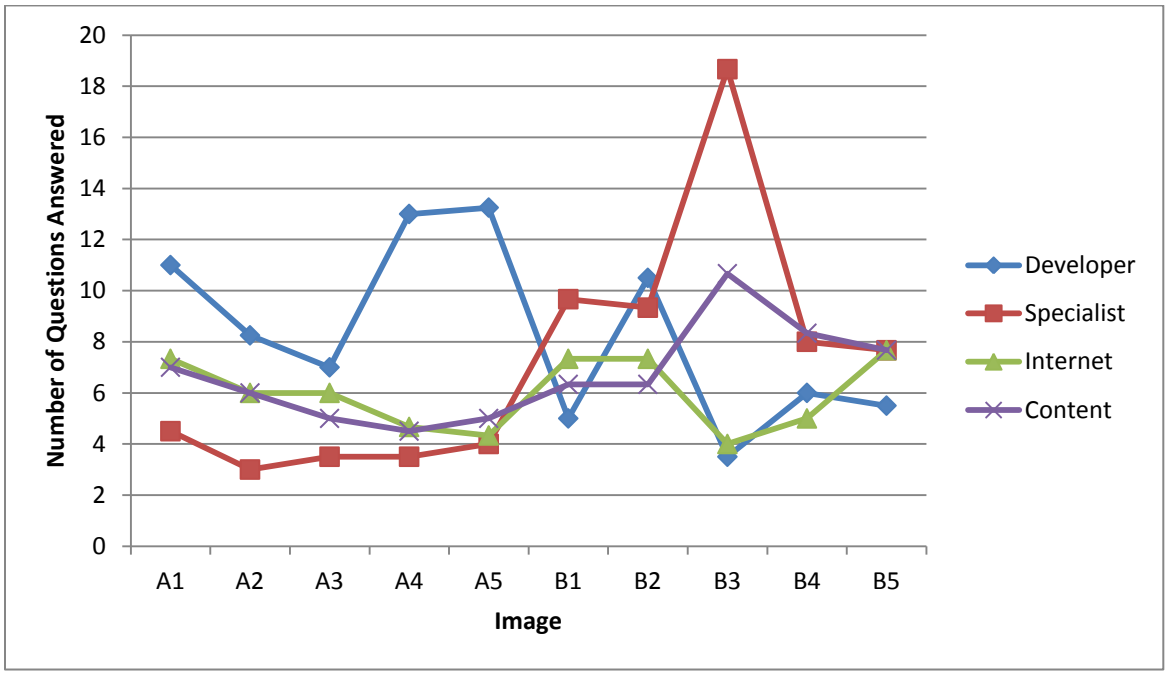


Figure 5.4 Average Number of Questions Answered for Each Image by Each User Group with the Use of the Procedure in Document Format

Table 5.5 Average Number of Questions Answered by Each User Group with the Use of the Procedure in Document Format

	All	Specialists	Developers	Content Providers	Internet Users
Overall	6.4	6.8	7.0	5.1	5.3
Set A	5.7	3.2	8.5	4.2	4.5
Set B	7.1	10.5	5.5	6	6.1

Table 5.6 illustrates the images in decreasing order of questions answered by each user group and overall. It can be shown that more questions were answered for certain images and fewer questions were answered for others. For example, A1 fairly consistently had more questions answered by the different user groups (showing a spike in Figure 5.4) while A3 fairly consistently had fewer questions answered (showing a drop in Figure 5.4). For Set B images, there is no consistency between user groups as to which images had more questions or fewer questions answered.

Table 5.6 Number of Questions Answered Per Image in Decreasing Order with the Use of the Procedure in Document Format

	Most									**Least**
Overall	B3	B2	A1	A5	A4	B1	B5	B4	A12	A3
Specialists	B3	B1	B2	B4	B5	A1	A5	A3	A4	A2
Developers	A5	A4	A1	B2	A2	A3	B4	B5	B1	B3
Content Providers	B3	B4	B5	A1	B1	B2	A2	A3	A5	A4
Internet Users	B5	A1	B1	B2	A2	A3	B4	A4	A5	B3

5.3.3 Amount of Time Spent

The participants were once again asked to record the amount of time it took for them to apply the procedure based on the document they read. The participants took an average of 15.9 minutes, a minimum of 2 minutes, and a maximum of 60 minutes to complete the procedure based on the document. Figure 5.7 shows a distribution of the results. The actual values can be found in Table C.11 in Appendix C. While a standard deviation is inappropriate since the data is not normal, 90% of the data is in the range of 2 to 36 minutes.

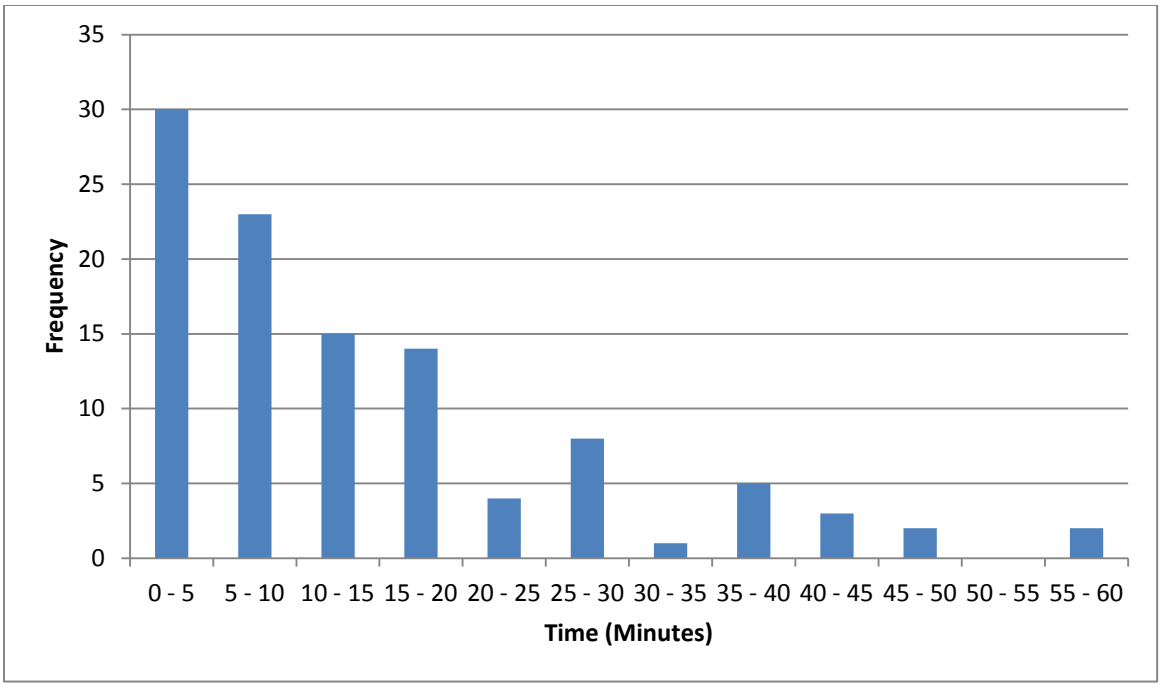


Figure 5.7 Amount of Time Spent with the Use of the Procedure in Document Format

Figure 5.8 shows the average amount of time each user group spent on applying the procedure to each image. The exact values can be found in Table C.13 in Appendix C. Table 5.9 shows the average times that each user group spent on all images overall, and for the individual sets of images. Based on these results, it can be seen that some user groups spent more time than the others. For Set A images (the left side of Figure 5.8), the Developers spent the most amount of time while the Specialists spent the least amount of time. The reverse was true for Set B images (the right side of Figure 5.8), that is, the Specialists spent the most amount of time while the Developers spent the least amount of time.

Table 5.10 illustrates the images in decreasing order of time spent by each user group and overall. It can also be seen that some types of images took relatively shorter or longer to apply the procedure than others for most or all user groups. For example, A2, B4, and B5 took less time to complete by each user group while each user group took longer to apply the procedure for A1, B1, B2, and B3. Overall, images in Set A (average 13.9 minutes) took less time to complete than images in Set B (average 17.8 minutes).

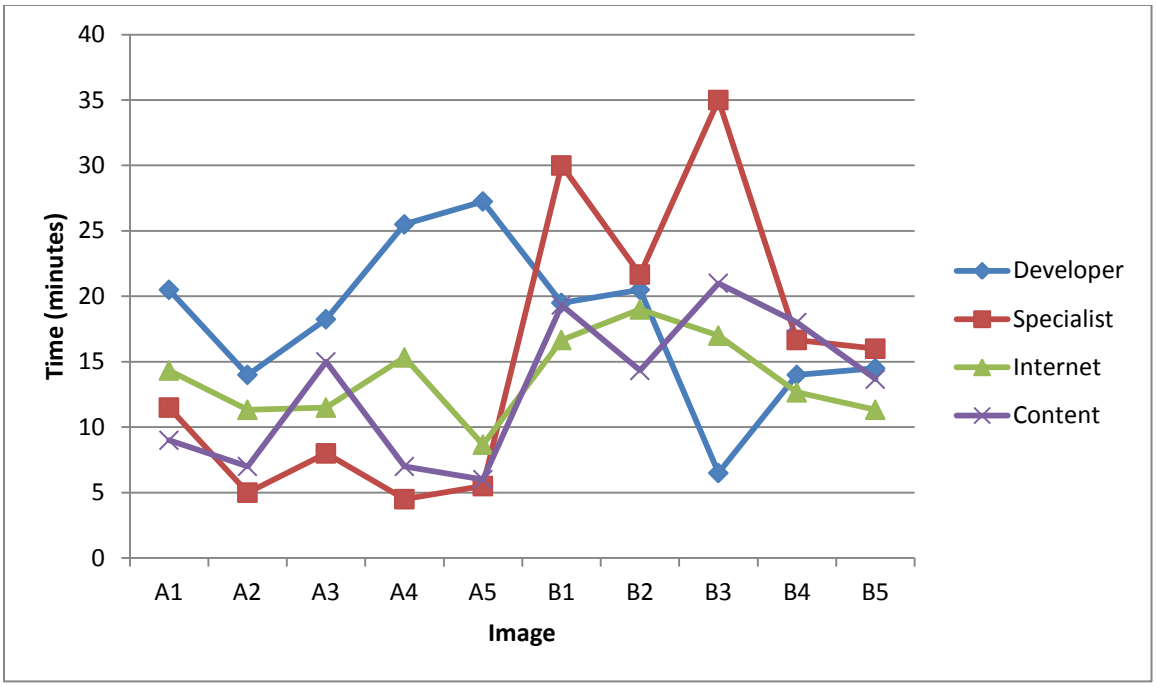


Figure 5.8 Average Time Spent on Each Image by Each User Group with the Use of the Procedure in Document Format

Table 5.9 Average Time Spent (in Minutes) by Each User Group with the Use of the Procedure in Document Format

	All	Specialists	Developers	Content Providers	Internet Users
Overall	15.9	15.4	18.1	13.0	13.8
Set A	13.9	6.9	21.1	8.8	12.2
Set B	17.8	23.9	15.0	17.3	15.3

Table 5.10 Time Spent Per Image in Decreasing Order with the Use of the Procedure in Document Format

	Most									Least
Overall	B1	B3	B2	A4	B4	A1	A5	A3	B5	A2
Specialists	B3	B1	B2	B4	B5	A1	A3	A5	A2	A4
Developers	A5	A4	A1	B2	B1	A3	B5	A2	B4	B3
Content Providers	B3	B1	B4	A3	B2	B5	A1	A2	A4	A5
Internet Users	B2	B3	B1	A4	A1	B4	A3	A2	B5	A5

5.3.4 Number of Words Written

A word processor program was used once again to count the number of words written by a participant for a given image. Based on these results, the participants wrote an average of 73.4 words, a minimum of 5 words, and a maximum of 323 words for a given image. Figure 5.11 shows a distribution of the results. The actual values can be found in Table C.19 in Appendix C. While a standard deviation is not appropriate due to the data being not normal, 90% of the data is within the range of 5 to 170 words.

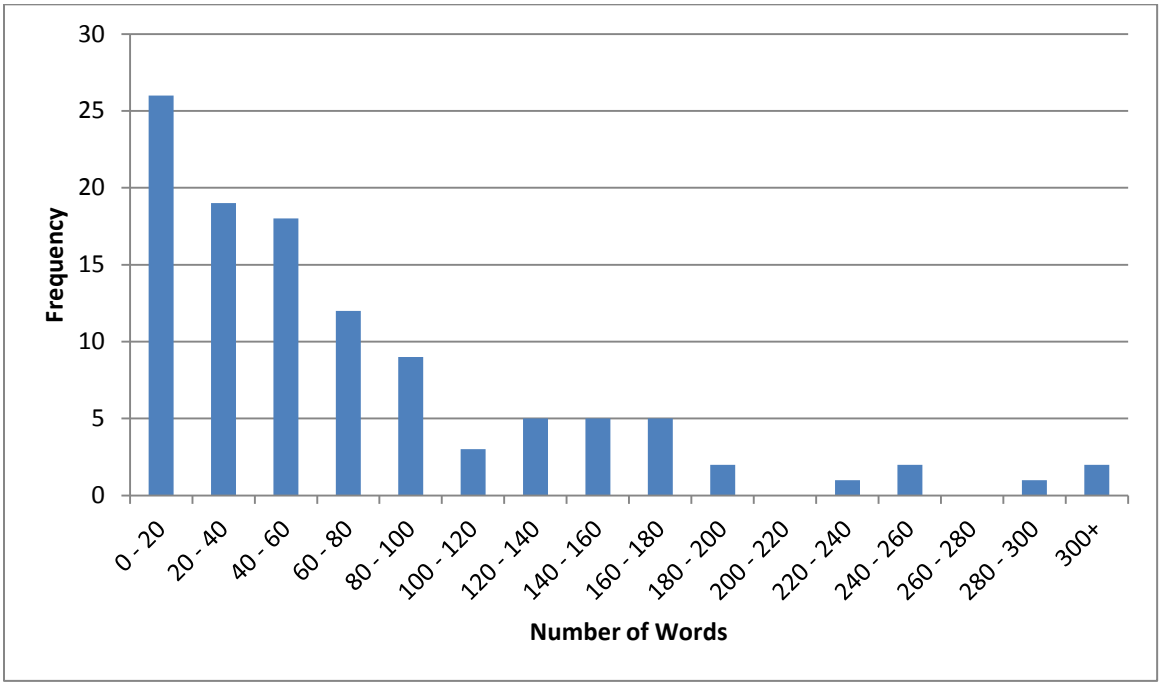


Figure 5.11 Number of Words Written with the Use of the Procedure in Document Format

Figure 5.12 shows the average number of words that was written by each user group for a given image. The exact values can be found in Table C.21 in Appendix C. Table 5.13 shows the average number of words that each user group wrote on all images overall, and for Set A images and Set B images. Based on these results, it can be seen that overall, more words were generated by Group B participants for Set B images (right side of Figure 5.12) than Group A participants for Set A images (left side of Figure 5.12). For both sets of images, Developers wrote the most number of words. For Set A images, the Specialists wrote the least number of words. For Set B images, the Internet Users wrote the least number of words.

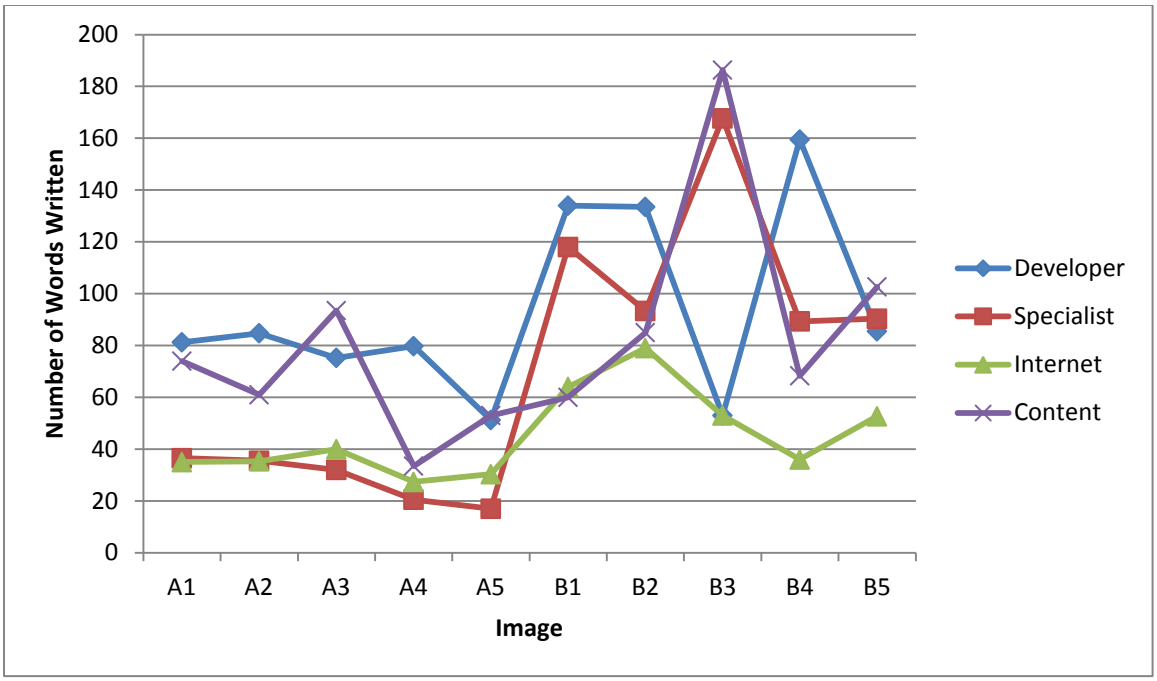


Figure 5.12 Average Number of Words Written for Each Image by Each User Group with the Use of the Procedure in Document Format

Table 5.13 Average Number of Words Written by Each User Group

	All	Specialists	Developers	Content Providers	Internet Users
Overall	73.4	70.0	93.8	81.7	45.3
Set A	52.8	28.3	74.5	63.0	33.6
Set B	94.0	111.7	113.1	100.5	56.9

Table 5.14 illustrates the images in decreasing order of words written by each user group and overall. It can also be seen that more words were written by all user groups about certain image types than others. For example, all user groups wrote more words about A1, A2, A3, B1, and B2 and fewer words about A4, A5, and B4. Overall, more words were written for images in Set B (average 94.0 words) than Set A (average 52.8 words).

Table 5.14 Words Written Per Image in Decreasing Order with the Use of the Procedure in Document Format

	Most									Least
Overall	B3	B2	B1	B5	B4	A3	A1	A2	A4	A5
Specialists	B3	B1	B2	B5	B4	A1	A2	A3	A4	A5
Developers	B4	B1	B2	B5	A2	A1	A4	A3	B3	A5
Content Providers	B3	B5	A3	B2	A1	B4	A2	B1	A5	A4
Internet Users	B2	B1	B3	B5	A3	B4	A2	A1	A5	A4

5.3.5 Quality of Identified Information

When the participants were provided with a procedure, they identified an average of 22.6% of the information considered as important to some extent based on the results of the Research Study of the Identified Information. This included an average of 35.4% of the Essential information, 24.8% of the Significant information, and 14.1% of the Helpful information. Participants also identified an average of 11.0% of the information that was later considered as Not Important. Table 5.15 provides more detailed results. These results were further broken down by user group. The average percentages of each user group identifying the four different levels of importance are illustrated in Figure 5.16. The values can be found in Table C.26 of Appendix C.

Table 5.15 Summary of Quality of Information Identified with the Procedure in Document Format

	Essential	Significant	Helpful	Not Important
Average (percent)	35.4%	24.8%	14.1%	11.0%
Average (identified/total)	33.1 / 94	20.8 / 84	23.5 / 167	5.5 / 50
Maximum (percent)	57.0%	54.2%	35.7%	62.6%
Minimum (percent)	12.5%	0.0%	0.0%	0.0%

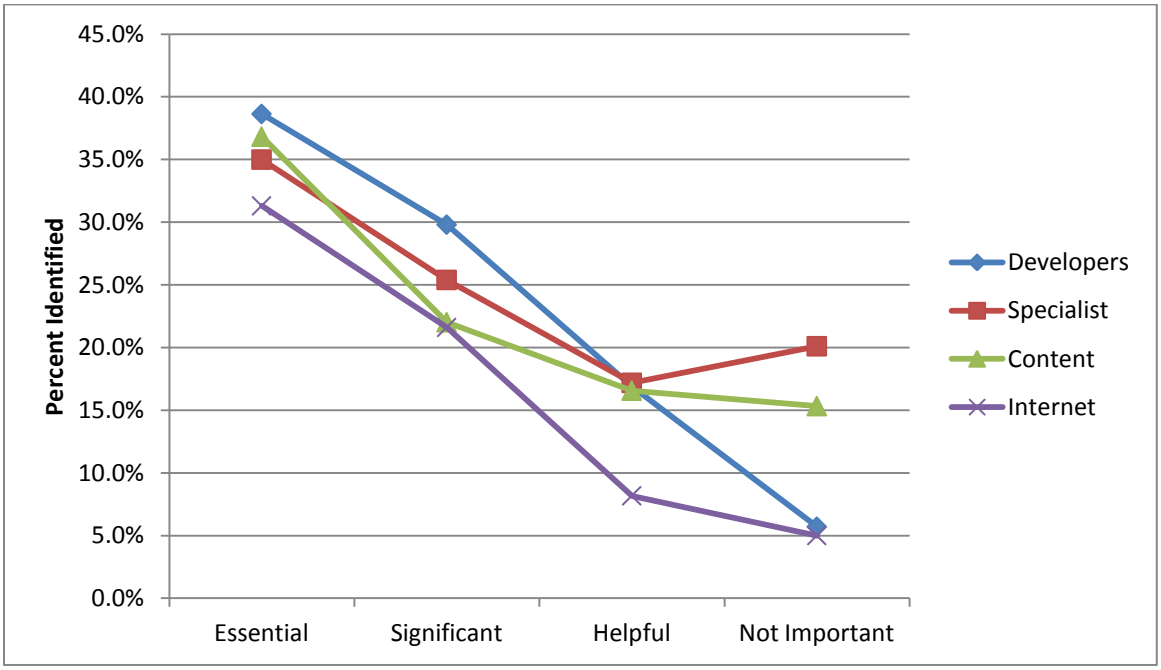


Figure 5.16 Average Percentage of Information Identified by Each User Group with the Use of the Procedure in Document Format

Table 5.17 lists the top five types of information that the participants often missed within each level of importance when they were provided with a procedure. Table 5.18 lists the top five types of information that the participants often identified within each level of importance when they were provided with a procedure.

Table 5.17 Types of Information Often Missed with the Procedure in Document Format by Importance Level

Essential	Significant	Helpful
1. What – Perceptual 2. What – Physical Objects 3. What – Subjective 4. When – Sequential Relationship 5. What – Textual	1. What – Physical Objects 2. How Much 3. What – Perceptual 4. What – Subjective 5. What – Logical Relationships	1. What – Perceptual 2. Where – Spatial Relationships 3. What – Subjective 4. When – Sequential Relationships 5. What – Physical Objects

Table 5.18 Types of Information Often Included with the Procedure in Document Format by Importance Level

Essential	Significant	Helpful
1. What – Classification 2. What – Physical Objects 3. Why 4. What – Subjective 5. Who	1. What – Perceptual 2. What – Textual 3. What – Physical Objects 4. Who 5. When – General	1. Who 2. What – Perceptual 3. Where – Spatial Relationships

Compared against the results of Research Study of the Identified Information, there were 216 instances where a greater percentage of participants identified a piece of information after they read the procedure. Of these 216 instances, 50 were Essential, 49 were Significant, 93 were Helpful, and 24 were Not Important. Table 5.19 breaks this down by types of information. There were also 55 instances where a lower percentage of participants identified a piece of information after they read the procedure. Of these 55 instances, 20 were Essential, 13 were Significant, 20 were Helpful, and 2 were Not Important. Table 5.20 shows the breakdown of these instances by types of information. This is further broken down by question in Table C.39 and Table C.40 in Appendix C.

Table 5.19 Instances When Greater Percentage of Participants Identified Information Using a Procedure in Document Format by Importance Level

	Essential	Significant	Helpful	Not Important
Why	4	0	1	2
What – Textual	6	2	3	1
What – Classification	4	1	0	0
What – Physical Objects	7	11	7	2
What – Perceptual	12	5	27	5
What – Subjective	5	7	21	0
What – Logical Relationships	0	2	0	0
Who	1	2	7	0
Where – Location/Place	1	0	0	0
Where – Spatial Relationships	2	4	15	14
When – General	1	1	1	0
When – Sequential Relationships	7	1	8	0
How Much	0	11	2	0
How	0	2	1	0
Total	50	49	93	24

Table 5.20 Instances When Lower Percentage of Participants Identified Information Using a Procedure in Document Format by Importance Level

	Essential	Significant	Helpful	Not Important
Why	3	0	1	0
What – Textual	0	0	0	0
What – Classification	5	0	0	0
What – Physical Objects	5	3	3	0
What – Perceptual	0	5	8	1
What – Subjective	2	1	3	0
What – Logical Relationships	0	0	0	0
Who	3	1	1	0
Where – Location/Place	1	0	1	0
Where – Spatial Relationships	0	0	3	1
When – General	0	2	0	0
When – Sequential Relationships	0	0	0	0
How Much	1	1	0	0
How	0	0	0	0
Total	20	13	20	2

5.4 Analysis of Research Study Results

This section analyzes the results and identifies possible improvements to the procedure and guidance. These improvements will be denoted using "IMP" followed by a number. For example, "IMP 1" and "IMP 2". Most of these improvements have already been implemented between research studies and during the evolution of ISO 20071-11.

5.4.1 Analysis of Feedback

The participant feedback identified several modifications to be made to the ISO document and problems to be resolved.

Executive Summary / Introduction

Some of the participants expressed that the current introduction in the document did not provide a clear understanding of what the document was about. By the time they finished reading the entire document, they forgot the purpose of the document. They felt that a more comprehensive introduction or executive summary would better prepare them for interpreting the information. Therefore, the document should give a general purpose or sense of the procedure.

IMP 1. Add content to the executive summary and introduction of ISO 20071-11 to give a general purpose or sense of the procedure.

Document Restructuring and Clearer Procedure Steps

Based on the participant comments and analysis of the time spent on each image, it was concluded that the procedure needed to be stated more clearly. Although the steps to describe an image were stated within the document, it was missed and the participants were confused as to what to do. Therefore, the procedure should be given a more prominent focus within the document. This might require restructuring the document. Originally, a summary of the steps to the procedure was placed at the end of the document and the participants may have gotten lost or confused by the time they reached the end of the document. It was better to provide the summary first, followed by the details of each step.

Based on the participant comments, it became clear that the current document structure did not support the participants in applying the procedure. The document provided to the participants was structured to support the guidance of generating good text alternatives but not the application of it. Therefore, the document should be restructured such that it supported the application of the procedure. This could improve comprehensibility, minimize the learning curve, and add emphasis on the procedure steps.

IMP 2. Rearrange ISO 20071-11 such that the procedure summary is provided prior to the details of each step of the procedure.

IMP 3. Structure and change the contents of ISO 20071-11 to support applying the procedure.

IMP 4. In ISO 20071-11, add emphasis on the procedure steps to clearly identify them.

Examples

The participant comments emphasized the importance of examples on how the procedure should be applied and how image descriptions can be generated. Examples can reassure users that they are applying the procedure correctly.

Also, some of the participants commented that they could not see this procedure and guidance being applicable to all types of images. They can easily see that it applied to charts and graphs, but had difficulty seeing it applied to other types of images.

One of the goals of this research study was to determine how people would interpret the document based on how it was written. Therefore, while full examples were included in the working draft being developed in SC35, they were purposely removed from the document that was evaluated in this study so as to not bias the research. The examples in the document should be expanded to include different types of images. However, the intent of the procedure and guidance is to be generic enough to be applicable even to image types that might not have been previously considered. The guidance must then also make note that the examples are to illustrate the procedure steps and the resulting text alternatives are not meant to be used a definitive text alternatives for every image of a particular type.

IMP 5. In ISO 20071-11, add examples that illustrate each step of the procedure.

IMP 6. In ISO 20071-11, provide additional examples of different types of images.

IMP 7. In ISO 20071-11, add a note that the examples are not meant to be definitive and the resulting example text alternatives are not meant to be definitive for every image of a particular type.

Context Specific

The participants shared the researcher's awareness that the context of the image would influence the text alternatives generated. The document had commented on and expressed the concept of the image`s context and its effects on the text alternatives. The document further expressed the idea that the purpose of the image depends on and may change with the context of use. Therefore, identifying the purpose requires considering the context of use. Despite the discussion regarding this in the document, the participant feedback showed that it was not expressed clearly enough. Due to the length of the document, this point may have been forgotten or lost along the way. The document should further emphasize how the context of the image affected the text alternatives that were generated and provide examples to illustrate this point.

IMP 8. Where applicable throughout ISO 20071-11, add reminders about the effects of the context and purpose of the image on the text alternatives and the importance of a piece of information.

Expertise Knowledge

Expert or domain knowledge of the image content, along with the ability to interpret the content, created a variance in the level of detail in the information identified in some instances, such as the flow chart (B3) and assembly diagram (A2). The effect of expert knowledge was discussed within the procedure document; however, the document lacked guidance on what to do when the user did not have the expert knowledge and expert knowledge was needed. The document should further discuss the importance of expert knowledge and what could be done when the user does not have expert knowledge about the image.

IMP 9. In ISO 20071-11, add further discussion regarding the importance of expert knowledge, including that it is not always needed.

IMP 10. In ISO 20071-11, add guidance on what could be done when expert knowledge is needed but not available.

Terminology

During the preparation of the initial draft of the ISO document, a large set of terminology was included in the document to ensure coverage of relevant definitions. However, not all of the terms were used in the ISO document's main content, resulting in a longer list of terminology than was necessary. Therefore, the ISO document should remove all of the terminology that was not used in the document in order to eliminate confusion between the terms and to shorten the document.

IMP 11. Remove terminology that is not used in the main content of ISO 20071-11.

Necessity of a Tool

The participants expressed that due to the length of the document and the difficulties they experienced in applying the procedure, a tool would be beneficial. It would remove the need to read the lengthy document, clarify how to apply the procedure, and shorten the time required to identify information about the image. To this end, a prototype tool was created and a second research study was necessary to evaluate the prototype tool and the type of information that would be identified as a result of using a tool.

Other Approaches

One participant suggested asking participants to write a set number of sentences about an image as the text alternative. Although this approach would greatly shorten the procedure document, it did not resolve the issue of helping people determine the information they should or

could provide in text alternatives. Also, it did not consider evaluating the text alternative with users.

The second method suggested by another participant used the same procedure as was described in the document provided to all of the participants. The participant may have suggested a similar approach because the document was unclear as to the steps of the procedure, as previously noted. Having a participant describe the same approach for identifying information and formulating text alternatives shows that they agree with the approach. The main issue is to modify the document so that people can use the procedure so that they will not need to recreate the procedure that was described.

Levels of Importance

Based on the participant comments, it was clear that the different levels of importance were confusing and the participants were uncomfortable about assigning an importance level when they did not choose the image themselves. They also did not understand the purpose of assigning the importance level. While some participants formulated text alternatives for the images, most participants did not and focused on identifying information about the images. Since it was the organization step that made use of the importance levels, it was understandable that they missed the significance of the importance levels.

The document and procedure should further emphasize the significance of the importance levels and why they were needed. The document should also provide additional guidance and examples of when each level should be applied. The guidance should also emphasize that it is not possible to be objective about assigning the importance levels because it is subject to the context and purpose. The guidance in the procedure document is suggestions about when a piece of information may be assigned a certain importance level and is not definitive or prescriptive.

IMP 12. In ISO 20071-11, clarify the levels of importance and their purpose.

IMP 13. In ISO 20071-11, add guidance on when a piece of information may be assigned a certain importance level.

General Application of Guidance

The original intent of the procedure and guidance was to be platform and technology independent. As technology changes over time, it would not be beneficial to provide guidance specific to HTML. Due to the prominence of HTML in providing text alternatives, it was often used as the example.

The document should be modified to emphasize the support of other forms of documents. The document should use the terms "primary alternative text" and "secondary alternative text" instead of "`alt` attribute" and "`longdesc` attribute", which are specific to HTML.

IMP 14. Where applicable in ISO 20071-11, replace "`alt` attribute" and "`longdesc` attribute" with "primary alternative text" and "secondary alternative text" respectively.

5.4.2 Analysis of Time, Words, and Questions

Similar to cataloguing an image properly, time is required to write informative text alternatives. Eakins reported that it took seven to forty minutes to catalog an image. [15] Since this is similar to the range of time spent to write text alternatives (which was 15.9 minutes on average), it can be concluded that the procedure took a reasonable amount of time to apply.

Table 5.21 summarizes the images in the research study that user groups consistently spent relatively more time, wrote relatively more words, and answered relatively more questions on particular images. The table also lists the images that the user groups consistently spent relatively less time, wrote relatively fewer words, and answered relatively fewer questions.

Table 5.21 Summary of Time, Words, and Questions Output with respect to Images with the Use of the Procedure in Document Format

	More	**Less**
Time Spent	A1, B1, B2, B3	A2, B4, B5
Words Written	A1, A2, A3, B1, B2	A4, A5, B4
Questions Answered	A1	A3

Perhaps because they were the first images in their sets, the participants spent relatively more time on images A1 and B1. Participants may have taken more time to familiarize themselves with the procedure and to ensure that they were following the procedure correctly.

The reason the other images took relatively more time or less time than others may be due to the complexity of the individual image, the participants' ability to comprehend the document, or the participants' ability to interpret the images. Given that the images that required more time and the images that resulted in more words and questions differed when the tool was used (Table 6.26 on page 129), it was likely due to the participants' ability to interpret the image and/or apply the procedure that caused some of the variance in the outcomes.

Within each user group, it was generally the case that the images in Set B required more time than the images in Set A. Table 5.22 summarizes the overall average time spent, words written, and questions answered for images in Set A and Set B. The Set B images resulted in higher values than those in Set A. This might be due to either the complexity of the image or the individuals themselves. When compared with the results from the Research Study of the Procedure as a Prototype Tool, the difference in values between Set A and Set B appear to be more likely due to the participants rather than due to the images' complexity. In the Research Study of the Procedure as a Prototype Tool, the participants had switched image sets and the results showed that the Set A images resulted in higher values than those in Set B (Table 6.27). This is the inverse of what happened in this research study with the use of the document. The same participants generated higher values in both research studies. Therefore, it appeared that the difference was due to the participants and not the images. This showed that the individual and the amount of effort they put into creating text alternatives influence the outcome.

Table 5.22 Comparison Between Set A and Set B Images (Overall Averages) When Using the Procedure in Document Format

	Set A Images	**Set B Images**
Time Spent	12.3 minutes	17.9 minutes
Words Written	49.8 words	95.6 words
Questions Answered	6.3 questions	7.7 questions

Table 5.21 showed that for some images, there was a relationship between the amount of time spent and the number of words written. For example, A1, B1, and B2 all showed that the more time spent on those images, the more words were written. Similarly, B4 showed that the less time spent, the fewer words were written about that image. However, this trend was not consistent across all images. In many cases, there was no correlation between time spent and number of words. In fact, A2 indicated that when less time was spent, more words were written.

There also does not appear to be any correlation between the amount of time spent and the number of questions answered.

However, there were more consistent trends when individual user groups were analyzed with respect to the time spent, words written, and questions answered. The comparison can be found in Section C.5 in Appendix C. Within each user group, there was a correlation between time, words, and questions for most of the images. As user groups spent more time on an image, more words were written and more questions were answered about that image.

The results showed that the Developers wrote the most words and answered the most questions compared with the other user groups. While the Developers might not consistently provide responses for above the average number of questions for all images, they did consistently write more words than average. Since the Developers can produce above the average amount of output when using the procedure, it shows that the procedure can and does help developers to identify information about an image that may end up as part of text alternatives.

Internet Users wrote the least words and answered the least questions compared to the other user groups. It is likely due to the highly technical nature of the document that was provided to them. The document made use of terms and was written as a technical specification. It is likely that they had difficulties comprehending the contents such that they could apply the procedure. The feedback and comments provided by the participants support this (Section 5.3.1).

5.4.3 Analysis of Quality of Information Identified

This section discusses the amount of the information reported as Essential, Significant, Helpful, and Not Important (in the Research Study of the Identified Information) that the participants identified with the use of the procedure in document form. The participants identified an average of 36.2% of the Essential information, 24.8% of the Significant information, 14.4% of the Helpful information, and 11.0% of the Not Important information. These percentages are still considered relatively low.

Although the procedure exposed the participants to more types of information that they could consider for text alternatives, they were not required to consider all of the types of information when identifying information. It is possible that the participants did not consider every single question within the set of questions, resulting in much of the Essential, Significant, and Helpful information being missed.

Even though the participants did not consistently identify more than 50% of the Essential or Significant information, there were some participants who identified more than 50% of the important information (as indicated by the maximum percentages shown in Table 5.15 on page 79). This showed that it was possible to identify greater percentages of information within each importance level when the procedure was used, with above 50% of information identified in almost all of the importance levels.

Many of the issues and problems discussed in Section 5.4.1 may have also contributed to the low percentages of information identified. Even with a clear set of questions to consider, not knowing what to do with it resulted in a lack of information being identified. With the large amount of information to process about the procedure and the set of questions, many participants were confused and did not know what to do. They had suggested that examples be included in the document to clarify what needed to be done. The examples would show that each question should be considered every time. Then, by following the examples, more types of information would be considered more consistently. The few participants who understood the procedure and made use of the set of questions identified a greater percentage of the important information.

As noted in Section 3.1.2, breaking down the image into its parts (image components) allows the person to focus on identifying information particular to a certain image component. Most participants did not do this, possibly because the steps of the procedure were not clear or because the participant did not feel it was necessary. Therefore, a large amount of important information was missed.

The participants had provided their personal opinions of the importance of the information they provided. The participants rarely provided information they had considered as Not Important. It was only later that the information was discovered to be Not Important. This shows the difference in what the users and text alternative writers consider as important to them and should be communicated.

With respect to the user groups, Table 5.21 shows that the Developers were the only group to identify above average of Essential, Significant, and Helpful information. The Developers also identified below average of Not Important information, indicating that Developers were able to use the procedure and identify the important information while limiting the amount of Not Important information. Being that the Developers are the main user group expected to use the procedure, these results show great promise that the procedure works.

The Specialists and Content Developers identified around the average amount of important information for all of the importance levels, except they identified significantly more Not Important information. These results show that the procedure can also be used successfully by the other user groups to identify important information. There needs to be additional guidance to help people limit the amount of Not Important information being identified.

While the procedure helps the participants to consider different types of information, there needs to be a greater focus on identifying Essential and Significant information and less focus on the Not Important information.

5.4.4 Comparison with Research Study Without a Procedure

There was a significant increase in the results when compared to the Research Study of Text Alternatives Without the Procedure with respect to time, words, questions, and quality. Table 5.23 summarizes these results.

Table 5.23 Results Comparison Between Document Format and Without a Procedure

	Without Procedure	Document Format
Time Spent (minutes)	2.9	15.9
Words Written (words)	23.2	73.4
Questions Answered (questions)	3.4	7.5
Essential Information Identified (average percent / number)	31.8% / 29.8	35.4% / 33.1
Significant Information Identified (average percent / number)	14.9% / 12.5	24.8% / 20.8
Helpful Information Identified (average percent / number)	4.8% / 7.9	14.5% / 24.1
Not Important Information Identified (average percent / number)	2.2% / 1.1	11.0% / 5.5
Total Important Information Identified (percent)	14.6%	22.6%

While not a desirable outcome, it took 5.5 times as long to complete the procedure for a single image while using the procedure than without. Based on the user feedback, a considerable amount of this time can be attributed to interpreting procedure to ensure that they have done it correctly. More time was also needed to consider the various types of information that were not considered before or that they normally would not have considered if it was not for the set of questions. Modifications are needed in order to lower the amount of time spent using the

procedure to be closer and comparable to the amount of time spent when the procedure was not available. These modifications were discussed in Section 5.4.1. It is anticipated that once the modifications are made, the amount of time needed to apply the procedure will decrease while the output (number of words, number of questions, types of information, and quality of information) will remain the same or improve.

While the participants from both research studies identified a higher percentage of the same types of information (such as Physical Object, Perceptual, Who), the participants in this research study also considered other types of information after reading the document. The percentages of information from each category and sub-category of information were more wide spread, indicating that more details and different types of information were being identified. This shows that the procedure can help to identify information that users may not otherwise consider.

The total amount of information considered of at least some importance increased by one-and-a-half times with the use of the procedure, from 14.6% to 22.6%. Although the average percentage and number of pieces of Essential information identified did not show much improvement with the use of a procedure, the amount of Significant and Helpful information doubled or tripled. It is important that the procedure help identify not only the Essential information, but also the Significant and Helpful information. There were 216 instances where more participants identified the information with the procedure than without the procedure. Of the 216, more than half of them were considered as Significant or Helpful. This indicates that the procedure is useful, especially for identifying Helpful and Significant information.

That is not to say that the procedure was not useful in identifying the Essential information. Even though there was only an approximately 4% increase of Essential information identified on average, there were fifty instances where there was a greater percentage of participants who identified a piece of Essential information. That is, 53.2% of the Essential information was identified better when the procedure was used. This shows that the procedure helps the participants identify the Essential information as well as the other important information.

The procedure was especially useful at identifying Perceptual, Subjective, Physical Objects, Spatial Relationships, Sequential Relationships, Textual, and How Much information. There were some instances within these types of information where the percentage of participants who identified a piece of information was lower when the procedure was used.

The participants in both studies sometimes identified different pieces of information more frequently within the same type of information, sometimes even for the same question. The participants identified different pieces of information or details presented in an image. This is likely attributed to the individual participants and what they felt was important to identify.

According to Table C.39 in Appendix C, the procedure was especially helpful for identifying the numerical data or quantity of objects, Subjective information, the facial expression and position of a person, textual content in the image, and Sequential information. It also greatly helped to identify when the image took place, the purpose of the image, and Perceptual information (such as texture and colour).

There were fifty-five instances where the procedure seemed to be not as useful in identifying information, listed in Table C.40 in Appendix C. That is, a lower percentage of participants identified a piece of information when the procedure was used. These instances were due to the small and different sample sizes between the research studies. The sample size of the Research Study of Text Alternatives Without the Procedure was approximately one-half that of this research study. Therefore, even though the same or a greater number of participants identified a piece of information when using the procedure, the percentage was still lower due to the sample size.

There was one instance where it may not have been due to the sample size. It was for the question to identify specific type or brand of the object in the image or image component. It is possible that the participants did not know the type or brand of the object or they simply did not consider that information to be important. The question may need to be modified to include the type of object since it currently only considers the brand or model of an object.

IMP 15. In the Physical Objects sub-category, modify the question regarding the brand or model of an object to also include the type or species of object. Examples should also be provided.

When the results of each individual user group were compared to the results of the Research Study of Text Alternatives Without the Procedure (Table 4.8 on page 56 and Table C.27 on page 216), the comparison shows that there was an improvement for all user groups. The Internet Users identified the least amount of important information but they identified more than

the participants who did not have the procedure. This indicates that all user groups can benefit from using the procedure.

5.5 Conclusion

Due to the frustration of figuring out what to do and the large amounts of time spent on applying the procedure, it was understandable that a low percentage of information was identified. However, having read about the different types of information to consider, a greater amount of important information was identified than without using a procedure. With the procedure, a lower percentage of important information would be missed.

This research study helped validate the procedure for creating text alternatives and identified improvements to ISO 20071-11. Many of these improvements have already been made to the ISO document to improve the applicability of the procedure. This research study showed that the procedure was useful in identifying more information that is important to users.

Chapter 6 Research Study of the Procedure as a Prototype Tool

In the Research Study of the Procedure in Document Format, the participants stated that a tool was necessary for the procedure to be feasible for everyday use. A prototype was created to facilitate the current state of the procedure in order to re-evaluate the procedure in a tool format. This chapter discusses the research questions, participants and methodology, and the results of this research study.

The results of this research study were presented and published at the following meetings and conferences:

- International Standards Organization's (ISO) User Interface group meetings, in Venice Italy 2010 [54] and in Belleview, WA, USA in 2011 [55],
- International Technology & Persons with Disabilities Conference (CSUN) 2011 [56],
- HCI International conference 2011 [58], and
- The Annual Meeting of the Human Factors and Ergonomics Society (HFES) 2011 [57].

6.1 Research Questions

Many of the objectives of this research study were similar to that of the previous research study. The difference was that the procedure was presented in a different format. The objectives of this research study were:

General

1. What, if any, are the issues with the prototype tool and potential improvements to be made?

Misinterpretation of questions (new research topic)

2. How often are the questions misinterpreted from their intended use?
3. What questions are more often misinterpreted?

Types of information and specific questions related to the types of information

4. What types of information do people include when they are required to consider every question?
5. How do the user groups compare with each other with regards to the number of questions answered?
6. How do the images compare with each other with regards to the number of questions answered?

Amount of time spent

7. How much time is spent on identifying information about an image using the prototype tool?
8. How do the user groups compare with each other with regards to the amount of time spent?
9. How do the images compare with each other with regards to the amount of time spent?

Quantity and quality of information

10. What is the quantity of information that people identify for images when using the prototype tool?
11. How do the user groups compare with each other with regards to the quantity of information?
12. How do the images compare with each other with regards to the quantity of information?
13. What is the quality of the information that people identify for images when they are forced to consider every question?

Across Research Studies

14. How do the results of this research study compare with the other research studies (in terms of the time spent, quantity of information, and number of questions)?

It was necessary to determine how the results differed between applying the procedure based on the document and with the use of a tool. At the time of this research study, the tool's intention was to ensure completeness in coverage for types of images and information.

Therefore, it was possible that the amount of time spent to describe image might not improve compared to the time spent while using the document.

6.2 Methodology

The methodology for this research study on the prototype tool was similar to the research study on the document format (Section 5.2). The participants from the previous research study were asked to also evaluate the prototype tool. As such, the participants were more familiar with the procedure and the set of questions since this was the second time they made use of the procedure. Due to the timing of the research study, many participants declined to participate. Therefore, at the end of the research study, there were five Usability and Accessibility Specialists, six Web Site Developers, two Content Providers, and four Average Internet Users who participated, for a total of seventeen participants who evaluated the prototype tool. Similar to the previous research studies, the participants may not be true representatives of their population and varying demographics was not possible due to the small sample size. However, the results can identify possible suggestions and recommendations that may be true of the population.

Although it was possible to locate other potential participants to replace the participants who declined, one goal of the research study was to compare the results of the document format with the prototype tool format. Therefore, new participants were not found at that time.

Once again, the research study was done without the presence of the researcher to avoid contaminating the responses. Each participant was provided with the following materials:

- A link to the prototype tool used for describing the images (described in Section 6.2.1), and
- A link to a Web page explaining how to use the prototype tool. This Web page can be accessed from every screen within the prototype tool. This document can be found in Section B.4 in Appendix B.

This research study consisted of three parts. First, the participant familiarized themselves with the prototype tool by reading the instructions page (Section B.3 in Appendix B), reading the help page (Section B.4 in Appendix B), and exploring the prototype tool itself. Second, the participant was given a different set of five images and they were asked to describe them using the prototype tool. Finally, the participant was asked to provide feedback regarding the prototype

tool and their experience, including any problems or any difficulties they experienced and any suggestions for improvement. The participants were encouraged to contact the researcher should they experience technical difficulties that prevented them from completing their tasks. These difficulties were also included as part of the research results.

6.2.1 Prototype Tool Used for Research Study

Writing text alternatives can be an art and it can take a lot of time to formulate an informative one. [8] Likewise, understanding how to apply a tersely worded international standard can be a daunting task. Given the amount of information to extract and consider when writing text alternatives, a support tool would simplify the procedure and help people gather the information being communicated by an image. Although a tool can be used to achieve these goals, the prototype tool used in this research study had another purpose.

The purpose of the prototype tool was to determine if and how much more information was identified when users were explicitly asked to answer all questions in the original document, which was used in Research Study of the Procedure in Document Format. Therefore, the prototype tool required (and verified) that the user consider every question and provide a response. The only sophistication included was to eliminate questions when the category or sub-category of questions did not apply to the image. The prototype tool automatically marked sub-questions as “Not Applicable” if the first question in the category or sub-category was marked as “Not Applicable”.

The prototype tool was developed to assist the user with the procedure outlined in Section 3.1. It was designed to have a simple layout and did not contain graphical content other than the images being described. The prototype tool allowed and helped users to do the following:

- See the set of images they have access to,
- See the image within its original context,
- Specify and delete image components in the image,
- Save the information entered,
- See the information entered previously,
- Provide a more complete set of information about the image, and
- Review the information entered.

Figure 6.1 illustrates the home screen of the prototype tool. In addition to a link to the instructions, the user can see a table of the images they have access to. For each image to be described, the user can: (1) see a thumbnail of the image; (2) view the Web page with the image; (3) add information about the image; (4) see whether or not they have completed their task of identifying information about the image ("Incomplete" and "Complete"); and (5) see the date and time that the user last updated the information (which is automatically recorded). All of the information they entered is saved between sessions. The user can log out at any time and return to add more information to any "Incomplete" images and review the information for any "Complete" images.

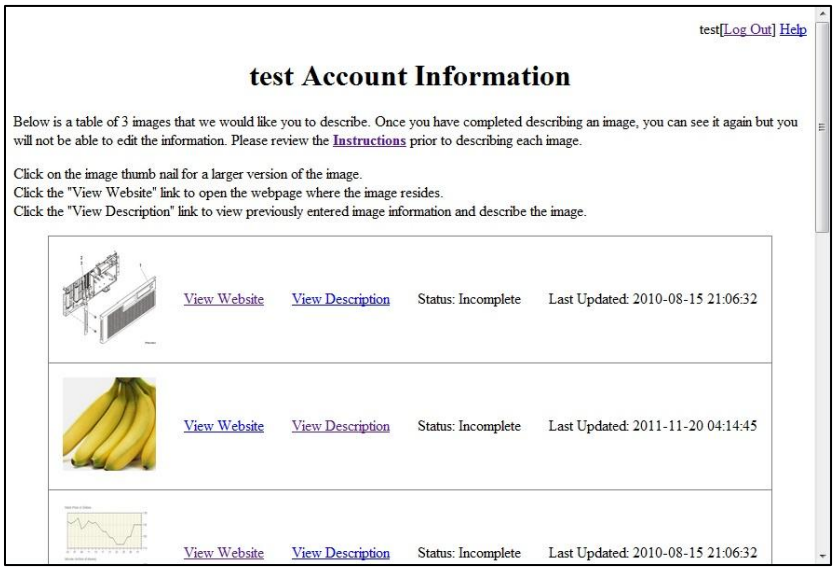


Figure 6.1 Screen Shot of the Prototype Tool Home Screen

For each image, the user can record the identification of image components. For the whole image plus each image component, the prototype tool provides a set of questions that the user was required to answer. For the purpose of this research study, the prototype tool forced the participants to consider and provide a response to every question (or at least every category or sub-category of questions) so as to determine the amount and types of information identified when forced to answer the questions. This is illustrated in Figure 6.2. This screen combines steps 1 to 4 so that the participant can modify the list of components while identifying information. Steps 3 and 4 (identifying general information and then expanding on the information) were

combined so that the participant could choose to use a breadth first approach, depth first approach, or a combination approach when identifying information about the image.

It was recognized that not all questions would be applicable to every image and certain types of information would not be applicable. Therefore, the prototype tool allowed the user to select “Not Applicable” for entire categories or sub-categories of questions automatically. Although the questions were marked as “Not Applicable”, they remained visible in case the user wished to alter their response.

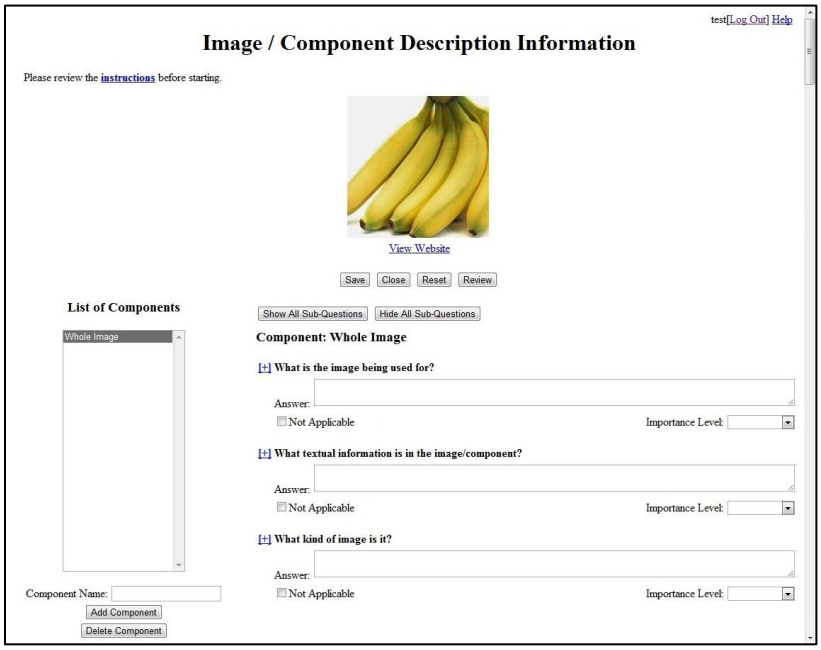
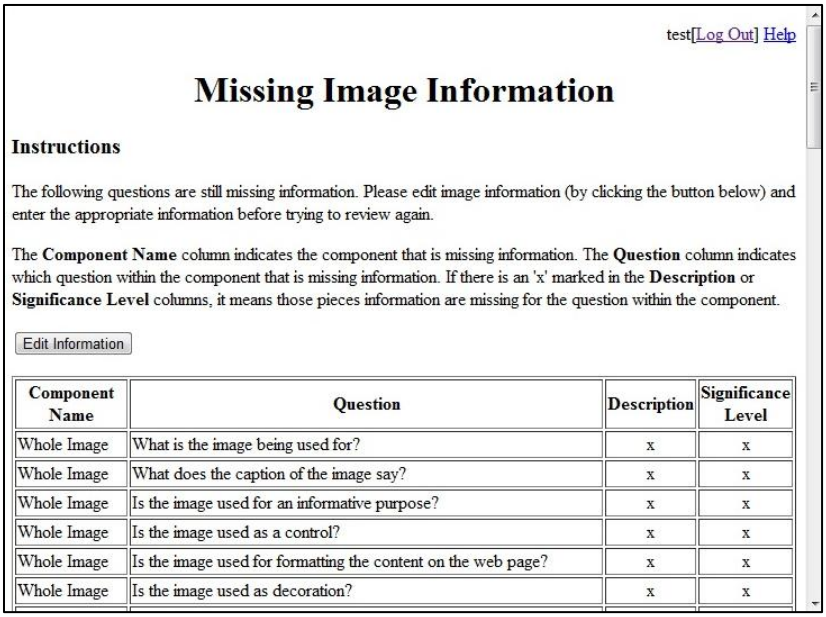


Figure 6.2 Screen Shot for Identifying Information About an Image

After the user finished identifying information, the user was asked to review the information before the image (and their task) was considered as completed. As part of this review process, the prototype tool verified that all of the questions in the set of questions were answered. If there were questions that were not answered, the user was provided with a table listing the questions that were unanswered. The table listed the component name, the question, and either the answer or the importance level was omitted. This is illustrated by Figure 6.3.

The user must go through the review process until every question about every component had a response. Once all questions were answered and the importance levels provided, the user could read over what they had written and record the amount of time it took for them to identify

the information. If the user believed they had missed some important information, they could return to editing the information. This is illustrated by Figure 6.4. After reviewing the information they entered, the user indicated that the information they entered was complete and the image's status was updated to "Complete" on the home page.



test[Log Out] Help

Missing Image Information

Instructions

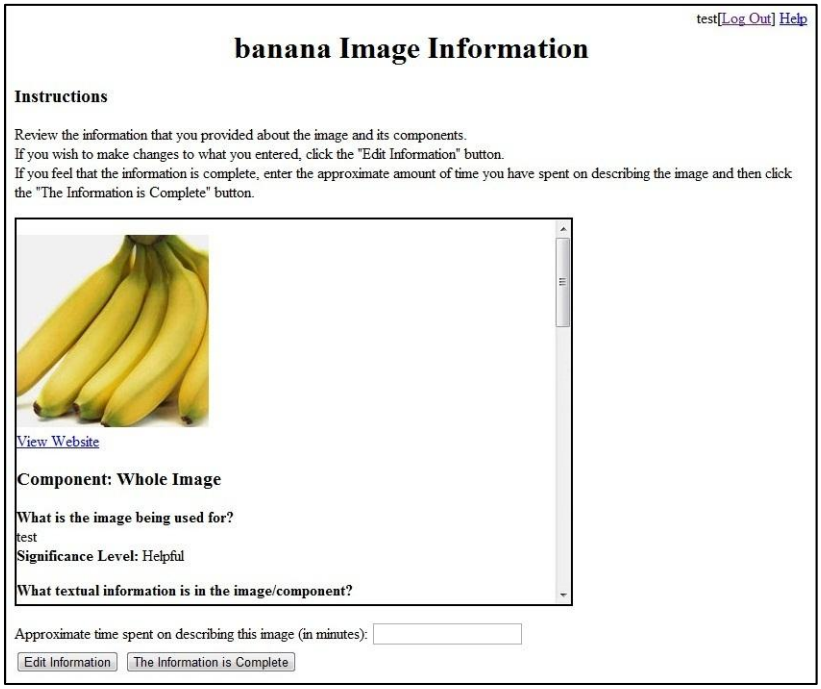
The following questions are still missing information. Please edit image information (by clicking the button below) and enter the appropriate information before trying to review again.

The **Component Name** column indicates the component that is missing information. The **Question** column indicates which question within the component that is missing information. If there is an 'x' marked in the **Description** or **Significance Level** columns, it means those pieces information are missing for the question within the component.

Edit Information

Component Name	Question	Description	Significance Level
Whole Image	What is the image being used for?	x	x
Whole Image	What does the caption of the image say?	x	x
Whole Image	Is the image used for an informative purpose?	x	x
Whole Image	Is the image used as a control?	x	x
Whole Image	Is the image used for formatting the content on the web page?	x	x
Whole Image	Is the image used as decoration?	x	x

Figure 6.3 Screen Shot of the Review Page Showing Missing Information



test[Log Out] Help

banana Image Information

Instructions

Review the information that you provided about the image and its components.
If you wish to make changes to what you entered, click the "Edit Information" button.
If you feel that the information is complete, enter the approximate amount of time you have spent on describing the image and then click the "The Information is Complete" button.

View Website

Component: Whole Image

What is the image being used for?
test
Significance Level: Helpful

What textual information is in the image/component?

Approximate time spent on describing this image (in minutes):

Edit Information | The Information is Complete

Figure 6.4 Screen Shot of Review Page

The prototype tool was a Web-based tool, using PHP and JavaScript with a mySQL database. Since it was a Web-based tool, it was accessible using a Web browser on any operating system. It was tested using Microsoft Internet Explorer, Mozilla Firefox, Apple Safari, and Google Chrome.

The prototype tool is currently available at http://userlab.usask.ca/AltTextTool. Anyone can create an account and try the procedure for themselves. Each account currently has access to ten images, which were used in evaluating both the procedure and the prototype tool. People can apply the procedure to any or all images. Any information and feedback provided will help improve the quality of the procedure and the prototype tool.

6.2.2 Method to Answering Research Questions

Research Question 1

To identify problems and improvements to be made to the procedure and prototype tool (research question 1), the comments and feedback from the participants were analyzed for commonalities. Quite a few issues were raised by multiple participants, which are summarized in Section 6.3.1.

Research Question 2 and 3

Unlike the previous research studies, the specific questions were provided to the participants and they were required to consider every question. Therefore, the questions that extracted each piece of information were explicitly stated. There were times when the participant provided information that should have been identified in a different question or the question was misinterpreted (research question 2 and 3). In this situation, the information was marked as wrongly classified and was re-classified to a more appropriate question(s) if one existed. These instances were considered and summarized in Section 6.3.2.

Research Question 4 to 6

The number of questions that were answered with something other than “Not Applicable” (research question 4) was determined, along with the average, minimum, and maximum.

Averages were also calculated for each user group within Group A and Group B participants for each image. These values were used to compare between user groups (research question 5) and between images (research question 6).

Within each user group, the images were ordered from the most questions answered to the least questions answered, with the consideration for each set of images and then for all images. The ordering of these images was used to determine if a consistent trend existed regarding the images.

The results to these research questions are summarized in Section 6.3.2. These values were used for comparison with other research studies.

Research Questions 7 to 9

Similar to the number of questions answered, the average, minimum, and maximum of the time spent was calculated (research questions 7). The averages were also calculated for each user group within Group A and Group B participants and for each image. These values were then compared between user groups (research questions 8) and between the images (research questions 9). The images were ordered from the most amount of time to the least amount of time to determine if any trends existed about certain images or image types. The ordering was done for each set of images and then for all images.

The results to these research questions are summarized in Section 6.3.4 and used for comparison with the other research studies.

Similar to the previous research studies, the amount of time spent on writing text alternatives for each image was reported by the participants. It was possible that the times were estimates of how long it took rather than the actual amount of time. However, the reported times were good enough for comparison between the research studies.

Research Questions 10 to 12

Similarly, the averages, minimum, and maximum values were calculated for the number of words written in this research study for all participants (research question 10), each user group within Group A and Group B participants, and for each image. These values were then compared

with respect to user groups (research question 11), images (research question 12), and research studies. The images were ordered within each user group (for each set of images as well as overall) in the order of decreasing number of words. These results were analyzed to identify trends regarding the images, if any existed.

The number of words written was used as a measurement for quantity of information about an image rather than quality of that information. It is possible that the same number of words be used while the quality of the words changes. There may be some correlation between the number of words written, the number of questions answered, and the types of information identified.

The results to these research questions are summarized in Section 6.3.5. The results were used to compare with other research studies.

Research Question 13

Like the previous research studies, the results from this research study alone were insufficient for evaluating the quality of the results (research question 13). The information needed to be evaluated for the level of importance, which was done in the Research Study of the Identified Information. The quality of the information identified while using the prototype tool was identified and analyzed by the same process as discussed in Section 4.2.3. The results are summarized in Section 6.3.6.

Research Question 14

The results of this research study were compared to the results of the previous two research studies (research question 14). Similar to the Research Study of the Procedure in Document Format, the analysis compared the amount of time, number of words, number of questions, types of information, and the quality of the information identified between the research studies. The analysis also compared the quality of information between user groups from when the document was used and from when the prototype tool was used.

The analysis considered the types of information where the tool was more effective. This was done by comparing the frequency that the participants identified a certain piece of

information. The tool was considered to be more useful for a particular question when a greater percentage of participants identified a piece of information derived from that question.

The results are summarized in Section 6.3.6.

6.3 Research Study Results

This section describes the analysis of the responses provided by the participants while using the prototype tool. It should be noted that the analysis with respect to particular user groups may not be representative of the user group because some user groups had only one or no participants available at the time of the research study.

6.3.1 Participant Feedback

This section discusses the comments made by the participants regarding the prototype tool, the process for describing the images, and the set of questions presented in Section 3.2.

More Usable Than Document Format

Most participants found that the prototype tool made it easier to understand the procedure and what to do. Unlike in Research Study of the Procedure in Document Format, the participants did not return with comments on being confused while using the prototype tool nor questions on how to apply the procedure. Two participants (11.8%) explicitly commented that the prototype tool was much more comprehensible and usable than the document.

Separate Pages or Wizard

Some participants (23.5%) felt that there was too much occurring on a single page, while others (17.6%) expressed that the set of questions was too long and overwhelming. As shown in Figure 6.2, the participants could add and remove image components as well as answer the set of questions about each image component on a single page. The prototype tool allowed the participants to collapse and/or expand the categories of questions so that the set was easier to manage. However, the categories are expanded by default, resulting in an extremely long page. The participants made two recommendations.

First, they recommended that the image components be identified on a separate page. Second, they recommended that the questions be separated onto several pages. To support this, the participants suggested a wizard-type application be used to guide the user through the procedure. The application might also customize the set of questions according to the type of image and present only the relevant questions for consideration.

Predefined Set of Components

One participant recommended that instead of asking them to identify the important image components to describe, the prototype tool should provide them with a list of image components. The participant was not sure which image components were important because they did not choose the image themselves. Therefore, providing participants with a predefined list of image components would eliminate the need for participants to make such a decision.

Image Visibility

The page for identifying information was set up such that the participant must scroll to view the image being described. The user could also click on the image to open a larger version of the image in a separate window. This method allowed the user to view the image at the same time as identifying information about an image. One participant commented that this method was not suitable for devices with small screens, such as laptops and netbooks. One possible solution suggested by two participants was to use frames in order to increase efficiency.

Saving the Input

The existing prototype tool reloaded the entire page every time an action request was made, such as loading a new component or saving the responses. If the participant performed an action prior to saving, their progress was lost and they were required to re-enter the information. Two participants commented that they could finish in half of the time if their progress was not lost due to this implementation issue.

The participants were instructed to save before loading other pages to prevent their progress from being lost. However, this resulted in another problem. Because the page reloaded after saving, the participant lost their place in the list of questions and had to relocate where they had left off. It was frustrating to have to locate their position again within the long list of questions. One participant suggested that the system state be saved during and between sessions.

Review Information Page

The prototype tool required that all questions in every category to have a response. When a question was missed, the participant was given a table of questions that require attention. The table listed the component name, the question, if a response was provided, and if an importance level was provided. Figure 6.3 (on page 101) illustrates the screen that participants saw. Some participants (17.6%) could not interpret the table nor understand what to do after seeing the table. This showed that additional help and instructions are needed in order for the table to be meaningful and useful. One participant suggested that the table link directly to each question that needed attention so that the participants did not need to locate the question(s) for themselves.

Importance Levels

As with the Research Study of the Procedure in Document Format, the importance levels remained an issue when using the prototype tool. The participants felt that the importance levels were highly subjective and wanted more guidance on how to correctly assign the importance levels. One participant also suggested that the levels of importance be reduced to “Important” and “Not Important” rather than varying degrees of importance.

Confusing or Duplicate Questions

The focus of the comments in this research study was regarding the set of questions (Section 3.2). A few participants (11.8%) commented that some questions were duplicates of each other and they felt that the questions were asking the same thing. As a result, they felt as if

they had answered the same question several times. Participants suggested that clearer distinctions be made between similar questions or the duplicates should be removed.

Help Regarding Individual Questions

Some questions may be too technical or specific for the users to know the answers. For example, one question asked for spatial information using x and y co-ordinates. To know something this specific may need additional work to calculate. One participant suggested that either provide additional help on identifying such information or the questions should be removed.

During the Research Study of the Procedure in Document Format, participants had commented that additional examples were needed to properly interpret how to answer the questions. Tool tips were provided in the prototype tool for this purpose. However, additional help and examples were not provided for questions that appeared very straight forward (such "What is the facial expression of the person?"). As such, some participants (17.6%) found that help was not always available when they needed it. For example, one participant needed help with interpreting the questions regarding State and found none was available. They recommended that additional help be provided for all questions.

Rephrasing Questions

The participants were required to consider every question as described in Section 3.2 even when the question was clearly not related to the image. For example, questions regarding slide shows must be answered even when the image was not part of a slide show. Although the prototype tool supported the ability to mark entire categories of information and questions as not applicable, the feature was not always helpful. Instead of "Not Applicable", sometimes the response may be "No" but this resulted in making the participant answer every sub-question in a category.

The majority of the participants (52.9%) commented that majority of the questions were not applicable. To improve the amount of time the participants spend describing an image, the participants (11.8%) recommended that the questions be rephrased to become yes/no questions

so that non-applicable categories of questions could be quickly eliminated. Some participants (29.4%) suggested that questions be automatically hidden if they were not applicable. The user could then concentrate on the applicable questions.

6.3.2 Misinterpretation of Questions

Each question within the set of questions was created with the intention of identifying specific types of information. Therefore, each set of responses to the questions was analyzed to determine if the questions successfully identified the intended information.

It was concluded that the questions were misinterpreted 8.3% of the time. Certain questions or categories of questions were misinterpreted more often than others. Some common misinterpretations include: (1) stating how the user can benefit from the image's content rather than actions performed on the image to initiate a response from the system (42.7%); (2) stating the type of information being communicated rather than the information itself (12.6%); and (3) confusing between perceptual and subjective information with regards to colour (8.0%). The remaining 36.7% did not have a common reason for misinterpretation. Examples of the uncommon misinterpretations include identifying information about when the image took place for a Where question and identifying objects for a Who question.

6.3.3 Number of Questions Answered

Table 6.5 illustrates the frequency that each type of information was identified by the participants of this research study. The questions most frequently answered were those regarding spatial relationships (15.5%); the colour, shape, and size of objects in the image (13.2%); the meaning and symbolisms represented in the image (9.4%), and how to interact with the image (8.2%). The types of information that was more often missed include State Information, Location/Place, When, and Classification.

Based on the results of this research study, it was calculated that the participants answered an average of 32.2 questions, with a standard deviation of 7.5 questions, a minimum of 8 questions, and a maximum of 89 questions while identifying information through the use of the prototype tool. Figure 6.6 shows a distribution of the results. The actual values can be found in Table C.4 in Appendix C.

Table 6.5 Usage of Categories and Sub-categories of Information With the Procedure in Tool Format

Category or Sub-category	Percentage
State Information	1.3%
Where – Location/Place	2.9%
When	3.4%
What – Classification	3.7%
What – Logical Relationships	4.0%
What – Textual	4.1%
How Much	4.3%
When – Sequential Relationships	4.9%
Why	5.7%
Who	6.3%
What – Physical Object	6.8%
How	8.2%
What – Subjective	9.4%
What – Perceptual	13.2%
Where – Spatial Relationships	15.5%

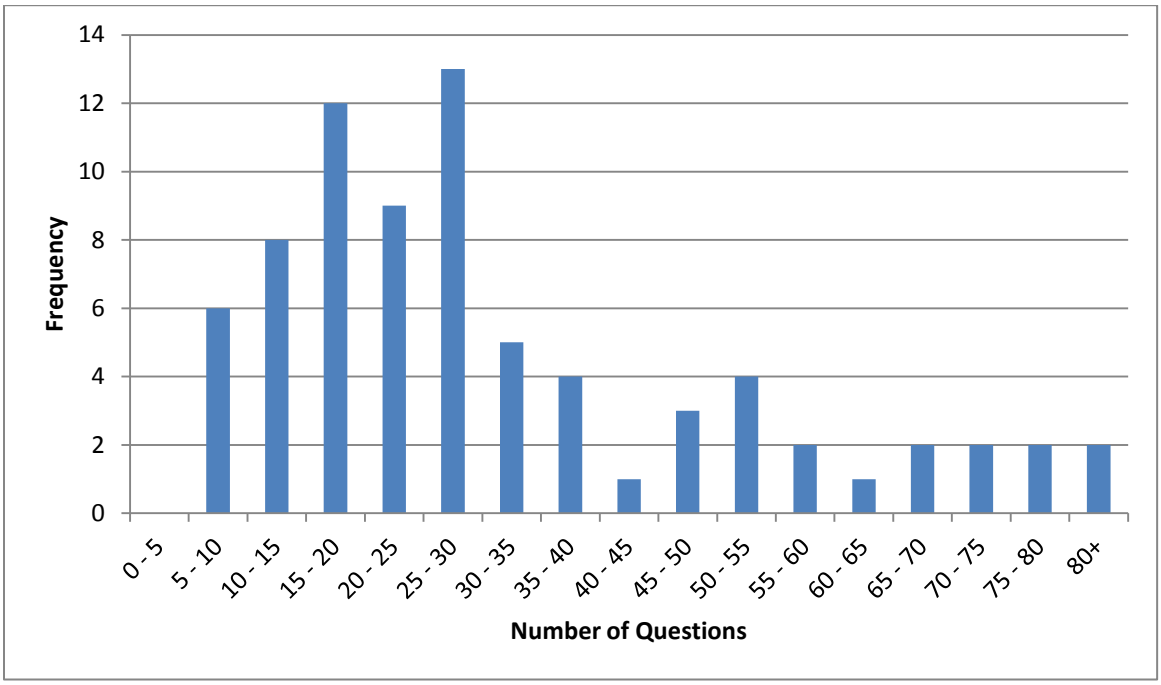


Figure 6.6 Number of Questions Answered With the Use of the Procedure in Tool Format

Figure 6.7 shows the average number of questions each user group answered while identifying information about each image. The exact values can be found in Table C.6 in Appendix C. Table 6.8 shows the average number of questions that each user group answered for all images overall, and for Set A images and Set B images. Based on these results, it can be seen that some user groups answered more questions than others. For Set A images, the Specialists

answered the most questions while the Developers answered the least questions. For Set B images, the Developers answered relatively more questions while Internet Users answered the least questions.

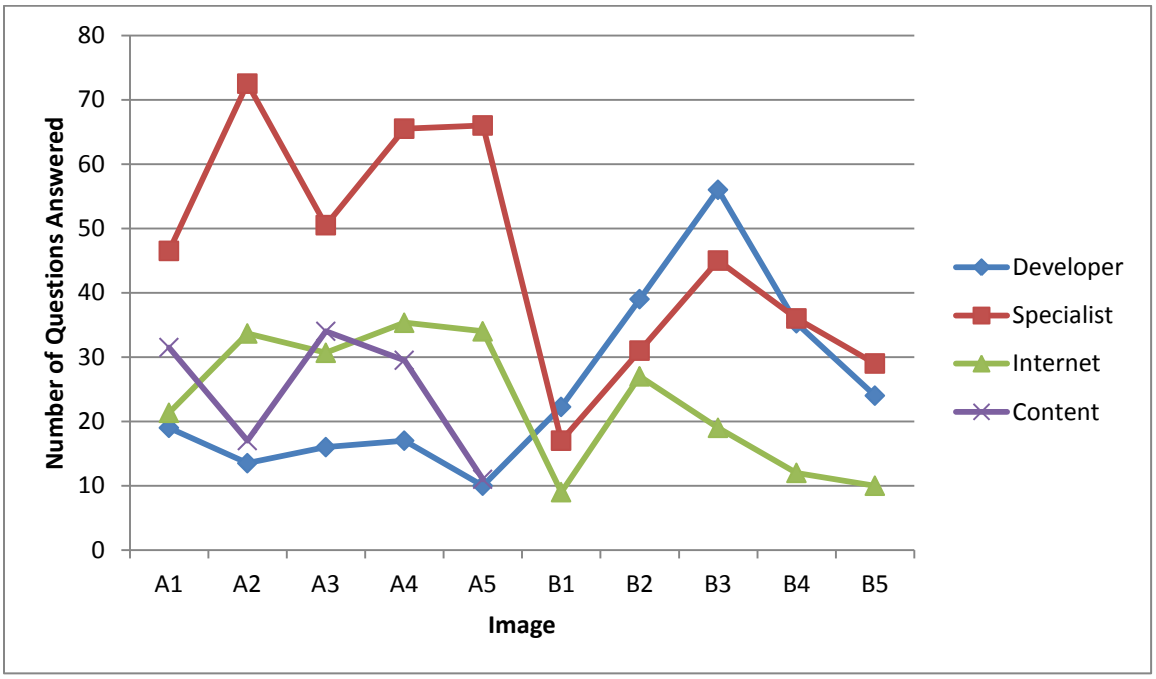


Figure 6.7 Average Number of Questions Answered for Each Image by Each User Group with the Use of the Procedure in Tool Format

Table 6.8 Average Number of Questions Answered by User Group with the Use of the Procedure in Tool Format

	All	Specialists	Developers	Content Providers	Internet Users
Overall	32.2	45.9	25.2	24.6	23.2
Set A	33.4	60.2	15.1	24.6	31.0
Set B	31.1	31.6	35.3	N/A	15.4

Table 6.9 illustrates the images in decreasing order of words written by each user group and overall. It can be shown that more questions were answered for some images and fewer questions were answered for others. For example, B2 and B3 fairly consistently had more questions answered by all user groups while B1 and B5 fairly consistently had fewer questions answered. For Set A images, there was no consistency between user groups as to which images had answered more questions or fewer questions.

Table 6.9 Questions Answered Per Image in Decreasing Order with the Use of the Procedure in Tool Format

	Most									Least
Overall	B3	A4	A2	B2	A5	B4	A3	A1	B5	B1
Specialists	A2	A5	A4	A3	A1	B3	B4	B2	B5	B1
Developers	B3	B2	B4	B5	B1	A1	A4	A3	A2	A5
Content Providers	A3	A1	A4	A2	A5					
Internet Users	A4	A5	A2	A3	B2	A1	B3	B4	B5	B1

6.3.4 Amount of Time Spent

After identifying information about each image, the participants were asked to provide the amount of time it took for them to perform their task. All 76 instances were reported. Based on the reported times, the participants took an average of 31.2 minutes, with a standard deviation of 8.4 minutes, a minimum of 3 minutes, and a maximum of 160 minutes to complete the procedure based on the prototype tool. Figure 6.10 shows a distribution of the results. The actual values can be found in Table C.12 in Appendix C.

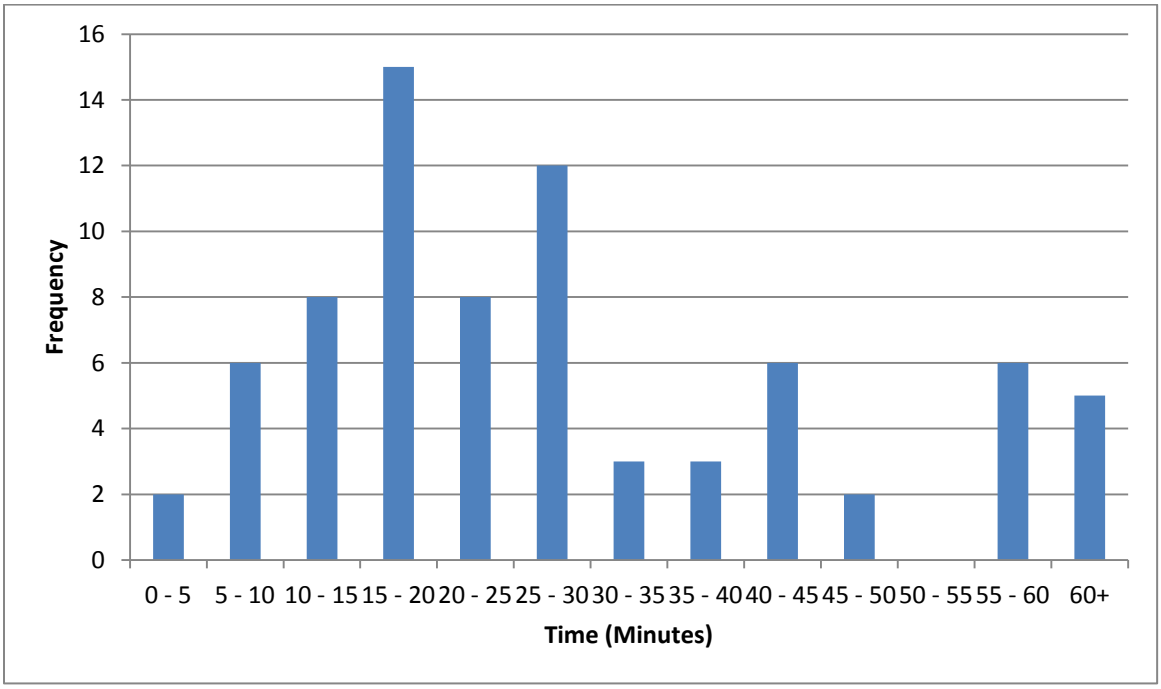


Figure 6.10 Amount of Time Spent with the Use of the Procedure in Tool Format

Figure 6.11 shows the average amount of time each user group spent on identifying information about each image. The exact values can be found in Table C.14 in Appendix C.

Table 6.12 shows the average times that each user group spent on all images overall, and for Set A images and Set B images. Based on these results, it can be seen that some user groups spent more time than others. For Set A images, the Specialists spent the most amount of time while the Developers spent the least amount of time. For Set B images, the Developers spent the most amount of time while the Internet Users and Specialists spent relatively less amount of time.

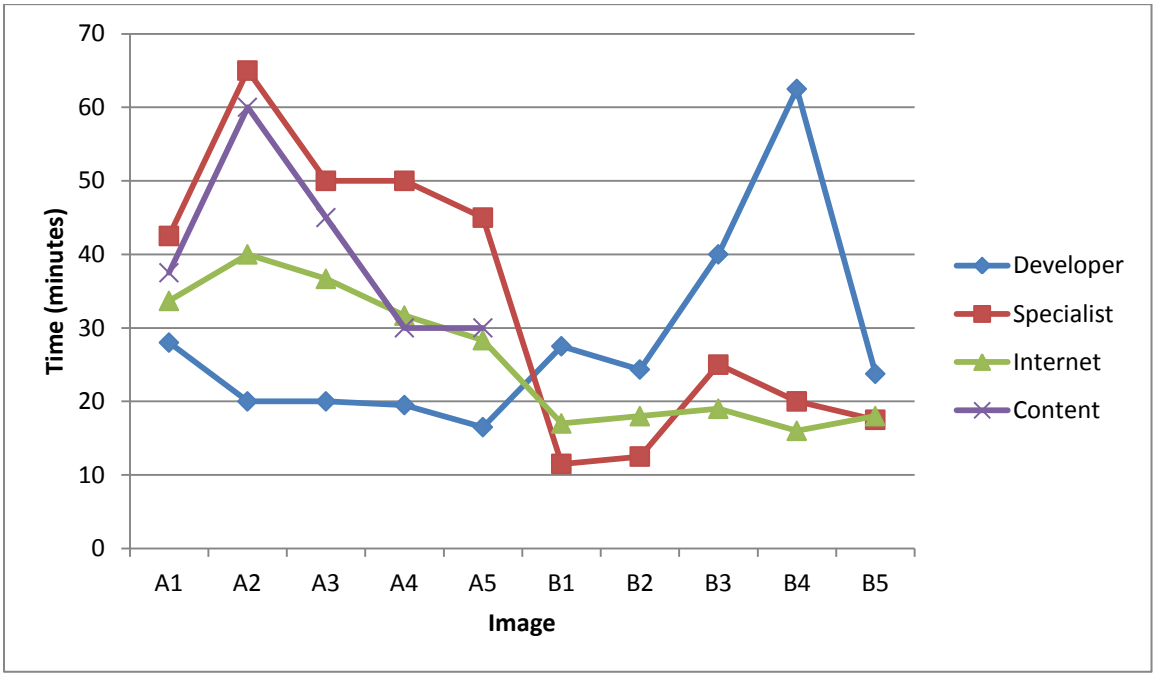


Figure 6.11 Average Time Spent on Each Image by Each User Group with the Use of the Procedure in Tool Format

Table 6.12 Average Time Spent (in Minutes) by User Group with the Use of the Procedure in Tool Format

	All	Specialists	Developers	Content Providers	Internet Users
Overall	31.2	33.9	28.2	40.5	25.8
Set A	35.7	50.5	20.8	40.5	34.1
Set B	26.8	17.3	35.6	N/A	17.6

Table 6.13 illustrates the images in decreasing order of words written by each user group and overall. It can also be seen that some types of images required relatively shorter or longer to apply the procedure than others for most or all user groups. For example, A4 and A5 took less time to identify information by the user groups while it took longer for A2, A3, and B3. Overall,

images in Set A (average 35.7 minutes) took more time than images in Set B (average 26.8 minutes).

Table 6.13 Time Spent Per Image in Decreasing Order with the Use of the Procedure in Tool Format

	Most									**Least**
Overall	A2	B4	A3	A1	A4	B3	A5	B1	B5	B2
Specialists	A2	A3	A4	A5	A1	B3	B4	B5	B2	B1
Developers	B4	B3	A1	B1	B2	B5	A2	A3	A4	A5
Content Providers	A2	A3	A1	A4	A5					
Internet Users	A2	A3	A1	A4	A5	B3	B2	B5	B1	B4

6.3.5 Number of Words Written

A word processor program was used once again to count the number of words written by a participant for a given image. Based on these results, the participants wrote an average of 154.8 words, with a standard deviation of 60.7 words, a minimum of 11 words, and a maximum of 654 words for a given image. Figure 6.14 shows a distribution of the results. The actual values can be found in Table C.20 in Appendix C.

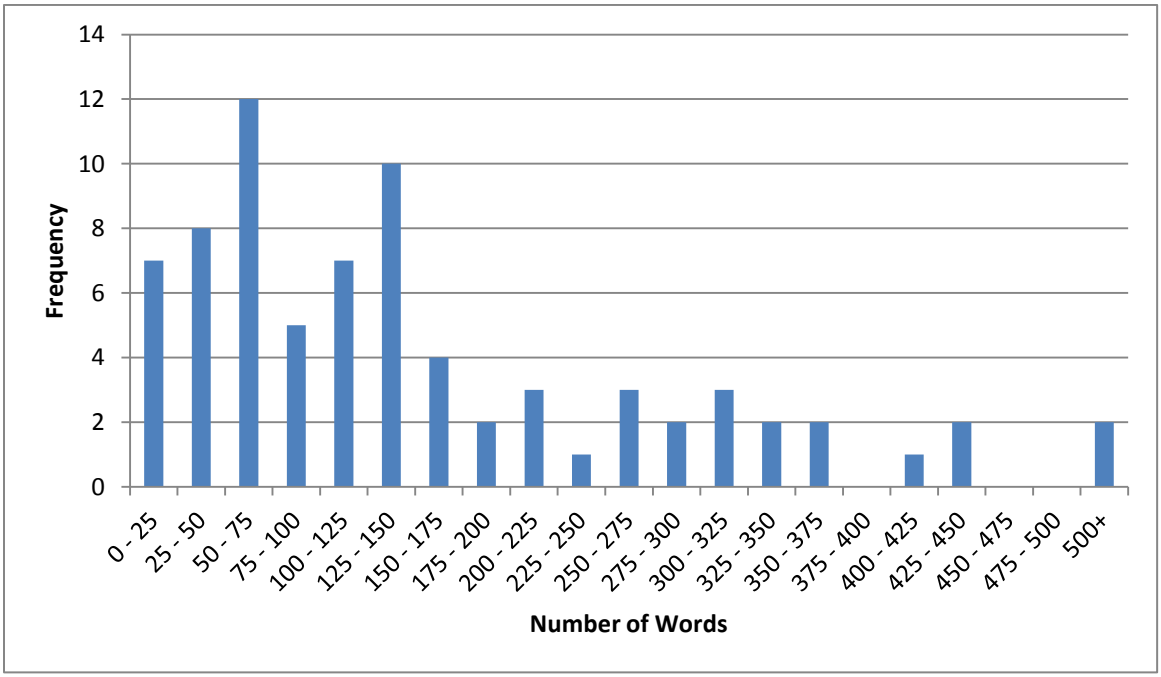


Figure 6.14 Number of Words Written with the Use of the Procedure in Tool Format

Figure 6.15 shows the average number of words that was written by each user group for a given image. The exact values can be found in Table C.22 in Appendix C. Table 6.16 shows the average number of words that each user group wrote on all images overall, and for Set A images and Set B images. Based on these results, it can be seen that overall, more words were generated by Group A participants than Group B participants. Within both groups, Specialist wrote the most number of words. For Set A images, Developers wrote the least number of words and for Set B images, Internet Users wrote the least.

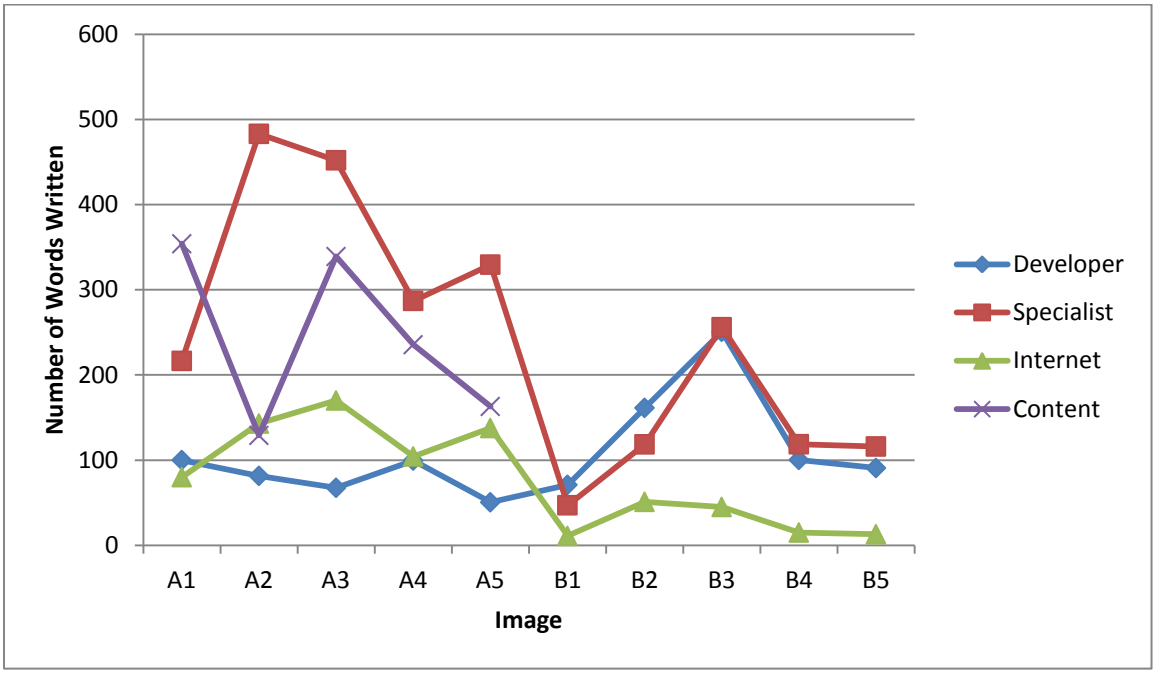


Figure 6.15 Average Number of Words Written for Each Image by Each User Group with the Use of the Procedure in Tool Format

Table 6.16 Average Number of Words Written by User Group with the Use of the Procedure in Tool Format

	All	Specialists	Developers	Content Providers	Internet Users
Overall	154.8	242.4	107.3	244.1	77.0
Set A	192.5	353.6	79.8	244.1	127.0
Set B	117.2	131.2	134.8	N/A	27.0

Table 6.17 illustrates the images in decreasing order of words written by each user group and overall. It can also be seen that more words written by all user groups about certain image types than others. For example, the user groups wrote more words about A3 and B3, while they

wrote fewer words about B1. Overall, more words were written for images in Set A (average 192.5 words) than Set B (average 117.2 words).

Table 6.17 Words Written Per Image in Decreasing Order with the Use of the Procedure in Tool Format

	Most									**Least**
Overall	A3	B3	A2	A1	A4	A5	B2	B4	B5	B1
Specialists	A2	A3	A5	A4	B3	A1	B5	B2	B5	B1
Developers	B3	B2	B4	A1	A4	B5	A2	B1	A3	A5
Content Providers	A1	A3	A4	A5	A2					
Internet Users	A3	A2	A5	A4	A1	B2	B3	B4	B5	B1

6.3.6 Quality of Identified Information

When the participants were given the procedure in the form of a tool, they identified an average of 39.4% of the information that was rated as important to some extent based on the results of the Research Study of the Identified Information. This involved an average of 65.1% of the Essential information, 40.0% of the Significant information, and 24.7% of the Helpful information. Participants also identified an average of 43.4% of the information that was later considered Not Important. Table 6.18 illustrates more details regarding the quality of information. These results were further broken down by user group. The average percentages of each user group identifying the four different levels of importance are illustrated in Figure 6.19. The values can be found in Table C.28 in Appendix C.

Table 6.20 lists the top five types of information that the participants often missed within each level of importance when they used the prototype tool. Table 6.21 lists the top five types of information that participants often included within each level of importance when they used the prototype tool.

Table 6.18 Summary of Quality of Information Identified with the Procedure in Tool Format

	Essential	**Significant**	**Helpful**	**Not Important**
Average (percent)	65.1%	40.0%	24.7%	43.4%
Average (identified/total)	60.9 / 94	33.6 / 84	41.3 / 167	21.9 / 50
Maximum (percent)	100.0%	69.3%	62.7%	100.0%
Minimum (percent)	37.1%	0.0%	2.1%	0.0%

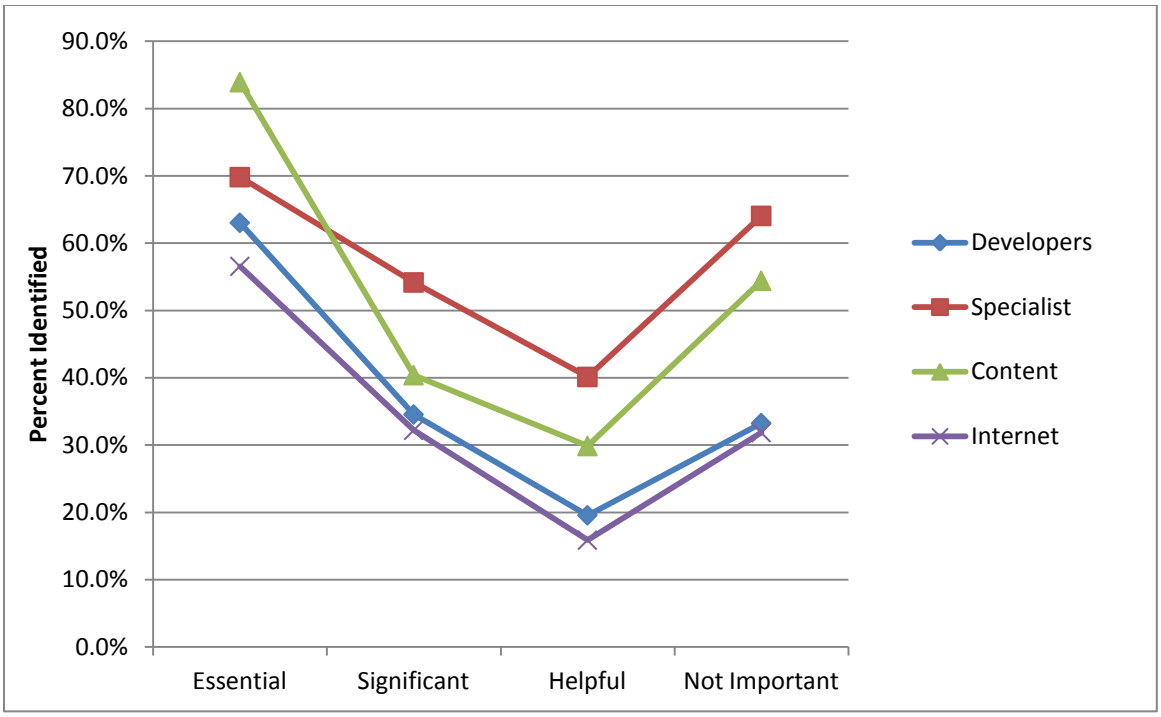


Figure 6.19 Average Percentage of Information Identified by Each User Group with the Use of the Procedure in Tool Format

Table 6.20 Types of Information Often Missed with the Procedure in Tool Format by Importance Level

Essential	Significant	Helpful
1. What – Perceptual 2. What – Physical Objects 3. When – Sequential Relationships 4. What – Subjective 5. What – Textual	1. What – Physical Objects 2. How Much 3. What – Perceptual 4. What – Subjective 5. Where – Spatial Relationships	1. What – Perceptual 2. Where – Spatial Relationships 3. What – Subjective 4. When – Sequential Relationships 5. What – Physical Objects

Table 6.21 Types of Information Often Identified with the Procedure in Tool Format by Importance Level

Essential	Significant	Helpful
1. What – Classification 2. What – Physical Objects 3. What – Perceptual 4. Why 5. What – Subjective	1. What – Physical Objects 2. What – Perceptual 3. When – General 4. Who 5. What – Textual	1. Where – Spatial Relationships 2. What – Subjective 3. Who 4. Why 5. What – Physical Objects

Compared against the results of the Research Study of Text Alternatives Without the Procedure, there were 291 instances where a greater percentage of participants identified a piece of information while using the prototype tool. Of these 292 instances, 69 were Essential, 65 were Significant, 115 were Helpful, and 43 were Not Important. Table 6.22 breaks this down by types

of information. There were also 29 instances where a lower percentage of participants identified a piece of information while using the prototype tool. Of these 28 instances, 6 were Essential, 9 were Significant, 13 were Helpful, and 1 was Not Important. Table 6.23 shows the breakdown of these instances by types of information.

Table 6.22 Instances When Greater Percentage of Participants Identified Information Using a Procedure in Tool Format than Without Using a Procedure by Importance Level

	Essential	Significant	Helpful	Not Important
Why	7	0	2	2
What – Textual	6	5	2	1
What – Classification	7	1	0	0
What – Physical Objects	13	12	3	1
What – Perceptual	17	6	37	14
What – Subjective	5	8	23	0
What – Logical Relationships	2	3	1	0
Who	1	2	7	0
Where – Location/Place	1	2	1	0
Where – Spatial Relationships	2	7	23	25
When – General	1	4	1	0
When – Sequential Relationships	6	1	7	0
How Much	1	11	2	0
How	0	3	6	0
Total	69	65	114	43

Table 6.23 Instances When Lower Percentage of Participants Identified Information Using a Procedure in Tool Format than Without Using a Procedure by Importance Level

	Essential	Significant	Helpful	Not Important
Why	0	0	0	0
What – Textual	0	0	0	0
What – Classification	0	0	0	0
What – Physical Objects	3	3	2	0
What – Perceptual	0	3	3	1
What – Subjective	1	0	2	0
What – Logical Relationships	0	0	0	0
Who	1	1	1	0
Where – Location/Place	0	0	1	0
Where – Spatial Relationships	0	0	3	0
When – General	0	0	0	0
When – Sequential Relationships	1	0	0	0
How Much	0	2	1	0
How	0	0	0	0
Total	6	9	13	1

Compared against the results of Research Study of the Procedure in Document Format, there were 269 instances where a greater percentage of participants identified a piece of information while using the prototype tool. Of these 268 instances, 76 were Essential, 52 were Significant, 103 were Helpful, and 38 were Not Important. Table 6.24 breaks this down by types of information. There were also 100 instances where a lower percentage of participants identified a piece of information while using the prototype tool. Of these 100 instances, 12 were Essential, 26 were Significant, 52 were Helpful, and 10 were Not Important. Table 6.25 shows the breakdown of these instances by types of information.

Table 6.24 Instances When Greater Percentage of Participants Identified Information Using a Procedure in Tool Format than in Document Format by Importance Level

	Essential	**Significant**	**Helpful**	**Not Important**
Why	8	0	2	3
What – Textual	6	3	2	0
What – Classification	9	1	0	0
What – Physical Objects	13	11	3	1
What – Perceptual	17	8	29	12
What – Subjective	7	6	23	0
What – Logical Relationships	2	3	1	0
Who	3	3	7	0
Where – Location/Place	1	2	2	0
Where – Spatial Relationships	2	6	20	22
When – General	1	4	0	0
When – Sequential Relationships	4	0	6	0
How Much	1	2	2	0
How	2	3	6	0
Total	76	52	103	38

Table 6.25 Instances When Lower Percentage of Participants Identified Information Using a Procedure in Tool Format than in Document Format by Importance Level

	Essential	Significant	Helpful	Not Important
Why	0	0	0	0
What – Textual	0	2	1	1
What – Classification	0	0	0	0
What – Physical Objects	4	5	7	1
What – Perceptual	3	3	19	4
What – Subjective	0	3	7	0
What – Logical Relationships	0	0	0	0
Who	0	0	2	0
Where – Location/Place	1	0	0	0
Where – Spatial Relationships	0	1	9	4
When – General	0	0	1	0
When – Sequential Relationships	4	1	5	0
How Much	0	11	1	0
How	0	0	0	0
Total	12	26	52	10

6.4 Analysis of Research Study Results

This section analyzes the results from the Research Study of the Procedure as a Prototype Tool to identify common trends, issues with the tool and the procedure, and possible solutions to those issues. The results were analyzed overall, between user groups, between images, between types of information, and between individual questions.

Possible improvements and modifications to the procedure, guidance, and tool that are identified throughout this section will be denoted using “IMP” followed by a number. For example, “IMP 1” and “IMP 2”. Many of the improvements to the ISO document could not be implemented prior to the publication of the document but will be implements in future versions of the document. The improvements to the tool itself are beyond the scope of this thesis but the results will be available to future tool development efforts.

6.4.1 Analysis of Feedback

Layout Design Improvements

A tool was supposed to accomplish two goals:

1. Simplify and clarify the procedure.
2. Decrease the amount of time needed to complete the procedure.

While the prototype tool succeeded in clarifying the procedure and the participants knew better what to do, the prototype tool did not lower the time to complete the procedure. This was likely due to the design of the prototype tool.

This research study identified several improvements that should be made in any future tool in order to satisfy users and improve the amount of time spent on the procedure, including separating content by making use of a wizard and/or with the use of frames.

A wizard could be used to guide the user through each step of the procedure. While this may be feasible for parts of the procedure (such as to identify the purpose and image components, Steps 1 and 2), it would not be feasible for identifying and elaborating on image content, Steps 3 and 4. Since the set of questions is meant to be answered for each image component, a wizard may not be the best method.

Instead, it would be more feasible to separate the tasks such that the user could focus on each step one at a time. The current prototype tool places Steps 1 to 4 in a single page, which resulted in participants stating that it was too much on a single page. The results showed that participants prefer for the steps to be separated. Separating each step into its own page would allow for clear separation of tasks and support the complexity of each step. Making use of frames (or a similar technique) would allow the user to see various elements of the page at the same time in order to complete their task. For example, it was commented that the participants wanted to see both the image and the questions about the image at the same time. The page could be separated such that it consists of one section for the image, one for the questions, one for the list of image components, and one for the navigation buttons. Depending on the separation of the task and the technology used, a different layout may be necessary.

IMP 16. In the tool, separate each step of the procedure into different screens. This may be done through the use of a wizard.

IMP 17. In the tool, make it possible to see the image being described and the questions at the same time constantly.

Default List of Image Components

Currently, online games exist where users are shown an image and they identify words to tag the image (e.g., Google Image Labeler [23]). While the games can be used to provide a basic idea of the image content, image tagging or indexing is not enough to fully describe the information portrayed in the image within its context. However, this tagging approach could be used as a starting point for identifying a default set of image components for a given image.

Providing a default list of image components for a given image is not desirable or feasible since the importance of an image component changes with each usage of the image. Therefore, there may be image components in the list that are not important due to the context. Using default image components would require the user to consider information about each image component even when the image component is not important. Similarly, there may be image components that are important due to the context but are not in the default list. Information about these components may not be identified as a result.

In future designs of the tool, additional guidance or other implementations for this step of the procedure may be needed to help users identify the image components. Additional research is necessary to consider if and how existing image tagging games could be used to help identify image components.

IMP 18. In ISO 20071-11 and the tool, add guidance to further help identify image components. This may involve providing additional examples.

Automatic Saving

During the design of the prototype tool, automatically saving responses was considered. However, it was not possible to do it unnoticed (that is, without the page reloading) due to the technologies used. It was also considered that data would be saved when an action was performed (such as selecting another image component). It was not possible to have the client

side of the system (Web browser) to execute functions performed by the server side of the system. That is, it was not possible to save the responses without fully reloading the page, which is noticeable. Frequently reloading the page for the sake of saving the responses may be frustrating to the user. It may also cause confusion since the user may not know why the page is reloading itself.

Therefore, it was decided that a button would be made available for users to save their responses. It required that users explicitly remember to save before performing other actions. As the feedback showed, this was not a desirable solution. The future design of the prototype tool should utilize technologies that allowed automatic saving of responses without disturbing the flow of work.

Each time that an action was performed, the page was reloaded and the participant had to relocate the question they had been working on. This was very frustrating and time consuming, therefore, the participants suggested that the tool save the state of the questions (either collapsed or expanded) as well as the location of the question that they were working on. These features were missed in the original design of the prototype tool, but should be supported in future implementations.

IMP 19. In the tool, the responses to questions should be automatically saved.

Reviewing Information

Some participants needed additional help to interpret the table of missing information that was presented to them (Figure 6.3 on page 101). This showed that users will need additional instructions and guidance on how to interpret the table as well as what they should do with the information. One suggestion made by the participants was to provide a link directly to the location of the question. The intention would be to help users to quickly locate the questions they had missed.

Another solution may be to allow the user to provide the responses directly in the table. This solution would eliminate the need for further navigation and it clearly indicates that a response is necessary. This may help users understand the information being presented on this page as well as inform them of what to do.

IMP 20. In the tool, responding to missed questions should be done at the same location as identifying the missed questions.

Clarify Importance Levels

The version of the importance levels given to the participants of this research study was different from the version provided in Section 3.1.6 (on page 34). Similar to the Research Study of the Procedure in Document Format, the participants were provided with only the definitions of each importance level, without the suggestions as to when each level may be assigned to a piece of information.

At the time, it was believed that the definitions were sufficient for the task since the importance of information changes with the context. It is subjective and the user's judgment and knowledge is required to decide the importance of a piece of information. The participants and other accessibility experts wished to have more concrete and definitive ways of assigning importance levels. They also suggested that the importance levels be reduced to "Important" and "Not Important".

Simplifying importance into "Important" and "Not Important" is insufficient because it does not allow for varying degree of importance. As noted in captioning and audio descriptions (Section 2.3), the most important information is provided first. If time and space is available, the next important information is provided. Therefore, even within "Important", there would need to be some distinction between the different information to determine which pieces were more important than others. Hence the varying levels of importance are necessary.

It is difficult to have definitive rules about assigning importance levels because the context and purpose may be different each time. Instead, guidance was added about possible situations or reasons for a piece of information to be considered a certain level of importance. The guidance took the following into consideration:

- Necessity of a piece of information for comprehending the image,
- Necessity of a piece of information for comprehending the document,
- Intention to provide or communicate a piece of information, and
- Conflicting information between the image and the document.

This improvement was previously identified in Research Study of the Procedure in Document Format (IMP 12 and IMP 13).

Repeated Questions

The set of questions consisted of some repeated questions between categories and within a category. Some questions (such as those involving actions) can apply to multiple categories (such as Who and Logical Relationships). Sometimes, it was logical for certain questions to be placed under multiple categories. It ensured that the question would be answered and the information would be identified.

Certain questions, such as Subjective, were sometimes phrased in several ways in an attempt to extract the same information. This was because different people may interpret the questions differently or they may not understand certain questions when phrased a certain way. Having multiple questions that targeted the same information ensured that the information would be identified.

It was noted in the document that the questions may overlap between categories; however, this was not made clear in the prototype tool. The tool chose to have information identified multiple times and removed later rather than to not have the information at all.

IMP 21. In ISO 20071-11 and the tool, clarify the need to have repeated or similar questions within a category an overlap of questions between categories.

Help and Examples for Individual Questions

There was a subset of questions that did not provide additional help in the form of tool tips. These questions were deemed to be straight forward and, therefore, would not require additional help to interpret the question and identify the information.

One sub-category that was misjudged was the State sub-category. This type of information was meant to be used by those with a technical background and understood concepts such as the current state of the system. However, the procedure is meant to also be used by those without a technical background. This was missed and should be added. The other questions

should be re-analyzed to provide help to people with and without expertise or necessary background.

IMP 22. In the tool, provide additional help and examples for every single question in the set of questions.

Sequence and Order of Questions

Streamlining the set of questions could greatly improve the speed of the procedure and quickly identify questions that were applicable. While some questions were already designed to support this (such as "Does the image/component contain perceptual information?"), there was no clear flow or path for users to follow such that they would be able to quickly identify important questions given the context or situation.

The use of Yes/No questions could quickly identify the appropriate questions. However, additional research is necessary to design the sequence for presenting the questions such that it ensured all possibly important information would be identified.

Instead of presenting the questions by category or sub-category, it is possible that the sequence of questions follow the logical order or train of thought when considering information. For example, asking the Perceptual question regarding shape followed by asking the Subjective question regarding symbolism. Further research is necessary to determine the best logical order of presenting the questions to quickly identify the important information.

IMP 23. In the tool, make use of Yes/No questions to quickly identify questions that apply to an image.

6.4.2 Misinterpretation of Questions

While answering questions to identify information about an image, there were times when the participants interpreted and used a question differently than intended. This resulted in identifying information that was not intended. Some common misinterpretations include stating how the user can benefit from the image's content rather than actions performed on the image to initiate a response from the system; stating the type of information being communicated rather

than the information itself; and confusing between perceptual and subjective information with regards to colour.

The intent of the How questions were to describe the control or technical function for which an image is being used, as well as outcomes when the function is executed. Instead, some of the participants interpreted the question as the intended result of viewing the image (such as to purchase a product for an advertisement or appreciate art for a painting). In the document, the questions are presented along with an explanation for that category of questions, providing a context for the question. However, the context is not presented in the prototype tool version unless the user makes use of and considers the help tool tips. Therefore, the questions should be rephrased to provide the context of using the image as a control or the prototype tool should provide the context when presenting the set of questions.

IMP 24. In the How category, modify the questions to indicate the intent or context of the question. For example, "What action is the user supposed to perform in order to interact with the image (or image component)?"

Within the set of questions, there is one question to identify the words in the image caption so that the same information would not be repeated in the text alternative. There is another question to identify the words presented in the image itself so that users would know the exact text that's given to sighted users. In both cases, the intent is to identify the exact words that are meant to be presented to users. However, there were many instances (21.1% of the time) where the participants stated the type of information (such as "company name") instead of the information itself (such as "iQmetrix"). Therefore, the questions should be rephrased to indicate that the exact words in the caption and image text should be stated instead of describing the type of information being communicated.

IMP 25. In the set of questions, change the question regarding captions to state "What are the exact words stated in the image's caption?"

IMP 26. In the Textual sub-category, change the question regarding text to state "What are the exact words shown in the image?"

There are two questions within the set of questions that deal with colour, one is categorized under Perceptual and the other is categorized under Subjective. The intent of the Perceptual question is to specify the colour(s) in the image component and the intent of the Subjective question is to identify the meaning or symbolism of the colour(s), if one exists. The Subjective question states "If the colour(s) of the image or image component is important, what is the colour(s) representing?" This clearly indicates that the intent is to identify the symbolism but many people used the question to specify the colour. One possible reason may be the Subjective question was presented first; therefore, the information perceptual information was identified under Subjective and then repeated when the Perceptual question was presented. The participants also may have misread the question and not notice the difference between the questions. One resolution is to present the Perceptual question first. Another resolution would be to ask the question one after the other so that there would be a clearer distinction between the two questions. This latter solution is more suitable as a future improvement to the prototype tool because the tool can present questions using a logical flow, whereas the document is meant to present the procedure in a logical manner.

IMP 27. In the set of questions, rearrange the questions such that the Perceptual questions appear before Subjective questions.

In addition to the resolutions described above, the questions could also benefit from additional examples to illustrate the intent of the questions and the types of responses that are expected.

IMP 28. Add example responses to each of the questions to the set of questions to illustrate the intent of the questions and the types of responses expected.

6.4.3 Analysis of Time, Words, and Questions

By this time, the participants had some experience with the procedure; therefore, the participants may show some improvement with regards to time and the information identified about the image. However, the amount of time spent on describing each image was much greater than anticipated.

In the Research Study of the Procedure in Document Format, a lot of time was spent on figuring out how to apply the procedure. With the tool, the participants had a better understanding of the procedure and the time was spent on applying the procedure more fully. Although the participants were more familiar with the procedure, a much longer amount of time was spent on completing the tasks fully, as was forced by the tool.

The participants were required to consider and respond to every single question as described in Section 3.2. With the increased number of questions to consider, more time was necessary to complete the procedure for each image. This was especially true when multiple image components were identified since each image component has its own set of questions to consider. One participant supported this view and commented that describing the image took a long time due to the number of components and the number of questions. This also had an effect on the number of questions answered and the number of words written. With more questions being considered, additional words were written in order to respond to those questions.

Table 6.26 summarizes the images in the research study that the user groups spent relatively more time, wrote relatively more words, and answered relatively more questions when identifying information about the image. It also summarizes the images where user groups spent relatively less time, wrote relatively fewer words, and answered relatively fewer questions.

Table 6.26 Summary of Time, Words, and Questions Output with respect to Images When Using the Procedure in Tool Format

	More	**Less**
Time Spent	A1, A3, B3	A4, A5
Words Written	A3, B3	B1
Questions Answered	B2, B3	B1, B5

The table shows that for some images, there was a relationship between the amount of time spent and the number of words written. There was one instance (B3) where more time was spent, resulting in more questions answered and more words written. However, similar to the results of Research Study of the Procedure in Document Format, this trend was not consistent across all images. In most cases, there was no correlation between time spent and number of words, between time spent and number of questions, nor between words written and number of questions.

Similar to Research Study of the Procedure in Document Format, there was a more consistent trend within individual user groups than overall for all participants. As users spent more time on an image, more questions were answered and more words were written. This was especially true for Specialists and Content Providers. While the trend was less noticeable for Developers and Internet Users, it still existed. It was then not surprising that Internet Users, who spent the least amount of time, to answer average number of questions and write below average number of words. Although Developers generated more output than Internet Users, it was still below average.

Specialists generated greater output than other user groups. They were the only user group that was above average in all three aspects (time, words, and questions). It may be because this research is within their field and was of more interest to them than other user groups. This user groups showed the strongest correlation between time, words, and questions being proportional to each other.

Table 6.27 summarizes the overall average time spent, words written, and questions answered for images in Set A and Set B. The images in Set A appeared to have taken more time to describe and generated more words than those in Set B. This is the opposite of the results reported in Research Study of the Procedure in Document Format. The discrepancy in the results between the two sets of images appears to be due to the individuals within each group.

When the ranges of values were compared between user groups for each measure (time, questions, and words), it can be seen that there is a great overlap between the user groups. Sections C.1 to C.3 illustrate the comparisons for each measure. In general, one user group would have a range large enough that the other user groups would be within the range. The large overlaps show that all of the participants fit within a similar range and the difference in output is due to the individuals within the group rather than the user groups themselves.

Similar to the results of Research Study of Text Alternatives Without the Procedure, the number of questions answered depended on the individuals within each user group. The fact that Developers answered the least questions for one set of images and the most questions for another set of images (rather than consistently answering the most or the least for both sets) indicates that it is the individuals within the user groups who are causing this discrepancy.

Table 6.27 Comparison Between Set A and Set B Images (Overall Averages) When Using the Procedure in Tool Format

	Set A Images	Set B Images
Time Spent	36.5 minutes	23.5 minutes
Words Written	201.1 words	97.7 words
Questions Answered	32.7 questions	27.4 questions

6.4.4 Analysis of Quality of Identified Information

This section discusses the quality of the information that was identified while using the prototype tool by analyzing the amount of information reported as Essential, Significant, Helpful, and Not Important (in the Research Study of the Identified Information). The participants identified an average of 65.7% of the Essential information, 40.0% of the Significant information, 24.7% of the Helpful information, and 43.4% of the Not Important information as rated by the participants in the Research Study of the Identified Information. Identifying 39.4% of the information rated as important is a relatively low amount. However, Table 6.18 (on page 116) showed that it was possible for at least one participant to identify 100% of the Essential information and majority of the Significant and Helpful information when the tool was used. The prototype tool needs to be modified to help people consistently identify all or most of the important information from the image. The modifications suggested in Section 6.4.1 may help accomplish this.

With the prototype tool, the participants were explicitly asked specific questions about the image. They had less freedom to write about whatever they wished in free form. They did not need to think about what type of information to provide. They were asked to about specific types of information and they provided the answers. Sometimes, this limited the details that were identified and the participants used fewer words and less descriptive words. Therefore, some detailed information was often missed when using the tool.

In the previous research studies, the participants had more freedom to identify information that did not have a corresponding question in the set of questions. These pieces of information were not or rarely identified when the participants used the prototype tool. It is likely because the participants were not explicitly asked to provide the information; therefore, they did not think to provide the information. Within the prototype tool, each category and sub-category of questions had a question asking if there was other information they wished to provide but was not previously asked. Participants rarely made use of such opportunities. This shows the

importance of having a set of questions that is comprehensive enough to identify most or all of the important information communicated by an image. It may mean adding more questions in order to identify more details about the image and its components.

IMP 29. Add the question "What are the x and y-axes of the graph or chart?" to the How Much category.

IMP 30. Add the question "What trend does the quantitative information suggest?" to the How Much category.

IMP 31. Add the question "What is the position of the object (facing left, slanted, sideways, etc.)?" to the Perceptual category.

The participants were once again asked to rate the importance level of the information they provided. Some participants seemed to provide information only when they felt it was important, while others provided information even when they felt it was Not Important. This explains the amount of information identified being rated as Not Important in the Research Study of the Identified Information. Since people identified information that they felt was Not Important simply because they were asked, the procedure and set of questions should provide guidance on when best to identify and include certain types of information into text alternatives. This could help decrease the amount of time spent on answering questions where information was Not Important. The tool should also help identify questions that may be Not Important and prevent those questions from being presented so that Not Important information is minimized.

It is possible that the participants thought that the information would be at least Helpful to the users. It was later determined that users had considered the information as Not Important. This shows that the importance levels should consider the points of view of the text alternative writer and as well as the user, which it does. However, the tool needs to help people focus on identifying important information (Essential, Significant, and Helpful) and limit the amount of unimportant information. Further research is necessary to determine how the tool can be better structured to support the focus of information. This may involve determining the sequence of questions to present to best identify the important information.

IMP 32. In the tool, the questions should be presented such that it helps the user focus on the important information while minimizing the unimportant information. Further research is necessary.

Many instances of the important (especially Essential and Significant) information were missed possibly due to two reasons: (1) the image was not broken down into components and (2) the different details that each participant considered important.

Many pieces of information, especially details about a particular component of the image, were missed because the participant did not separate the image into components. They would answer each question with respect to the entire image and did not focus on providing details about a particular component. For example, for a painting or photograph, the participants would answer the Perceptual question regarding colour with respect to the entire image (e.g. "reds and orange") rather than with respect to each object or person in the image (e.g. "red and orange" for leaves, "dark brown" for tree trunks, and "blue" for the sky). It would take more time to answer all of the questions with respect to each individual image component, but the participants would have been able to identify a greater amount of the important information. Many details regarding people's and objects' appearance were missed.

IMP 33. In ISO 20071-11 and the tool, reinforce the importance of separating the image into its components in order to identify details particular to each image component.

Even when the participants do break the image down into components and/or provide a more detailed description of the components, each person may focus on a different aspect and provide different information. For example, one person may describe a dog's breed while another may describe the way that the dog was posed. Both descriptions discuss the dog's appearance but regarding different aspects. Therefore, it is possible to miss information that another person identified, resulting in a lot of information that was missed. It may be difficult to ensure consistency regarding information identified for a single image by different people because each person will consider different information as important and each person may know different information about the image. One possible solution would be to create a question for every possible detail that could be important. For example, instead of simply asking "What is the

appearance of the person?", there could be questions such as "What is the height of the person?" and "What is the person wearing?" Another solution would be to add guidance and suggestions as to the information that users may wish to know. While it is important for the results to be consistent between different people, it is of greater importance that the procedure be capable of identifying the various details that may be important to users, which the results show the procedure can do.

IMP 34. In the Who category, add guidance and suggestions about information that could be included when describing a person's appearance, such as height, clothing, hair style, and jewelry.

The amount of time necessary to identify information about an image was a limiting factor on the amount of information identified by each participant. With each component, there was another set of questions, requiring more time to answer. It is understandable that some people would choose not to break the image down into components due to time constraints. With fewer questions being asked about individual components, the details about each component were often missed by most participants. The procedure should suggest that questions be answered for a component if the response would differ from or is more specific than what has already said about the entire image. For example, if the entire image is a black and white image, it is not necessary to answer the question regarding colour for both the entire image and for each individual component. However, if the image consists of "primary colours", it may be beneficial to specify the particular colours at the component level (such as "blue" for one component and "red" for another).

IMP 35. In ISO 20071-11 and the tool, add guidance for when information applies to all or multiple image components. The guidance should suggest that a question be answered for a larger component if the response applied to the smaller components within the larger component. Otherwise, the question should be answered for the smaller components and skipped for the larger component.

Participants misinterpreting questions also affected the amount of important information identified. For example, stating the actual text in the image instead of the type of information the text represents could have resulted in more important information being identified. Since the participants misinterpreted the question, they may have missed information derived by that question if they had interpreted the question properly.

The prototype tool did not automatically save responses, so participants may have re-entered their responses multiple times. When they re-entered their responses, some information that was entered the first time may have been missed the second time. This may have caused some information from being identified, lowering the percentage of important information that was identified. Automatically saving responses would eliminate this occurrence.

With respect to the user groups, Figure 6.19 (on page 117) shows that the Specialists were the only group to identify above average of Essential, Significant, and Helpful information. The Content Providers identified an exceptionally high percentage of Essential information as well. Developers and Internet Users both identified below average for all levels of importance, with Internet Users identifying the lowest percentages amongst the user groups. This result is not surprising since Developers and Internet Users had spent less time, wrote fewer words, and answered fewer questions compared to Specialists and Content Providers. Therefore, Developers and Internet Users identified less information (and less of the important information). It is believed that if Developers and Internet Users spent as much time on identifying information as Specialists and Content Providers, then they would also identify more of the important information.

6.5 Analysis Across Studies

This section considers and compares the results gathered within the three research studies. It considers the procedure's ability to identify information for text alternatives, the types of information identified, and improvements to be made to the procedure.

6.5.1 Comparison Regarding Time, Words, and Questions Between Research Studies

Figure 6.28, Figure 6.29, and Figure 6.30 summarize the comparison of the results between the three research studies with regards to the amount of time spent, number of words

written, and the number of questions answered. It can be seen across all three figures that there is a consistent increase in all three aspects with each research study.

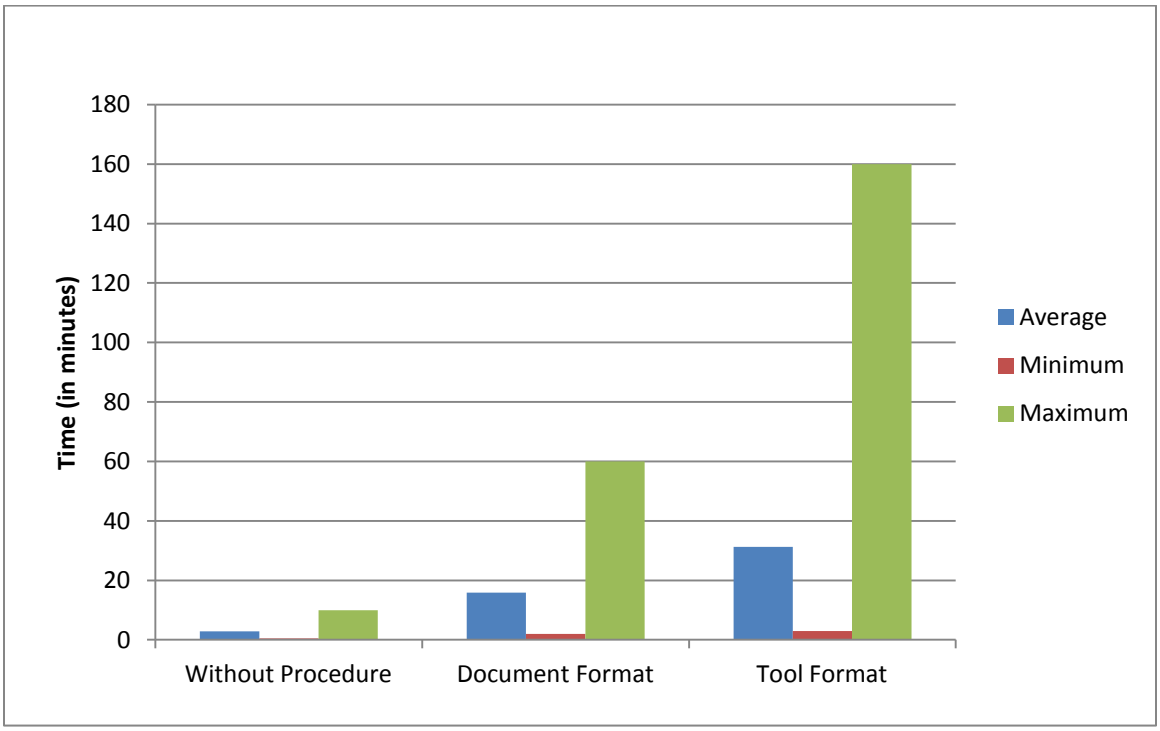


Figure 6.28 Comparison Between Research Studies on the Amount of Time Spent

It is understandable that the amount of time increased when a procedure was introduced. Time was taken to interpret and apply all of the steps of a procedure. It is desirable that the time needed to create text alternatives be approximately the same with or without a procedure but it takes time to create text alternatives that are informative and useful to users. It can be seen that due to the extra time, more words were written, more questions were answered, and more types of information were consistently considered. To improve the current state of text alternatives, it is necessary for more time to be spent on creating informative text alternatives. However, the amount of time spent needs to be kept within a range that people will be willing to spend. This means that optimization of the time taken needs to be a major objective of further developments.

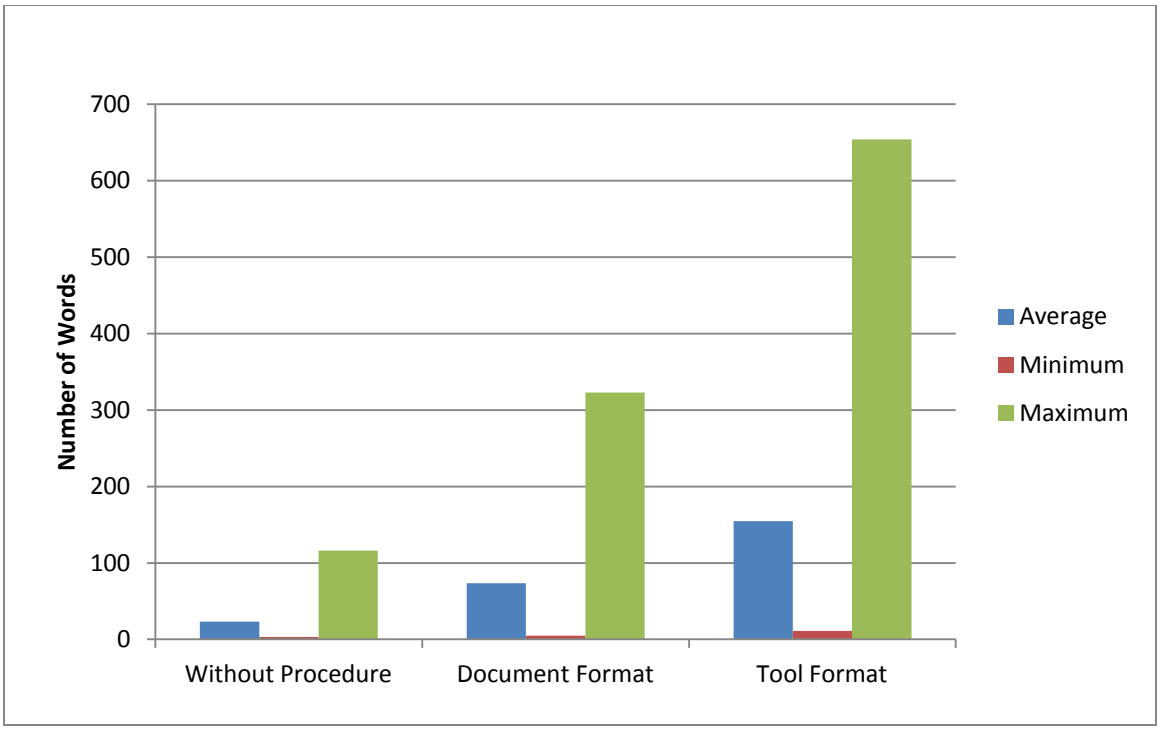


Figure 6.29 Comparison Between Research Studies on the Number of Words Written

While more words were written about each image with each research study, as shown in Figure 6.29, it is not possible to comment on the quality of the information without considering the importance of the information. This is further analyzed in Section 6.5.2, where the importance of the information is considered.

Figure 6.30 shows that the average number of questions answered more than tripled when the prototype tool was used to force users to provide answers. With each research study, the increased number of questions also increased the types of information being identified by participants. When forced to consider every question or category of questions, the participants attempted to provide as much information as they knew regarding each category. This shows that prompting users for responses can help identify more and different types of information about an image.

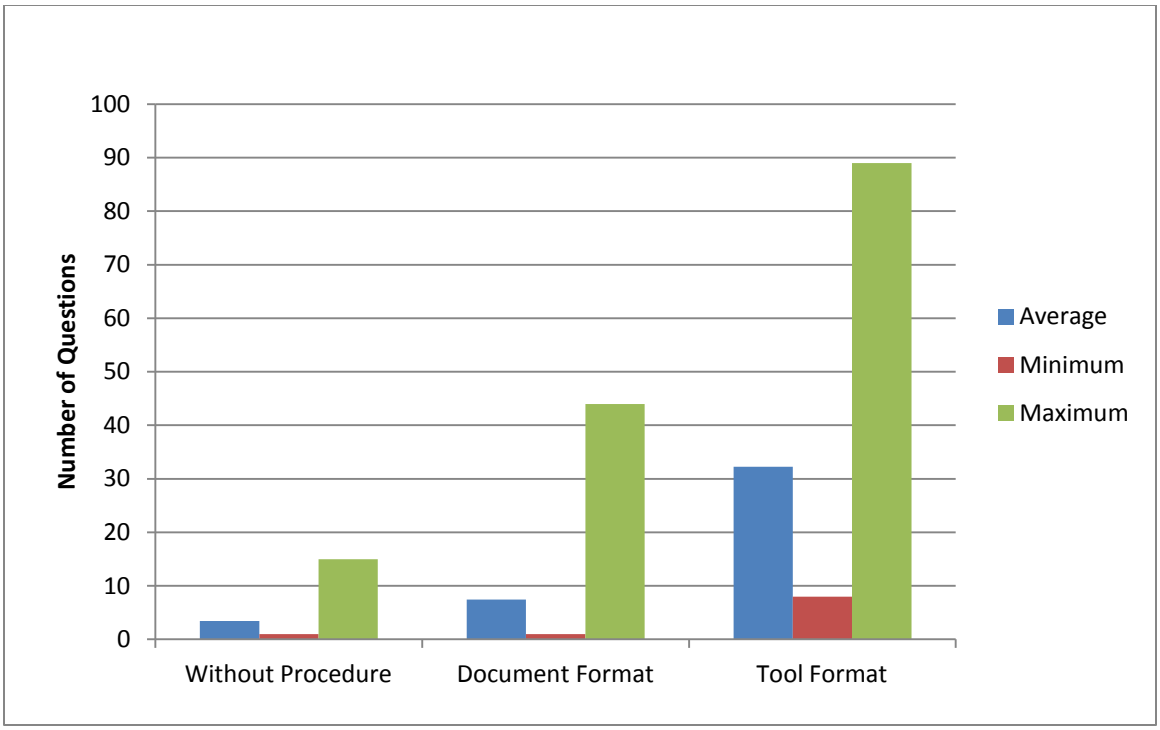


Figure 6.30 Comparison Between Research Studies on the Number of Questions Answered

There was indication that when more time was spent, more words were written and more questions answered. Ideally, the tool could help generate the same level of output in a shorter amount of time. Figure 6.31 compares the output rates in the three research studies. The rate of words written with the use of the procedure was not as good as without the procedure. This is due to the extra time necessary for participants to familiarize themselves with the procedure and the prototype tool. The rate improved slightly using the prototype. The prototype tool doubled the rate of questions answered compared to when the document format was used. This shows that a tool helps to apply the procedure and improves the efficiency of the procedure. The rate of questions answered might be greatly improved if the suggestions described in Section 6.4 were implemented while the amount of time spent on the procedure decreased.

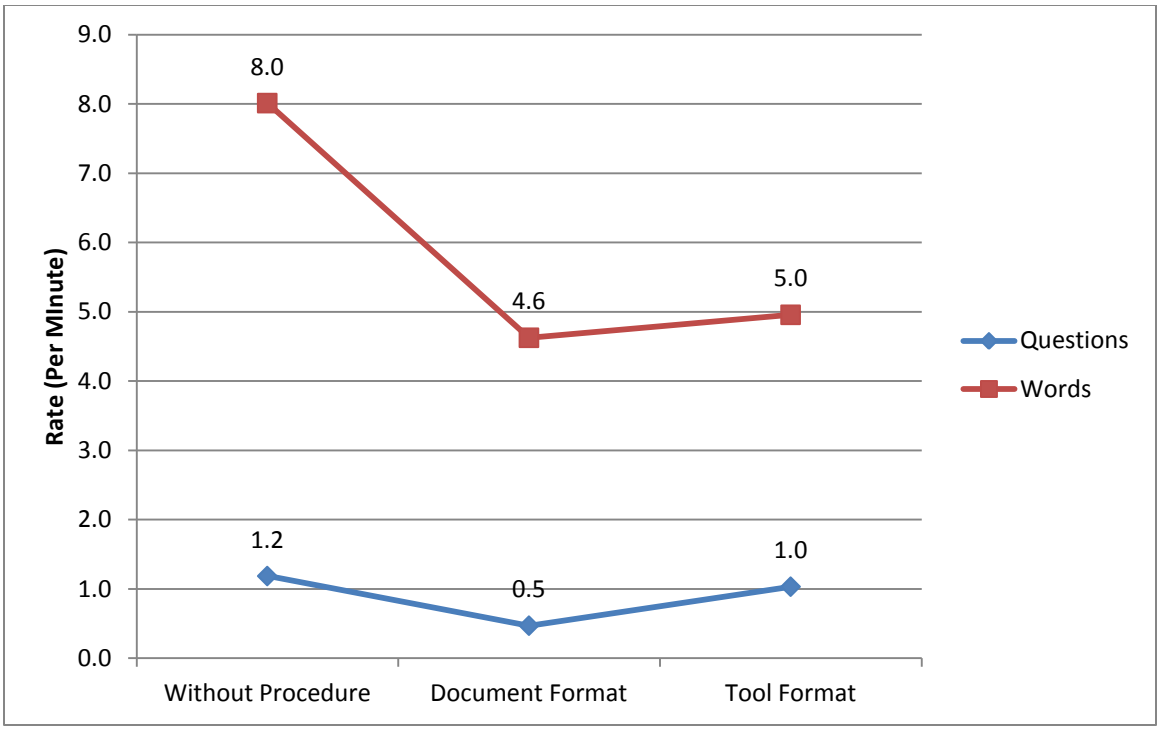


Figure 6.31 Comparison Between Research Studies on the Rate of Output

The types of important information that were more commonly missed and those that were commonly identified were analyzed. It appeared that the same types of information were more often missed with and without the procedure. This was true for Essential, Significant, and Helpful information. Similarly, most of the commonly identified types of information were the same across the research studies. This indicates that people will identify and miss similar types of information regardless of the tools available to them. The difference is in the amount of information that they identify and miss.

For example, without a procedure and when the procedure was in document format, participants identified fewer than five types of Helpful information, while more than five types of Helpful information were identified with the prototype tool. With the prototype tool, participants answered the most questions and wrote the most number of words. This resulted in more types of Helpful information being identified. Therefore, even though all three research studies identified similar types of information, the prototype tool identified more of the important information and information that users want to know about.

6.5.2 Quality of Information Identified in Research Studies

Figure 6.32 summarizes the average percentage of Essential, Significant, Helpful, and Not Important information identified within all three research studies. The importance level ratings were provided by participants in the Research Study of the Identified Information, who represented the users of text alternatives. The percentage of information identified for each importance level increased with each research study. This meant that the quality of the information identified improved with the use of the procedure, even more so when the participants were forced by the prototype tool to consider every question. The participants identified Essential and Significant information that would have otherwise been missed.

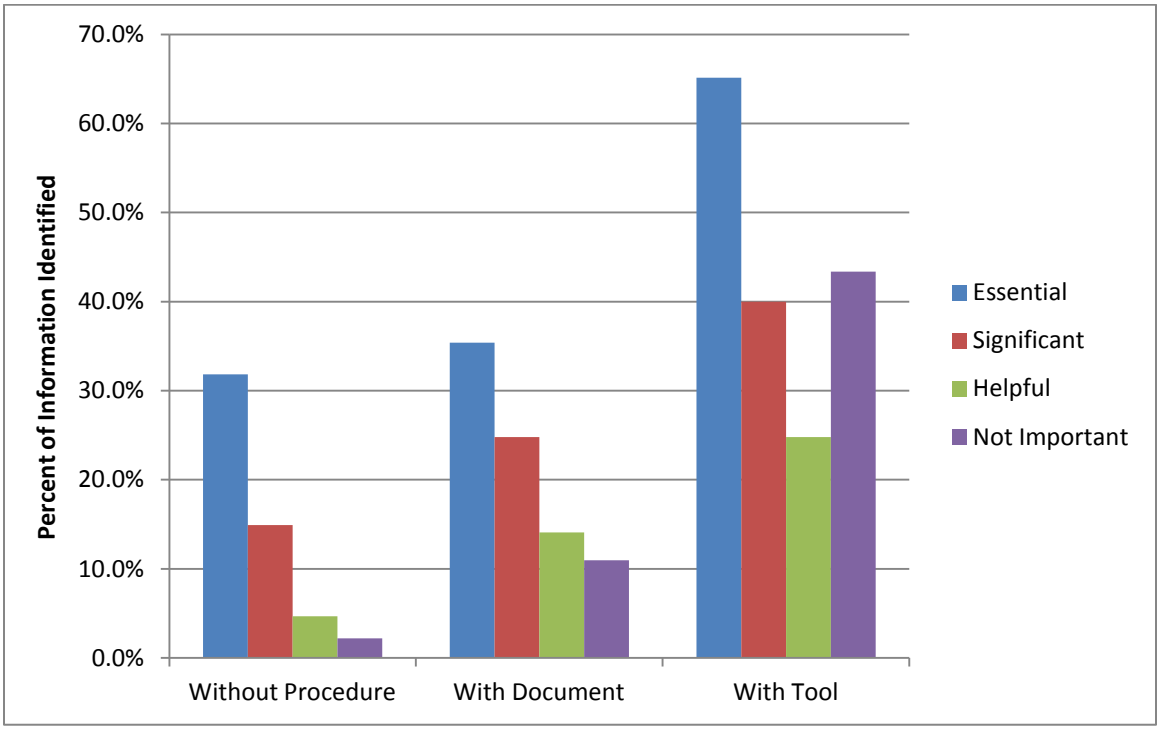


Figure 6.32 Percent of Information Identified in Each Research Study

Conversely, more Not Important information was also identified with each research study. This was especially true of the Research Study of the Procedure as a Prototype Tool. However, the Research Study of the Procedure as a Prototype Tool forced the participants to consider and respond to every question. Therefore, it was understandable that a greater amount of the information identified would be considered as Not Important. This indicates that the procedure and tool can be further improved to minimize the amount of Not Important

information being identified so that the focus can be on identifying information that users would consider as Essential, Significant, or Helpful. Additional guidance on the important information people could focus on is discussed in Section 7.4.4.

The information identified by the participants in the first three research studies was considered important enough for the participants to provide the information. It was rarely the case that the participants would provide information they considered Not Important information even when they were forced to consider every question in the set of questions. With each research study, an increased amount of information was identified. This information was thought to be of some importance. It was only later that the information was discovered to be Not Important to users. This may explain why there was a large increase in information rated as Not Important being identified by the participants once the procedure was used. These results also show the importance of evaluating text alternatives since the information that was thought to be important might not actually be important to users.

When forced to consider every question, there were 292 instances where a higher percentage of participants identified a piece of information than without a procedure and 268 instances where a higher percentage of participants identified a piece of information than when not forced to consider every question. This shows that a lot of information is often missed in text alternatives when people are not required to make a conscious choice regarding the existence of information in an image. Overall, the prototype tool was especially useful in identifying Subjective, Why, Perceptual, Classification, Sequential Relationship, and How information. Table C.41 and Table C.43 in Appendix C list the specific questions that show the greatest improvement.

When compared to the results generated without a procedure, there were 28 instances where a greater percentage of participants identified a piece of information without a procedure than when the prototype tool was used. Majority (58.6%) of these were likely affected by the sample size of the research studies. Even though the same or greater number of participants identified a piece of information when using the prototype tool, the percentage indicated otherwise. Also, in 31.0% of the instances, there was only one additional participant who identified a piece of information without the procedure than with the prototype tool. Both of these cases were likely due to the research study conditions (sample size) and the individual participants, giving the impression that the information was identified better without the

procedure. The remaining instances presented possible improvements and considerations, such as adding a question regarding the placement of objects and make use of the Physical Object questions to answer other questions.

One instance where the prototype did not improve was regarding the spatial relationship between one image component and all other image components in the image. However, the information could also be derived from a question regarding the position or placement of an object. Although a question existed to identify the positioning of a person, a question did not exist regarding objects. This question had since been added to the set of questions.

There were instances where the question "What is the object in the image or image component?" was left blank because the information was identified elsewhere, such as the Classification, Why, Who, or Where questions. There were instances where the reverse had also been true. Because the What question is so versatile, some participants chose to provide the information in one question or the other but not both. It may be beneficial to add guidance so that the information is identified using the more appropriate question. This may prevent the same information from being identified multiples, which saves time, and additional (other) information could be identified about the image.

IMP 36. In ISO 20071-11, add guidance that the type of image (such as painting, flow diagram, and graph) should be provided through the Classification sub-category instead of the Physical Object sub-category.

IMP 37. In ISO 20071-11, add guidance that identifying and describing individual people who appear in an image should be provided through the Who category instead of the Physical Object sub-category, where each person is a separate image component. A summary of all the objects and people depicted in an image could be provided through the Physical Object sub-category.

IMP 38. In ISO 20071-11, add guidance that identifying the location depicted in an image (such as a park) should be provided through the Location/Place sub-category instead of the Physical Object sub-category. The objects that appear at the specified location (such as trees, benches, and paths) should be provided through the Physical Object sub-category

When compared with the results generated from procedure in document format, there were 100 instances where a greater percentage of participants identified a piece of information while using the document than while using the prototype tool. In majority of these instances, there was only a difference of one or two participants identifying the piece of information. The difference is negligible, therefore, it was concluded that these instances were likely due to the individual participants choosing to provide the information or not. Many of the other instances lead to possible modifications or considerations. Some of these instances were due to:

- Questions missing from the set of questions,
- The type or brand of an object,
- The type of information was identified rather than the information itself,
- The image was not broken down into components, and
- The What question was missed since the information was identified through the Classification, Why, Who, or Where questions.

These issues and their resolutions were previously discussed above and in Section 6.4.4.

One new issue is the answering of a question for the entire image versus answering a question per each individual component. This concern was especially obvious for the image containing sequential relationships. Based on the information from individual components (or steps in a sequential relationship), it is possible to generate a summary or general answer for a particular question for describing the entire image. However, the participants who had identified the individual steps in a sequential relationship did not provide such a summary. It was likely believed to be unnecessary until the text alternative was to be written. Similarly, this can also occur for Logical Relationship, where there are questions to break down the relationship into parts.

It appears that in some cases, it may not be necessary to separate these relationships into their parts if the person can build the relationship themselves. Some people may need to break the relationship into its parts to help organize the information, but others may not. The guidance should note that if a complete summary of the relationship(s) is provided, it may not be necessary to answer the questions for the individual parts of the relationship. For example, if the logical relationship was identified as “the girl is walking the dog”, it is not necessary to explicitly identify the subject, object, and interaction of the relationship (for which there are separate questions). It should be noted that this would apply to the individual relationships and not to the

entire image. It is still recommended to break an image down into components in order to identify details that may be otherwise missed.

IMP 39. In ISO 20071-11, add guidance stating that if a complete summary of the relationship(s) is provided, it may not be necessary to answer the questions for the individual parts of the logical and sequential relationships. It should also be noted that this would apply to the individual relationships and not to the entire image.

Table 6.33 compares the quality of information identified by each user group when using the prototype tool against the quality of information when using the document format. All user groups identified a greater amount of information across all levels of importance. The results show that the Specialists benefitted the most from the use of the tool (has the greatest increase in the percentage of information across all of the importance levels) and the Developers benefited the least. Since Developers identified the most amount of important information when using the document, they would have to generate more output than the other user groups in order to show the same level of improvement as the other user groups. The results were further analyzed by comparing the results of Set A and Set B images separately. It was discovered that Specialists showed the greatest improvement for Set A images while the Developers showed the greatest improvement for Set B images. This indicates that Developers benefit from the use of a tool as much as the other user groups. This once again shows the variance between the Group A and Group B participants within the user groups.

Table 6.33 Quality of Information Comparison by User Group

		Overall	Specialists	Developers	Content Providers	Internet Users
Essential Information	Document	35.4%	35.0%	38.6%	36.8%	31.3%
	Tool	65.1%	69.8%	63.0%	83.9%	56.6%
Significant Information	Document	24.8%	25.4%	29.8%	22.0%	21.6%
	Tool	40.0%	54.2%	34.6%	41.5%	32.2%
Helpful Information	Document	14.1%	17.0%	16.0%	16.2%	8.0%
	Tool	24.7%	38.9%	19.8%	28.4%	15.4%
Not Important Information	Document	11.0%	20.1%	5.7%	15.3%	5.0%
	Tool	43.4%	64.1%	33.3%	54.4%	31.9%

Chapter 7 Research Study of the Identified Information

Once the information about the images was collected, it was then important to determine the importance of the information to the users of text alternatives. This in turn determined how well people identified information that users want to know about, with and without the help of the procedure.

This chapter describes the research questions, participants, methodologies, results, and the analysis of results for this research study.

7.1 Research Questions

This research study was designed to answer the following research questions:

1. What important pieces of information were identified as missing from the provided sets of information?
2. How reliable were the levels of importance that were provided by the participants?
3. Was there greater consistencies within specific user groups (such as the visually impaired or people with a technical background)?
4. How important was each of the pieces of information (obtained from the previous studies)?
5. How did the quality of the information identified within each research study compare with each other?

The final step of the procedure is to evaluate the resulting text alternatives. Section 3.1.5 recommended that three evaluations be conducted: one with visually impaired users, one with sighted users without seeing the image, and one with sighted users while seeing the image. This research study focused on the first two evaluations. The purpose of evaluating with sighted users who see the image is to confirm the validity and accuracy of the information presented in text alternatives. While creating the statements used in this research study, the researcher validated the statements to be accurate.

One of the research questions of this research study was to learn the importance of the information so as to identify initial guidance on information that users want or need to know about. The importance of information about an image may differ if the user can see the image. Information may have lower importance if the user can get the information by seeing the image. To evaluate the true level of importance, it was necessary for the participants to not see the images.

7.2 Methodology

7.2.1 Participants

This research study consisted of visual impaired participants (using screen readers) and sighted users (with images being unavailable). While text alternatives may be used by sighted users who can see the image, the main intent is to be used by those who cannot. Therefore, it was important to have participants who were visually impaired and made use of screen readers. The research study was designed such that it was accessible to both the sighted and the visually impaired.

Unlike the previous research studies, there was no requirement that the participants be developers, content providers, or usability specialists. Text alternatives are meant to be usable by everyone with and without disabilities or expertise knowledge. Therefore, the information about images should be evaluated by people with different abilities and knowledge.

The participants were once again separated into two groups. One group evaluated the images from Set A and another group evaluated the images from Set B. The groups were selected at random while balancing the number of visually impaired participants in each group.

At the conclusion of this research study, there were 42 participants, 9 of which were visually impaired (and made use of screen readers) and 33 were sighted. Of the 42 participants, 18 participants were from an IT or technology background while 24 participants were from non-technical backgrounds. Given the small sample, the participants may not be true representatives of their population and varying demographics was not possible. However, the results can identify possible suggestions and recommendations that may be true of the population.

7.2.2 Materials and Execution

The research study was conducted using an online Web site. Each participant was given login information to access the research study. The research study was designed to be completed without supervision and at the participants' convenience and pace. This allowed the participants to complete the research study over several sittings at their convenience rather forcing them to complete it in one sitting.

The online Web site was designed such that both the visually impaired and sighted users received the same materials for evaluation, which were:

- Instructions regarding the tasks they were to perform. The instructions included a description of the four importance levels (as described in Section 3.1.6) as well as the layout of the research study Web site. A copy of the instructions were provided to the participants in a document as well as presented as the first screen of the evaluation tool. The document can be found in Section B.4 in Appendix B.
- A copy of the original Web pages that contained the images used in the research study, modified such that most of the images and graphical content were not visible. This provided participants with the context of the images.
- A link that led to the image being evaluated within the modified copy of the original Web sites. The `alt` attribute of the each image being evaluated stated "USERLab is evaluating this image."
- A set of sentences about each image.

The participants were instructed to consider the information expressed by the sentences rather than the phrasing and wording of the information. The intent of this research study was to determine the importance of the information rather than the length or presentation of the information. The important information can be worded or phrased to suit any length requirements and the flow of the document.

Visually impaired users used screen readers for the research study while sighted users were not required to use screen readers. Both user groups were presented the content in the original Web page without seeing the images on the Web page. The participants were encouraged to explore the original Web page to familiarize themselves with the content and context surrounding the image. This is illustrated in Figure 7.1.

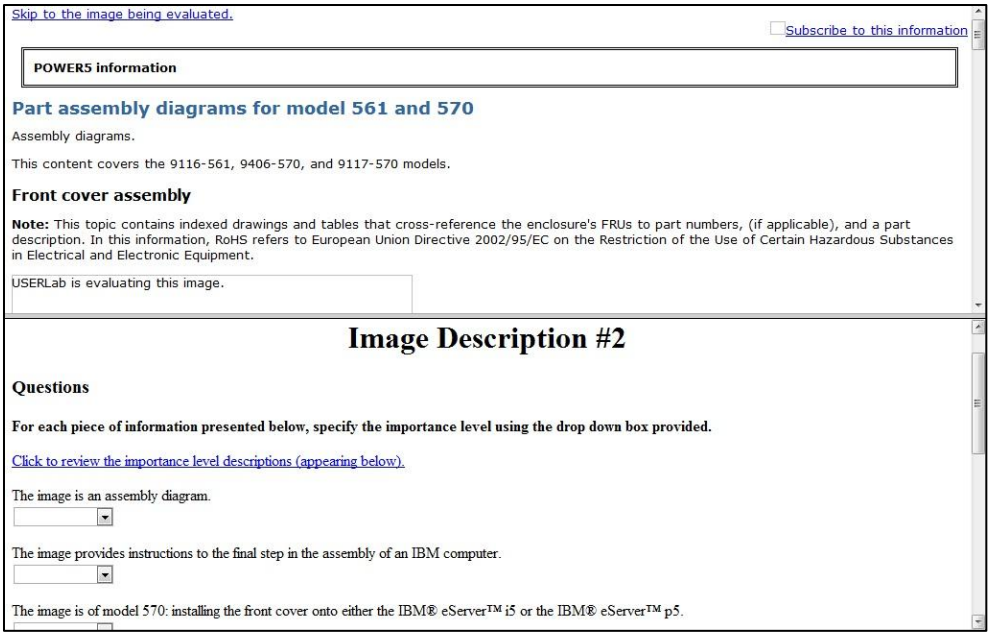


Figure 7.1 Screen Shot of Evaluation Tool Layout

Each Web page presented to the participants contained one image that was being evaluated. This image had the text alternative "USERLab is evaluating this image." The participants were then given a set of sentences about that particular image. This was done for a total of five Web pages and images.

The sets of sentences about each image were developed from the output of the first three research studies. One or more sentences were written based on the pieces of information extracted from each question within the set of questions. When there was a conflict between the information, the sentences attempted to encompass what was written by the majority of the people. In some instances, a sentence encompassed several pieces of information (sometimes from different questions) in an attempt to minimize the number of sentences that the participants had to evaluate. For example, instead of "It was a sunny day." and "It was autumn.", the sentence "It was a sunny day during autumn." was used. The sentences usually combined information from within the same category or sub-category of questions. The participants were asked to rate the importance level of each of these sentences. The source(s) of the information (the questions) for each sentence were recorded so that the results could be traced back to the source.

After reviewing all of the sentences and information that was presented about each image, the participants were asked to identify any information that they would have liked to

know or learn about the image but the information was not provided to them. They were asked to do this without having seen the image.

After the participants provided ratings to the sentences for their set of images, they were asked to share any comments, concerns, or feedback they had about the research study, the research project, the evaluation tool and environment, and general comments to the researcher.

7.2.3 Method to Answer Research Questions

This section describes the methods used to extract answers to the research questions from the output generated by the participants. The method to answer each research question will be discussed separately.

Research Question 1

For the set of information that the participants said they wanted to know but was missing from the sentences presented (research question 1), each piece of information was analyzed by the researcher using the following questions:

1. Can this information be found on the Web page content or linked Web pages?
2. Can this information be identified using the existing set of questions?
    a. If so, why was it not identified and presented?
    b. If not, should a new question be added to the set of questions?

The results of this process determined the changes to be made to the procedure, if necessary. The results are summarized in Section 7.3.1.

Research Question 2

Before considering the importance of the information, it was necessary to consider the reliability of the results (research question 2). An overall importance rating needs to be calculated for each piece of information in order to determine the quality of the information that was identified by each of the previous research studies. The overall importance rating for a question would be considered reliable if most of the participants agree with that rating or its neighbouring rating(s).

The mode and median was calculated for each piece of information presented to the participants. The percentage of participants who responded the same as the mode and median was calculated, along with the percentage of participants who responded with neighbouring values of the mode and median. It is unreasonable to expect the participants to provide the same rating for a piece of information in order for the rating to be considered reliable. Therefore, it was decided that the neighbouring values be considered as well when determining the reliability.

These percentages were analyzed to determine the reliability and consistency of the responses between participants. The instances where more than 70% of the participants selected the mode (or median) or its neighbouring values were considered as reliable since most participants selected similar values. The instances where less than 70% of the participants selected the mode (or median) or its neighbouring values were further analyzed to identify trends. Instances where all of the participants agreed with the mode (or median) or its neighbours were analyzed to identify additional trends. The average percentage where more than 50% of the participants selected the mode or median was also determined.

The results to this research question are summarized in Section 7.3.2.

Research Question 3

In addition to calculating the reliability and consistency of results with respect to all participants, consistency was also calculated with respect to the group of only the visually impaired and to the group of only people with technical backgrounds (research question 3). The mode and median were calculated based on the responses of the participants within the group and the same calculations were done for each group.

The percentages calculated for each group were compared to the percentages calculated for all participants in the research study to determine if there was more consistency in responses for a particular group. If the percentages improved, it showed that the specific group had a more consistent view of what information was important, which was different from the general public, and required special attention to ensure the procedure also supported their needs. If the results did not improve, it showed that the specific group did not have a consistent view of what was important and it would be difficult to identify the particular needs of that group at this time.

The results to this research question are summarized in Section 7.3.2.

Research Question 4

An overall importance rating for a piece of information or question was determined (research question 4) based on the mode. If there was a tie for the mode (for example, the same number of people rated a piece of information as Essential and Significant), the median was used to determine the overall rating. This rating was then mapped back to each question that extracted the information.

The total number of Essential, Significant, Helpful, and Not Importance information was tallied to calculate the percentage of Essential, Significant, Helpful, and Not Important information identified by each of the previous research studies. The number of times that a question was considered as Essential, Significant, Helpful, and Not Important was also tallied to analyze the importance of that question generally. This was similarly done for each sub-category and category of questions. These results were analyzed to possibly suggest guidance regarding when certain information may be of higher importance and should be considered and when certain information may be of lower importance.

The results to this research question are summarized in Section 7.3.3.

Research Question 5

To determine the quality of the input from the previous research studies (research question 5), the number and percentage of people who identified a piece of information (or answered a question that arrived at the information) was tallied for each piece of information for each image for each research study. These values were compared between research studies to determine when and if there was improvement in identifying a piece of information when different materials were provided to the participants. Each instance that did not show improvement was analyzed for a possible reason and way of improving the procedure.

The percent of Essential, Significant, Helpful, and Not Important information identified by each participant was also calculated and averaged. These values were then compared between research studies to determine if there were improvements when the procedure was used. These values were also compared between user groups within a research study and between research studies.

The types of information that participants tended to identify and tended to miss were determined for each research study and for each importance level. This was done by first identifying the pieces of information identified by majority (more than 50%) of the participants in each research studies. Then the importance level and type of information was determined for the pieces of information that was identified by the majority. The types of information were then arranged with respect to each importance level. These types of information were compared between the research studies to determine if it would change when the participants had different materials available to them.

The results to this research question were summarized in Sections 4.3.4 (without the procedure), 5.3.5 (procedure in document format), and 6.3.6 (procedure in tool format).

7.3 Research Study Results

7.3.1 Missing Information

Participants were asked to identify any information that they wished to know about an image but was not provided within the set of sentences. Participants identified a total of 49 different pieces of information as missing about the ten images, 7 of which were identified by more than one participant. The minimum pieces of information reported as missing for any given image was 2 and the maximum was 13. Of these 49 pieces of information, 36.7% were rated as Essential, 55.1% as Significant, and 8.2% as Helpful.

The 49 pieces of information were analyzed to determine how the procedure could be modified in order to include the information that was missed. Of these 49 pieces of information, 36.7% was already identified within the information that was provided to the participants or it was information that was provided as part of the Web site text and context; 55.1% can be already identified as by the existing set of questions; 8.1% are new pieces of information and may require new questions to be developed.

7.3.2 Reliability of Results

In order to determine the reliability or consistency of a participant's responses with the rest of the group, the percentages of participants who responded the same as the mode (or median) or its neighbouring values were calculated. Overall, 46.8% of the instances had more than half of the participants agree with the mode and 35.3% of the instances where they agreed

with the median. Table 7.2 shows the frequency that less than 25%, 25% to 39%, 40% to 59%, and more than 60% of the participants agreed with the mode or median for a piece of information. Table 7.3 shows the frequency that less than 70%, 70% to 79%, 80% to 89%, and over 90% of the participants selected a rating equal to or neighbours of the mode or median.

Table 7.2 Frequency that Participants Agreed with the Mode or Median

	< 25%	**25% - 39%**	**40% - 59%**	**≥ 60%**
Mode	0.0%	21.1%	56.4%	22.5%
Median	7.8%	32.1%	45.9%	14.2%

Table 7.3 Frequency that Participants Agreed with the Neighbours of the Mode or Median

	< 70%	**70% - 79%**	**80% - 89%**	**≥ 90%**
Mode Neighbour(s)	12.4%	17.4%	23.9%	46.3%
Median Neighbour(s)	1.8%	11.5%	31.7%	55.0%

For the visually impaired participants, 94.5% of the instances had more than half of the participants agree with the mode and 77.5% of the instances where participants agreed with the median. Table 7.4 shows the frequency that the participants agreed with the mode or median for a piece of information. Table 7.5 shows the frequency that the participants selected a rating equal to or neighbours of the mode or median.

Table 7.4 Frequency that Visually Impaired Participants Agreed with the Mode or Median

	< 25%	**25% - 39%**	**40% - 59%**	**≥ 60%**
Mode	0.5%	0.0%	13.3%	86.2%
Median	5.5%	4.6%	17.4%	72.5%

Table 7.5 Frequency that Visually Impaired Participants Agreed with the Neighbours of the Mode or Median

	< 70%	**70% - 79%**	**80% - 89%**	**≥ 90%**
Mode Neighbour(s)	7.3%	6.9%	16.1%	69.7%
Median Neighbour(s)	7.3%	5.0%	17.4%	70.2%

For the sighted participants, 45.0% of the instances had more than half of the participants agree with the mode and 35.8% of the instances where participants agreed with the median. Table 7.6 shows the frequency that the participants agreed with the mode or median for a piece of information. Table 7.7 shows the frequency that the participants selected a rating equal to or neighbours of the mode or median.

Table 7.6 Frequency that Sighted Participants Agreed with the Mode or Median

	< 25%	25% - 39%	40% - 59%	≥ 60%
Mode	0.4%	18.3%	58.7%	22.5%
Median	8.3%	30.3%	48.2%	13.3%

Table 7.7 Frequency that Sighted Participants Agreed with the Neighbours of the Mode or Median

	< 70%	70% - 79%	80% - 89%	≥ 90%
Mode Neighbour(s)	19.3%	15.6%	37.2%	28.0%
Median Neighbour(s)	5.1%	9.7%	45.2%	40.1%

For the technology industry participants, 85.8% of the instances had more than half of the participants agree with the mode and 83.0% of the instances where participants agreed with the median. Table 7.8 shows the frequency that the participants agreed with the mode or median for a piece of information. Table 7.9 shows the frequency that the participants selected a rating equal to or neighbours of the mode or median.

Table 7.8 Frequency that Technology Participants Agreed with the Mode or Median

	< 25%	25% - 39%	40% - 59%	≥ 60%
Mode	0.0%	4.1%	38.5%	57.3%
Median	1.4%	9.2%	23.9%	65.6%

Table 7.9 Frequency that Technology Participants Agreed with the Neighbours of the Mode or Median

	< 70%	70% - 79%	80% - 89%	≥ 90%
Mode Neighbour(s)	6.9%	12.4%	22.5%	58.3%
Median Neighbour(s)	1.4%	5.0%	18.3%	75.2%

For the non-technology industry participants, 50.1% of the instances had more than half of the participants agree with the mode and 42.7% of the instances where participants agreed with the median. Table 7.10 shows the frequency that the participants agreed with the mode or median for a piece of information. Table 7.11 shows the frequency that the participants selected a rating equal to or neighbours of the mode or median.

Table 7.10 Frequency that Non-Technology Participants Agreed with the Mode or Median

	< 25%	25% - 39%	40% - 59%	≥ 60%
Mode	0.5%	17.0%	55.5%	27.1%
Median	7.8%	31.2%	40.8%	20.2%

Table 7.11 Frequency that Non-Technology Participants Agreed with the Neighbours of the Mode or Median

	< 70%	70% - 79%	80% - 89%	≥ 90%
Mode Neighbour(s)	18.8%	15.6%	23.4%	42.2%
Median Neighbour(s)	9.2%	8.7%	25.7%	56.4%

7.3.3 Overall Importance of Information

Of the 395 pieces of information presented in the set of sentences, the participants rated 94 pieces as Essential, 84 pieces as Significant, 167 pieces as Helpful, and 50 pieces as Not Important. These were further broken down by categories and sub-categories of questions (Table 7.12) and individual questions. The results for each research study can be found in Sections 4.3.4 (without the procedure), 5.3.5 (procedure in document format), and 6.3.6 (procedure in tool format).

Table 7.12 Overall Instances (and Percentages) of Importance of Information by Category or Sub-category

	Essential	Significant	Helpful	Not Important	Total
Why	8 (61.5%)	0 (0.0%)	2 (15.4%)	3 (23.1%)	13
What – Physical Object	17 (37.0%)	16 (24.8%)	11(23.9%)	2 (4.3%)	46
What – Classification	9 (90.0%)	1 (10.0%)	0 (0.0%)	0 (0.0%)	10
What – Textual	6 (40.0%)	5 (33.3%)	3 (20.0%)	1 (6.7%)	15
What – Perceptual	20 18.3%)	14 (12.8%)	57 (52.3%)	18 (16.5%)	109
What – Subjective	7 (15.2%)	9 (19.6%)	30 (65.2%)	0 (0.0%)	46
What – Logical Relationships	2 (33.3%)	3 (50.0%)	1 (16.7%)	0 (0.0%)	6
Who	3 (18.8%)	4 (25.0%)	9 (56.3%)	0 (0.0%)	16
Where – Location/Place	2 (33.3%)	2 (33.3%)	2 (33.3%)	0 (0.0%)	6
Where – Spatial Relationships	2 (2.8%)	9 (12.7%)	34 (47.9%)	26 (36.6%)	71
When – General	1 (16.7%)	4 (66.7%)	1 (16.7%)	0 (0.0%)	6
When – Sequential Relationships	14 (63.6%)	1 (4.5%)	7 (31.8%)	0 (0.0%)	22
How Much	1 (5.9%)	13 (76.5%)	3 (17.6%)	0 (0.0%)	17
How	2 (16.7%)	3 (25.0%)	7 (58.3%)	0 (0.0%)	12

7.4 Analysis of Research Study Results

This section analyzes the results from the Research Study of the Identified Information to identify information that may have been missed by the set of questions, to identify improvements

to the set of questions, and to identify trends in the importance of the different types of information. The results were analyzed overall, between user groups, between types of information, and between individual questions.

Possible improvements and modifications to the procedure and guidance that are identified throughout this section will be denoted using "IMP" followed by a number. For example, "IMP 1" and "IMP 2". Majority of these improvements have not been implemented because they were identified after the publication of the latest version of ISO 20071-11.

7.4.1 Missing Information Analysis

Participants identified 49 pieces of information that they wanted to know about but was not provided to them within the research study. The pieces of information were placed under three categories: information already provided in the Web site content, information that can be identified by the set of questions, and information that cannot be identified by the set of questions.

General text alternative guidelines (such as WCAG [67]) state that information from a document's textual content should not be repeated. However, 36.7% of the information identified as missing could be found on the Web page or other parts of the original Web site. The participants' responses indicate that they want the information stated again in the text alternative. Part of the reason for this may be is that they did not fully explore the Web page and consider the content of the Web page.

The majority of the "missing" information could be identified using the existing set of questions. Most of the pieces of information identified as missing were concerned with additional details about the appearance of people, presence or absence of objects, and location and time that the image was taken. This type of information mainly had a rating of Significant or Helpful. Some of the requested information can only be known by the owner or creator of the image, while others were details that could have been identified if additional examples were provided for consideration. Questions could be added to identify specific details about a person's appearance. However, a large number of questions would be necessary to fully describe the appearance of a person. Rather than adding questions to the set of questions, additional examples or suggestions of details to describe may be more beneficial.

IMP 40. In the set of questions, add more examples of possible detailed or specific responses for Location/Place sub-category and When category.

The presence or absence of objects is a difficult problem to resolve. Unless explicitly told that certain objects and information is not available, users will not know if the objects and information are actually not part of the image or if the objects and information exist in the image but the user was not informed about it. However, it is not possible to provide a full listing of all objects or information that is not part of the image. Instead, the guidance should encourage the practice of explicitly stating that information is purposely being omitted or cannot be disclosed. While this will not inform users of when objects and information are not part of the image, it can let them know that a decision was made about the information that they should know and that information has been provided.

IMP 41. In ISO 20071-11, add guidance encouraging the practice of explicitly stating (in text alternatives) that information is purposely being omitted or cannot be disclosed.

Of the 8.1% of information identified as missing and was not considered by the existing set of questions, some instances would benefit from being added to the set of questions. This information pertained to reporting about changes to the current state of the Web page due to interacting with an image, that is, the How category. For example, indicating static or dynamic content on the Web page or pop-up windows as a result of an action performed on the image. While creating the tool for conducting this research study, it was recommended by a screen reader user to clearly explain changes to the Web page content prior to performing an action. This was so that the user would know the consequences of performing an action. The same information should be described when an image is involved.

IMP 42. In the How category, add the question "Does a new window or application appear as a result of the interaction? If so, which? For example, interacting with the image will result in a pop-up window of the product's Web site." This question further identifies a possible outcome of interacting with an image, including changes to the computing system resulting in additional windows to appear.

IMP 43. In the How category, add a note stating "If the computer system's focus would change as a result of interacting with an image, the change should be made aware to the user. For example, an e-mail composition window will appear and will have focus." Users may not be aware that the computer system's focus has changed and may continue interacting as if the focus has not changed. The user should be informed of such unexpected changes.

IMP 44. In the How category, add the question "How does the interaction with the image change the current document? For example, the text size increases or additional explanatory text appears." Also, add a note to indicate where in the document that these changes occur. Interacting with an image could result in modifying the document that the user is viewing. These changes may occur at the section of the document the user is viewing or may occur in a different part of the document. Users should be made aware of such changes so that they know what to expect as well as where to locate the changes.

Some of the other instances of information not covered by the set of questions, however, would not benefit from adding a question to the set of questions; specifically, the information with regards to specific details about objects or products portrayed in the image. For example, image B5 contains different models of iPods and some participants wanted information on the product availability and features. Such types of information would be better supported by providing a link to the official product Web site for information. The procedure should include guidance on situations when providing links for additional details may be beneficial.

IMP 45. In ISO 20071-11, add guidance that text alternatives could reference other documents for information and link to such references, also indicating the types of information to expect from the reference. For example, product information.

Majority of the "missing" information rated as Essential and Significant were already considered by the existing set of questions. The new questions to be added regarding dynamic changes to the document would improve the set of questions such that all information identified as Essential and Significant are identified. All of the information that was considered as Helpful was already part of the existing set of questions.

Since majority of the "missing" information was already considered by the set of questions, it shows that the set of questions is comprehensive enough to identify information regarding a wide range of images.

7.4.2 Reliability of Results Analysis

Less than half of the instances involved majority of the participants agreeing with the mode or median (46.8% and 35.3% respectively). However, once the neighbouring values were also included, all of the instances had majority of the participants' agreement. This indicates that the results are sufficiently reliable for use within these studies.

When the neighbouring values were considered as well, there were 12.4% of instances where there was less than 70% agreement with the mode and its neighbour(s). There were many instances where an importance level was selected almost as frequently as the mode and that importance level was not a neighbouring value of the mode. This explained most of the 12.4% of instances that had less than 70% agreement. This occurred most frequently with Spatial Relationship information and Perceptual information. This indicates that there is greater variance on the importance of these types of information for a given instance.

Another explanation for the variance may be because the participants did not know the context or purpose with which to evaluate the importance of the information. When browsing the Internet, a person may arrive at a Web page that contains images. They arrive at that Web page for a particular reason and they may be looking for specific information. However, during the evaluations, the participants were asked to evaluate the information without having a reason for arriving at the Web page and may not have particular interest in the Web page content. B1 and B5 may have had less variance in importance because the image's intent and purpose (a logo and an advertisement) was clearer to the participants than the other images.

Twelve point four percent of the instances had 100% agreement when the neighbouring values were included. There was indication that Subjective and Who information tended to be either Significant or Helpful. The guidance should make note of this to guide people regarding the importance of these types of information.

IMP 46. Add a note in the guidance for Subjective and Who information that they tend to be considered Significant or Helpful.

7.4.3 Consistency of Results Within User Groups Analysis

Certain groups of participants' responses were considered to see if responses were more consistent within the group, indicating that there may be information that the group consider as more important or as less important. This could lead to additional guidance for information to include in text alternatives. The percentages for sighted participants and non-technology participants did not show much improvement. This indicates that the guidance for information that is important to all users would be sufficient for these particular groups. The percentages for the visually impaired participants and technology participants had doubled, indicating that there may be additional guidance needed for these particular groups.

The visually impaired participants and the technology industry participants showed greater consistency in their responses for information within the Why category, Textual sub-category, Physical Objects sub-category, Perceptual sub-category, and Spatial Relationship sub-category.

Information explaining the image's purpose (Why) was fairly consistently rated as Essential and Significant by the visually impaired participants. This shows that the reason the image is presented is very important to these users. The guidance should reflect that this type of information should be given to users the majority or all of the time.

IMP 47. For the Why category, add guidance that Why information is considered Essential or Significant by visually impaired users, therefore, it should be provided in text alternatives the majority of the time.

Textual information was always considered at least Helpful, with it being rated as Essential or Significant majority of the time. This indicates that textual information is always of some importance to users. The guidance, therefore, should recommend that this type of information almost always be given to users through text alternatives.

IMP 48. For the Textual sub-category, add guidance that textual information is likely considered Essential or Significant when the text is the main focus of the image.

IMP 49. For the Textual sub-category, add guidance that textual information is likely considered Helpful when the text is not the main focus of the image.

While there was greater consistency with respect to Perceptual information for the visually impaired users, the ratings were evenly spread across the importance levels. There did not appear to be a consistent trend with the ratings such that additional guidance can be provided. The same was true for Physical Objects and Spatial Relationship information.

7.4.4 Overall Importance of Information Analysis

The importance ratings from this research study can give insight into the importance of certain types of information in general. It must be kept in mind that the importance ratings are specific to the context and purpose of the image. Therefore, the importance level for a type of information may change with the context and purpose. The discussion here is a generalization to improve the usability of the procedure, increase the identification of important information, and provide better guidance as to when certain types of information is beneficial to users.

Table 7.12 showed that some types of information were generally considered more important than others to users. To improve the quality of text alternatives, the procedure should provide additional guidance and suggestions regarding situations when certain types of information may be of more importance and when it may be of less importance.

The information regarding Why was most often considered as Essential. The purpose of the image generally should always be provided to users. The visual appeal of a decorative image or painting tended to vary in importance. This is expected since visual appeal is subjective to each user and context. Even though existing text alternative guidelines recommend that no text alternative be provided for decorative images, the guidance for the procedure presented in this thesis should recommend that this information be provided so that users can at least know the impression and feeling that the decorative image is attempting to achieve.

IMP 50. For the Why category, add guidance that the purpose of the image (Why) is generally considered Essential and should be provided to users.

Textual content in images have generally been considered at least Helpful, but mainly Essential. Text was considered less important when it was in the background and it could not be read clearly. Therefore, the guidance should recommend that text within the image be presented in primary alternative text whenever the text is in the foreground and is the focus of the image.

IMP 51. For the Textual sub-category, add guidance that textual information be provided as stated in the image through the primary alternative text when the text is in the foreground and is the focus of the image.

When the caption of an image was identified, users considered the information as Not Important. This was expected because the information was identified so that it would not be repeated in text alternatives. Step 5 of the procedure (which is organizing the information into text alternatives) states that redundant information be removed. This means not only redundant information that was identified, but also information that is already part of the document's textual content. This should be discussed further in the guidance.

IMP 52. In ISO 20071-11, add a discussion regarding information that has been provided through other parts or sections of the document should not be repeated in text alternatives. Identifying the information in the image caption helps to specify the information not to include in text alternatives.

The image classification or image type was consistently rated as Essential. It is likely because the image type can prepare and inform the user of the information presented by the image. Since this information was never rated below Significant, it indicates that users always want to know this information and the guidance should reflect this.

IMP 53. For the Classification sub-category, add guidance that the image classification is most often considered Essential or Significant, therefore, it should always be provided to users in text alternatives.

The Physical Objects information was generally considered as of some importance (at least Helpful). More specific information about the object's brand and type were considered of higher importance. This supports the guidance that specific information and details be identified whenever possible instead of identifying generic information (Section 3.2.1). Although an object

can be described in both generic and specific terms, specific terms and information can give users more details.

IMP 54. For the Physical Objects sub-category, add a note that the more specific the information about an object, the higher the importance of the information. More generic information tends to be of lower importance.

Within the Perceptual sub-category, the size (or relative size) of an object was most often considered as Not Important by users. The highest rating it received was Helpful. Even though the size of objects was thought to be of importance, users had a different opinion. This question should be considered for removal. However, this information was reported as "missing" by a participant and this information was rated as Significant, indicating that the size can be of importance to users. Therefore, this question should remain, but the guidance should recommend that the question be used less often.

IMP 55. For the Perceptual sub-category, add note that the size or relative size of an object tends to be considered as Helpful or Not Important and should rarely be provided to users.

The colour of an object was rated as each of the levels of importance approximately the same number of times. Due to the high frequency of identifying information regarding colour, a high number of this information considered as Not Important. Upon closer inspection, it appeared that colour was frequently rated as Not Important for technical diagrams since the image was black and white and colour did not have particular significance. The procedure should add guidance that colour should not be considered for technical diagrams and images unless the colour had meaning.

IMP 56. For the Perceptual sub-category, add guidance that colour be omitted from the text alternatives of technical diagrams and images if colour does not have significance or meaning.

The way that objects (and people) look was most often rated as Helpful. A lot of information could be collected about the appearance of objects and people, ranging from the general shape and look to specific details about hair style, clothing, and pose or position. The times when information about appearance was rated higher than Helpful were when the appearance had meaning. For example, a diamond shape represented a decision in a flow chart or the colour white represented innocence. Therefore, the guidance should indicate the importance of describing appearance when there is a special meaning to the appearance.

IMP 57. For the Subjective sub-category, add guidance that the meaning or symbolism of an object's or person's appearance should be communicated to users.

The Subjective information was most often considered as Helpful. This is most likely because the information is one person's opinion or interpretation. While it was helpful to have, Subjective information was considered to be less important for users to know about. The guidance should recommend that this type of information be identified later so that more time can be spent on the more important information.

IMP 58. For the Subjective sub-category, add guidance that Subjective information tends to be considered as Helpful and could be identified later, allowing other more important information be identified first.

Within the Logical Relationships sub-category, the participants were more interested in the final outcome of the action and how the action was performed rather than the subject and object involved. In most cases, identifying the outcome and how the action was performed also identified the subject and object of the action. This suggests that the questions regarding the subject and object of an action or relationship should be considered for removal. However, there may be times when the action or relationship will be identified without the subject or object. Therefore, there should be questions to explicitly identify them. While the questions should not be removed, there should be guidance that it would not be necessary to explicitly answer the questions if the subject and object of the interaction are identified in the interaction itself.

IMP 59. For the Logical Relationships sub-category, add a note that it may not be necessary to explicitly identify the subject and object of the action if they are already identified through the question "What interaction is taking place?"

Within the Who category, identifying the person and their actions were considered to be Essential and Significant. However, the way that the person looks (facial expression, appearance, and position) were mainly considered as Helpful. Therefore, the guidance should recommend that the person and their actions always be identified and provided to users, while the remaining questions can be answered sparingly. If the person's appearance is to be described, it should be done as fully as possible since much of the information that participants identified as "missing" were due to missing details regarding appearance.

IMP 60. For the Who category, add guidance that if a person's appearance is to be described, it should be done as fully as possible so as to minimize the amount of information users will consider as missing.

Within the Location/Place sub-category, the information was of less importance when only a general location (for example, "a room" or "a park") was mentioned instead of an exact location (for example, "Sydney"). Therefore, the guidance should note that the place or location depicted in the image is of greater importance when a specific location is known.

IMP 61. For the Location/Place sub-category, add a note that the more specific the information about the location or place, the higher the importance of the information. More generic information tends to be of lower importance.

Spatial Relationship information was often rated as Helpful or Not Important. One participant had commented that it is often not necessary or useful to know exactly where certain objects were in an image since they cannot see the image. It was more helpful to identify locations of objects relative to other objects in order to get a better idea of the relationships between objects. This suggests that perhaps the question regarding the absolute location of an object should be removed.

Even though the question specifying the absolute location of objects might not always be useful, there are times when it can be. For example, text alternatives can be used by people who can see the image and knowing the absolute locations of objects may help to quickly identify an object. Also, if there were multiple instances of the same object, the absolute location may be a quicker reference than a relative location. While the question should not be removed, the guidance can recommend that the information be identified when there are many instances of the same or similar object.

IMP 62. For the Spatial Relationship sub-category, add guidance recommending that the absolute location of objects be identified if there are multiple instances of the same or similar object in the image.

The information regarding When was generally rated as Significant. The information was never rated as Not Important, indicating that it will generally be of some importance to users. The guidance should mention that this information tends to be Significant when applicable and therefore should be provided to users.

IMP 63. For the When sub-category, add guidance that When information tends to be considered as Significant when applicable and should be provided to users.

Within the Sequential Relationship sub-category, the question regarding the basis for the ordering in the sequential relationship was often rated as Helpful. The time when it was rated as Significant was when the response described the basis for the ordering with regards to all of the sequential relationships in the image. Users seemed to prefer to know the overall response rather than a response with respect to each individual sequential relationship. Therefore, the guidance should recommend that this question be answered with respect to the entire image.

IMP 64. In the Sequential Relationship sub-category, add guidance that the question "What is a suitable basis for logically ordering of the individual steps or components?" be answered with respect to the entire image rather than each individual step in the sequential relationship.

The information regarding How Much was generally considered as Significant. This was especially true when the value had special meaning beyond stating the number of times a particular object appears in an image (for example, the number of people). Otherwise, the information was considered as Helpful. The guidance should recommend that the information be presented most of the time, especially when the value has a special meaning.

IMP 65. In the How Much category, add guidance that a quantitative value be presented to users most of the time, especially when the value has a special meaning or significance.

While the information regarding How was most often considered as Helpful, it was originally assumed that information regarding any physical interactions with the image would be of higher importance since it informs the user of the outcomes and changes due to such interactions. It was previously expressed to the researcher that at least visually impaired users would greatly benefit from this information since they may not otherwise be aware of the changes to the computing environment. It is possible that such information would be of higher importance if the outcome was out of the ordinary or altered the contents of the document. Additional research may be necessary to clarify when it is necessary to provide information regarding How.

7.5 Conclusion

This research study provided a general sense of the importance level for the different types of information. It helped to identify guidance on when certain types of information may be of higher importance and users would benefit from the information. It also helped to determine when certain types of information may be of lower importance so that less time can be spent on identifying the information. Additional research could be conducted in the future to further understand the conditions of when a type of information is of higher importance and when it is of lower importance.

Chapter 8 Conclusions and Future Work

Images and other graphical content are being used more frequently as the method for delivering content. For people who cannot see the images (due to visual impairments or technological limitations and/or choices), they would not receive the information communicated by the image without the information being presented in another manner, such as text. Document editor programs allow people to add or attach text alternatives to the graphical content. Some of the programs place length limitations on the text alternatives while others do not. Guidelines (such as WCAG [66]) require that text alternatives exist for all images in a Web page, but text alternatives are rarely available.

This thesis showed that one possible reason that people would not provide text alternatives (making use of the empty text instead) is if the person could not produce text alternative quickly and easily. While people may understand the importance of text alternatives, they would choose to use the empty text or copy the caption if they cannot generate what they believe to be a good text alternative. This implies that guidance is needed to help people determine the information that may be important to include.

From the review of the literature, a procedure for producing text alternatives was created, including the types of information being communicated by images. The steps of the procedure related to identifying information were evaluated to improve its usefulness and effectiveness. The results showed that some improvements are needed to increase the efficiency and clarify the technical details of the procedure. Overall, the results showed that the procedure helped to identify information that people considered as important, which were otherwise missed when left to their own devices.

The tool, which was used to evaluate the procedure, successfully eliminated confusion regarding the steps of the procedure and the tasks to complete. While a tool should also improve the efficiency of the procedure, the tool used for this thesis was essentially used to determine if the quality and quantity of information improved when people were required to consider all of the types of information. It helped to identify questions that were often misinterpreted as well as

questions that were missed by the set of questions. It also identified design considerations for an efficient tool.

The results presented in this thesis show that with the use of the procedure, all user groups identified a greater amount of information that was important to users. While some user groups identified less information than the average, all user groups identified more of the information that was considered as Essential, Significant, and Helpful with the use of the procedure than without. This shows that the procedure is successful in helping people identify information for text alternatives.

The results also indicated that some types of information tended to be more important to users than others. Based on the importance ratings provided by users, further guidance was identified regarding the importance of each type of information. Such guidance could guide people towards identifying more of the information that is of importance to users.

This thesis showed that the amount of time spent on applying the procedure was an issue for all participants. As previously stated, the ability to create text alternatives quickly and easily affected when text alternatives were provided to users. With respect to the document format, the problem was due to the interpretation of the procedure in order to apply it. With respect to the tool format, the problem was due to the technical implementations of the tool as well as the necessity to consider every question in the set of questions. The participants' comments as well as the analysis of all research studies identified numerous improvements that could be made to improve the amount of time spent applying the procedure in both the document and the tool format. These suggestions were enumerated throughout the previous chapters and are summarized in Section 8.2.

8.1 Lessons Learned

The research studies showed that people are very different in the way they identify information and the amount of time they put into creating text alternatives. Although all of the participants volunteered, which may have created a feeling of obligation in some participants, there was a great range in the results with respect to the amount of time spent and the information identified. Some participants followed the procedure fully (identifying components and seriously answering all of the questions) while others applied the procedure quickly (answering some questions without identifying components). Some participants identified only

the information that was asked of them while others considered information that was not directly asked (making use of questions like “What other information about the person is important for users to know?”) and identified information that was not in the set of questions. The procedure (and a tool supporting the procedure) needs to be flexible such that it helps people who want to apply the procedure quickly as well as those who wish to apply it more fully. The procedure and tool needs to also support users who would freely identify information based on a general question (such as “What is the appearance of the person?”) as well as those who need direct questions to be asked (such as “What is the age of the person?” and “What is the person wearing?”).

In the Research Study of the Procedure in Document Format, the questions were presented within their categories while the questions were presented out of context in the Research Study of the Procedure as a Prototype Tool. To minimize misinterpretations, it is preferable to indicate the context of each question or the questions should be rephrased to include the context.

Rather than consistently focusing on one aspect of an image (either describing the entire image or describing just a single component for all questions), the participants would frequently switch between answering some questions with respect to the whole image and other questions with respect to a component. Again, this was likely because the questions did not specify the context. This showed that participants would create their own mental model of the context when it is not specified to them. In the future, it would be preferable for a context to be defined prior to asking a question or within the question.

The participants requested additional examples and help for questions that seemed to be straightforward. The examples and help could help the use to properly answer the questions. This, along with misinterpretation of questions, showed the importance of having examples so that users know that they are doing things correctly. This would minimize the amount of time users spend trying to determine how they are expected to answer questions.

It is very difficult to definitively measure the amount of information identified. The research methods in this thesis made use of number of questions answered, number of types of information, and number of words to measure the quantity of information. These measures are closely linked since each type of information has a number of questions and each question would have a number of words in its response. None of these measures qualify the words and

information identified; however, the concern should not be on how well the words are written, but on how informative and useful the information can be to users.

It is possible for one person to use fewer words to say the exactly the same thing as another person. It is also possible for a person to provide the same words for multiple questions. For this research project, it was more important that people consider the information identified by a question than what information was provided. It showed an effort to apply the procedure and identify information about an image. While measuring the number of words is not a definitive way of measuring the amount of raw material, it is a helpful and convenient way of measuring it.

The research studies allowed the participants to complete each research study on their own and at their own pace using online tools. Even though the participants were given a week to complete each study, many required constant reminders. It appeared as though the research studies were completed at the last minute rather than spanned over a period of time, which was recommended. This behaviour was expected since it was volunteer-based and the participants were not scheduled to conduct the research study at a specified time. One reason for using online tools instead of conducting in-person experiments was to reach a wider audience and include people (such as usability and accessibility specialists and people with visual impairments) who would not otherwise be able to participate due to distance and location. In the future, it may be preferable to conduct a localized research study in order to ensure completion in a timely manner.

While the research questions and methodologies were known prior to each research study, they were not as specific and detailed as was needed. Some research questions were not specific enough or phrased such that it was clear whether or not have been answered. There were other times when the methods previously defined needed to be modified slightly in order to answer the research questions. Along the way, it became clear the importance of clearly defining detailed research questions and methodology so that a definitive answer could be reached from a research standpoint.

8.2 Future Work

This thesis laid out the basis and foundation for improvements on producing informative text alternatives for images. There are many improvements that can be made in the future in order for the procedure to be more usable. The following is a list of potential future work:

Modifications to the Document

- Restructure the document to clarify and emphasize the steps of the procedure so that it is easier to apply.
- Provide a more detailed introduction to the document to better emphasize the purpose and importance of the procedure.
- Add emphasis on the importance of the image's purpose and its context of use when applying the procedure.
- Provide guidance on how to apply the procedure when the person may not have the necessary expertise to describe or interpret the image.
- Include examples illustrating the steps of the procedure.
- Eliminate duplicate questions in the set of questions.
- Rephrase questions to minimize misinterpretation.
- Add questions and further examples to identify details for certain types of information users want to know but were not identified.
- Add guidance that suggests the importance of certain type of information, as was identified in Sections 7.4.3 and 7.4.4.

Tool Considerations

- Separate each step of the procedure into a separate Web page in the tool.
- Automatically save input from the user.
- If the tool allows the user to review information, the information should be presented in a way that allows for easy access to edit or add information to specific questions.
- Make use of existing image tagging software to suggest image components that exist in the image.

- Make use of existing libraries or thesauruses to suggest more descriptive terms to use in text alternatives.
- Incorporate the procedure into existing document authoring tools.

Additional Research

- Determine the amount of time that would be acceptable to people for applying the procedure to produce text alternatives.
- Determine the flow or sequence of questions such that it would maximize the identification of important information, minimize the identification of unimportant information, and minimize the number of questions to be answered.
- Customize the set of questions based on image type or image purpose.
- Determine how to formulate text alternatives from the identified information.

8.2.1 Modifications to Procedure Document

Many of the suggestions on modifying the procedure document came from user comments and results of the research study. In preparation to submit the procedure document to the International Standards Organization, some of the suggestions were made prior to the completion of this thesis. These modifications include restructure of the document, expansion of the introduction, emphasis on purpose and content of use, guidance for when expertise is necessary, inclusion of examples, and removal of duplicate questions. The updated and modified ISO 20071-11 document can be found in Appendix D.

Some suggestions (which were identified through the later research studies) have yet to be made into the document already submitted to the International Standards Organization. This includes rephrasing questions to minimize misinterpretation, adding questions to include more details, and adding guidance on the general importance level for a type of information (such as the purpose and image classification tends to be considered as Essential).

8.2.2 Tool Considerations

The tool developed and used in the research study in this thesis was designed to validate the procedure and the set of questions described in the procedure document. While the tool successfully eliminated confusion regarding how to apply the procedure, many improvements and considerations are needed to develop a usable tool. These improvements include separation

of tasks, automation, and incorporation of existing technologies for individual steps of the procedure.

The tool in the research study forced the participants to answer every question in the set of questions, which applied the procedure thoroughly. It is not feasible to expect people to always apply the procedure so thoroughly. Realistically, a tool implementing the procedure should allow people to apply the procedure quickly (answering only a small selection of questions), apply the procedure fully and thoroughly (answering all of the questions for every component), or somewhere in between the two extremes. In this way, the procedure can help people write information text alternatives regardless of the amount of effort they are willing to put in or the amount of time available to them.

The Research Study of the Identified Information identified some trends regarding the general importance levels of certain types of information and details. Making use of the guidance, a tool can make suggestion regarding the types of information that may be important to users. This can help people identify information of higher importance more quickly and save time. Conversely, a tool can also make suggestions regarding types of information that may not be very important to users, which can save people time from identifying and considering information that users may not wish to know about.

During discussions with accessibility experts, many have suggested making use of the various games or tools that already exist for image tagging. These image tagging tools could be used to suggest a default list of components that may benefit from being identified in a particular image and described in the text alternatives. Inclusion of such technologies could expedite the step for identifying image components. The terms that an image was tagged with may also be used to automatically provide answers to specific questions, such as “smile” or “smirk” for the question regarding facial expressions. Similarly, image recognition technologies may be included to identify image components and some details about the component, such as colour. It would be necessary to review such automations to ensure that the identified components and responses to questions are relevant to the context of the image.

Audio description and captioning make use of descriptive words to provide more information in a short amount of time or limited space. Text alternatives could also benefit from using descriptive words. Making use of existing libraries or thesauruses to suggest words that might be more applicable would help produce more vivid and succinct text alternatives.

Many technologies have specific containers for storing text alternatives. Document authoring tools provides access to those containers and sometimes default text alternatives are inserted. The procedure could also be incorporated into these document authoring tools so that when an image is inserted into a document, a wizard could appear to help create text alternatives for that image.

While the inclusion of such tools and technologies could help improve the amount of time that is spent applying the procedure, it does not replace the procedure.

8.2.3 Appropriate Amount of Time

The amount of time required to apply the procedure was the major issue identified by the research studies. While it took a short amount of time to write text alternatives without the use of the procedure, the information identified may not be useful to the users, especially when it repeats the image's caption. However, it is unreasonable to expect that people spend excessive amounts of time on writing text alternatives.

It is recognized that people can spend large amounts of time to write text alternatives if they wished, but the procedure should also be able to be completed in a short amount of time. Additional research is necessary to determine the amount of time people currently spend on text alternatives, the amount of time people spend when forced to write text alternatives, and the amount of time people believe is reasonable for applying the procedure.

8.2.4 Flow of Questions

The tool used in this thesis to apply the procedure forced people to consider and answer every question in the set of questions. As a result, it took a lot of time to consider and provide responses to all of the questions. The tool and procedure could benefit from guidance to focus on the important questions and to filter out types of information that might not apply.

While the tool used in this thesis attempted to filter out some categories of information with the use of Yes/No questions for certain categories and sub-categories of information, additional research is needed to determine the sequence of questions to be presented. It may become necessary that the questions be presented based on previously answered questions rather than on the types of information.

Additional research is also required to determine when questions might not be applicable for an instance of the image and should be removed, allowing people to focus on the important

information. This may require research into customizing the set of questions to those suitable only for specific types of images. Although the procedure itself should be generic so that it is applicable to a wide range of images, the tool could benefit from customization to image types.

8.2.5 Customized Set of Questions

The set of questions described in Section 3.2 is meant to be generic so that it can be applied to a wide range of images. It is recognized that every image is unique and may not contain information from every category or sub-category. The set of questions can then be customized such that only a subset of questions would be answered. This would decrease the amount of time spent on identifying information about an image. The field of Audio Description similarly created customized guidance for specific situations. [3, 4]

This customization can be done based on the image purpose (e.g., decoration, bullet, informative) and/or image type (e.g., painting, comic, map). Further research is needed to identify the subset of questions that would be appropriate and necessary for each customization. This is part of future work.

While the customizations can be created, it should be kept in mind that there may be situations where the customizations may not apply. For such situations, the full set of questions should be made available for users to choose their own subset of questions.

8.2.6 Text Alternative Formulation

After identifying the important information being communicated by an image, the information needs to be organized and turned into text alternatives. This thesis focused on ensuring the important information was identified. Organizing the information into text alternatives was left at the discretion of the person writing the text alternative. Additional research is needed to provide guidance on how to organize and present the information.

This may involve creating standard sentence structures that state which information and in which order (e.g., placing the classification of the image in square brackets at the start of every text alternative). These sentence structures may be in fill-in-the-blank formats, allowing people to easily create text alternatives. A tool could also make use of such sentence structures to automatically create text alternatives based on the answers to questions.

8.3 Contributions

This thesis made four major contributions: (1) identification of information that may be communicated by an image (Section 3.2); (2) a general set of questions to identify information being communicated by an image (Section 3.2); (3) a four-level structure for evaluating the importance of the information communicated by an image (Section 3.1.6); and (4) a procedure for creating text alternatives (Chapter 3).

The existing guidance on text alternatives provided examples of what possible text alternatives could be for those particular example images. However, there was no guidance on the types of information to consider when producing those examples or on a procedure to apply. This thesis compiled a set of types of information that an image could communicate based on existing research in the fields of library science, audio description, captioning, and art description. These types of information were then categorized based on Why, What, Who, Where, When, How Much, and How. This list of types of information can help people consider information that might have been missed otherwise. This thesis recognizes that not all images would contain all of the different types of information.

With the identification of the types of information, it became necessary to define a way of extracting that information from people about an image. A set of questions was developed to explicitly ask the person for the information and extract that information if it was available. This set of questions was modified based on the results of the research studies so that the set of questions could be applied to a wide range of images.

With the large amount of information that could be communicated by an image, it was necessary to develop a structure for evaluating the importance of the information so that the most important information can be presented to the user first. A three-level structure was presented in this thesis, along with a list of properties to help people determine if the information was important for a particular instance. The list of properties considers importance from the points of view of the content provider as well as the user. People can then decide the most appropriate manner for presenting the information based on the importance, may it be in the main document text or as text alternatives.

The biggest contribution of this thesis is the procedure for creating text alternatives. Although there is a requirement to provide text alternatives, there was previously no explicitly defined process or method for creating text alternatives. While much of the existing guidance is

focused on Web technologies, the procedure in this thesis can be applied to any document format that may contain images. The procedure was presented to the International Standards Organization's user interfaces group to become a technical specification (ISO 20071-11), which will be made available internationally. Developers around the world can then make use of this technical specification to create informative text alternatives or tools to support the creation of text alternatives.

Bibliography

[1] American Foundation for the Blind (AFB). n.d. *Improving Your Web Site's Accessibility*. Retrieved 6 January 2010, from http://www.afb.org/Section.asp?SectionID=4&TopicID=167&DocumentID=2176

[2] Antonacopoulos, A., Karatzas, D., and Ortiz Lopez, J. 2001. Accessing textual information embedded in internet images. In *Proceedings of SPIE Internet Imaging II*. San Jose, USA, 24-26 January 2001, 198-205.

[3] Audio Description Project (ADP). 2009. *Audio Description Standards*. Retrieved 30 September 2009, from http://www.acb.org/adp/docs/ADP_Standards.doc

[4] Audio Description Project (ADP). 2009. S*tandards for Audio Description and Code of Professional Conduct for Describers*. Retrieved 30 September 2009, from http://www.audiodescriptioncoalition.org/adc_standards_090615.pdf

[5] Bennett, D.J. 1999. *Presenting Diagrams in Sound for Blind People*. PhD Thesis, University of York, York, U.K.

[6] Bennett, D.J. and Edwards, A.D.N. 1998. Exploration of Non-seen Diagrams. In *International Conference on Auditory Display 1998*. Glasgow, UK, 1-4 November 1998.

[7] Bigham, J.P. 2007. Increasing web accessibility by automatically judging alternative text quality. In *Proceedings of the 12th international Conference on intelligent User interfaces*, Honolulu, Hawaii, USA, 28-31 January 2007, IUI '07. ACM, New York, NY, 349-352.

[8] Bigham, J.P., Kaminsky, R.S., Ladner, R.E., Danielsson, O.M., and Hempton, G.L. 2006. WebInSight: making web images accessible. In *Proceedings of the 8th international ACM SIGACCESS Conference on Computers and Accessibility*, Portland, Oregon, USA, 23-25 October 2006. Assets '06. ACM, New York, NY, 181-188.

[9] Brown, A., Stevens, R., and Pettifer, S. 2004. Issues in the Non-Visual Presentation of Graph Based Diagrams. In *Proceedings of the information Visualisation, Eighth international Conference*, 14-16 July 2004. IV. IEEE Computer Society, Washington, DC, 671-676.

[10] Canadian Association of Broadcasters (CAB). 2008. *Closed Captioning Standards and Protocol for Canadian English Language Television Programming Services*. Retrieved 4 January 2010, from http://www.cabacr.ca/english/social/captioning/captioning.pdf

[11] Craven, T.C. 2006. Some features of alt texts associated with images in Web pages. In *Information Research, 11*(2) (January 2006). Retrieved 6 January 2010, from http://informationr.net/ir/11-2/paper250.html

[12] Defense Visual Information (DVI) Directorate. n.d. *Captioning Basics*. Retrieved on 30 September 2009 from http://www.defenseimagery.mil/learning/captionstyle/basics.html

[13] Description Key. 2008. *What to Describe*. Retrieved 6 January 2010, from http://www.dcmp.org/descriptionkey/what_to_describe.html

[14] DiSimone, C. 2010. *How to Describe Art*. Retrieved 22 September 2010, from http://www.ehow.com/how_6146108_describe-art.html

[15] Eakins, J.P. 2001. Retrieval of still images by content. In *Lectures on information Retrieval, Springer Lecture Notes In Computer Science Series, 1980*. M. Agosti, F. Crestani, and G. Pasi, Eds., Springer-Verlag, New York, New York, NY, 111-138.

[16] English Grammar Online. 2010. *Picture Description*. Retrieved 22 September 2010, from http://www.ego4u.com/en/cram-up/writing/picture-description

[17] Eriksson, Y. 1999. How to make tactile pictures understandable to the blind reader. In *Proceedings from 65th International Federation of Library Associations and Institutes (IFLA) Council and General Conference*. Bangkok, Thailand, 20-18 August 1999.

[18] Esaak, S. n.d. What Is Art?. *Art History About.com*. Retrieved 10 January 2011, from http://arthistory.about.com/cs/reference/f/what_is_art.htm

[19] Flavell, A.J. 1999. Use of ALT texts in IMGs. *Web Design Group at University of Washington*. Retrieved 20 September 2009, from http://htmlhelp.com/feature/art3.htm

[20] Fredj, Z.B. and Duce, D.A. 2006. GraSSML: accessible smart schematic diagrams for all. In *Proceedings of the 2006 international Cross-Disciplinary Workshop on Web Accessibility (W4a): Building the Mobile Web: Rediscovering Accessibility?*, Edinburgh, U.K., 22 May 2006. W4A '06, *134*. ACM, New York, NY, 57-60.

[21] Gagnon, L., Foucher, S., Héritier, M., Lalonde, M., Byrns, D., Chapdelaine, C., Turner, J., Mathieu, S., Laurendeau, D., Nguyen, N.T. and Ouellet, D. 2009. Towards computer-vision software tools to increase production and accessibility of video description for people with vision loss. *Universal Access in the Information Society*, *8*(3) (August 2009). 199-218.

[22] George Eastman House: International Museum of Photography and Film. n.d. *Analyze a Photo Document*. Retrieved 20 September 2009, from http://education.eastmanhouse.org/discover/kits/files/11/analyze.php

[23] Google, 2011. Google Image Labeler. http://images.google.com/imagelabeler

[24] Heller, R.S. and Martin, C.D. 1995. A media taxonomy. In *Multimedia, IEEE*, *2*(4) (Winter 1995), 36-45.

[25] High Tech Center Training Unit. n.d. *Creating Tactile Graphics*. Retrieved 20 September 2009, from http://www.htctu.fhda.edu/trainings/manuals/alt/Tactile_Graphics.pdf

[26] Hixson, C. 2003. *Analyzing the Subject of a Picture: Guidelines*. Retrieved 10 January 2011, from http://libweb.uoregon.edu/inc/data/catdept/meta/Image_Subj_Analysis.pdf

[27] Hunter, B., Crismore, A., and Pearson, P.D. 1987. Visual Displays in Basal Readers and Social Studies Textbooks. In *The Psychology of Illustration, Instructional Issues, Vol. 2*, D.M. Willows and H.A. Houghton, Eds., Springer, New York, 116-135.

[28] Image of the Young Lady and Old Woman Optical Illusion. Retrieved on 27 September 2009, from http://www.bobjude.co.uk/entertain/illusion/illp.html

[29] Ina, S. 1996. Presentation of images for the blind. In *ACM SIGCAPH Computers and the Physically Handicapped*, (56) (Sep. 1996), 10-16.

[30] ISO/FDIS 14915-3. Software ergonomics for multimedia user interface – Part 3: Media selection and combination.

[31] ISO/IEC 20071-11. Information technology – User interface guidance – Guidance on text alternatives for images.

[32] Jaimes, A. 2003. Conceptual Structures and Computational Methods for Indexing and Organization of Visual Information. *Thesis Dissertation*. ProQuest.

[33] Japan Electronics and Information Technology Industries Association (JEITA). EXIF 2.2. Retrieved 18 October 2011, from http://www.exif.org/Exif2-2.PDF

[34] Kardoulias, T. n.d. *Guidelines for Making Tactile Diagrams and Accompanying Narratives*. Retrieved 9 January 2012, from http://www.artbeyondsight.org/handbook/acs-tactileguidelines.shtml

[35] Korpela, J. 2010. *Guidelines on alt texts in IMG elements*. Retrieved 20 September 2009, from http://www.cs.tut.fi/~jkorpela/html/alt.html

[36] Library of Congress, 2011. Prints & Photographs Online Catalog. http://www.loc.gov/pictures/

[37] Ma, X., Boyd-Graber, J., Nikolova, S., and Cook, P.R. 2009. Speaking through pictures: images vs. icons. In *Proceedings of the 11th international ACM SIGACCESS Conference on Computers and Accessibility*, Pittsburgh, Pennsylvania, USA, 25-28 October 2009. Assets '09. ACM, New York, NY, 163-170.

[38] McEwan, T. and Weerts, B. 2007. ALT text and basic accessibility. In *Proceedings of the 21st British HCI Group Annual Conference on HCI 2008: People and Computers Xxi: Hci..But Not As We Know It - Volume 2*, United Kingdom, 3-7 September 2007. British Computer Society Conference on Human-Computer Interaction. British Computer Society, Swinton, UK, 71-74.

[39] Metatla, O., Bryan-Kinns, N., and Stockman, T. 2007. Using hierarchies to support non-visual access to relational diagrams. In *Proceedings of the 21st British HCI Group Annual Conference on HCI 2008: People and Computers Xxi: Hci..But Not As We Know It - Volume 1*, United Kingdom, 3-7 September 2007. British Computer Society Conference on Human-Computer Interaction. British Computer Society, Swinton, UK, 215-225.

[40] Metatla, O., Bryan-Kinns, N., and Stockman, T. 2008. Constructing relational diagrams in audio: the multiple perspective hierarchical approach. In *Proceedings of the 10th international ACM SIGACCESS Conference on Computers and Accessibility*, Halifax, Nova Scotia, Canada, 13-15 October 2008. Assets '08. ACM, New York, NY, 97-104.

[41] National Braille Association, Inc. (NBA). 2000. *Excerpts from the NBA Tape Recording Manual, Third Edition*. World Wide Web Consortium (W3C). Retrieved 19 September 2009, from http://www.w3.org/2000/08/nba-manual/Overview.html

[42] Olsen, M.G., Snaprud, M., and Nietzio, A. 2010. Automatic checking of alternative texts on web pages. In *Proceedings of the 12th international conference on Computers helping people with special needs: Part I (ICCHP'10)*, K. Miesenberger, J. Klaus, W. Zagler, and A. Karshmer, Eds., Springer-Verlag, Berlin, Heidelberg, 425-432.

[43] Paek, S. and Smith, J. R. 1998. Detecting image purpose in World-Wide Web documents. In *Proceedings of the IS&T/SPIE Symposium on Electronic Imaging: Science and Technology*, San Jose, California, USA.

[44] Panofsky, E. 1962. *Studies in Iconology: Humanistic Themes in the Art of the Renaissance*. Oxford University Press, New York.

[45] Patil, S.R. 2007. Accessible image file formats: the need and the way (position paper). In *Proceedings of the 2007 international cross-disciplinary conference on Web accessibility (W4A) (W4A '07)*. ACM, New York, NY, USA, 40-43.

[46] Petrie, H., Harrison, C., and Dev, S. 2005. Describing images on the web: a survey of current practice and prospects for the future. In *Proceedings of Human Computer Interaction International (HCII) 2005*, Las Vegas, Nevada, USA, 22-27 June 2005.

[47] Petrie, H., Schlieder, C., Blenkhorn, P., Evans, G., King, A., O'Neill, A.-M., Ioannidis, G.T., Gallagher, B., Crombie, D., Mager, R., and Alafaci, M. 2002. TeDUB: a system for presenting and exploring technical drawings for blind people. In *Lecture Notes in Computer Science 239: Computers Helping People with Special Needs*. K. Miesenberger, J. Klaus and W. Zagler, Eds., Springer Verlag, Heidelberg.

[48] Poley, S. 2007. *The ALT attribute in HTML*. Web Matters at XS4ALL. Retrieved 20 September 2009, from http://www.xs4all.nl/~sbpoley/webmatters/alt.html

[49] Roland, C. 2010. Looking At Art: Seeing Questions. *Incredible @rt Department*. Retrieved 10 January 2011, from http://www.princetonol.com/groups/iad/Files/crit.htm

[50] Shatford, S. 1986. Analyzing the Subject of a Picture: A Theoretical Approach. *Cataloging & Classification Quarterly, 6*(3) (Spring 1986).

[51] Shen, H. T., Ooi, B. C., and Tan, K. 2000. Giving meanings to WWW images. In *Proceedings of the Eighth ACM international Conference on Multimedia*, Los Angeles, California, United States. 30 October – 3 November 2000. ACM, New York, NY, 39-47.

[52] Sierkowski, B. 2002. Achieving web accessibility. In *Proceedings of the 30th Annual ACM SIGUCCS Conference on User Services*, Providence, Rhode Island, USA, 20-23 November 2002. SIGUCCS '02. ACM, New York, NY, 288-291.

[53] Slatin, J. 2001. The art of ALT: Toward a more accessible Web. *Computers and Composition, 18*, 73-81.

[54] Tang, L. and Carter, J. 2010. *Guidance for Creating Alt-Text for Images*. Presentation presented to ISO/IEC JTC1/SC35 User Interface at the Meeting of the User Interface Accessibility Working Group, Venice, Italy, August 2010.

[55] Tang, L. and Carter, J. 2011. *Creating Alternative Text: Prototype Tool Evaluation Results*. Presentation presented to ISO/IEC JTC1/SC35 User Interface at the Meeting of the User Interface Accessibility Working Group, Bellevue, Washington, USA, February 2011.

[56] Tang, L. and Carter, J. 2011. Alternative Text for Images. In *The 26th Annual International Technology & Persons with Disabilities Conference (CSUN)*, San Diego, California, USA, 14-19 March 2011.

[57] Tang, L. and Carter, J. 2011. Communicating Image Content. In *55th Annual Meeting of the Human Factors and Ergonomics Society (HFES)*, Las Vegas, Nevada, USA, 19-23 September 2011.

[58] Tang, L. and Carter, J. 2011. *Verbalizing Images*. Poster session to be presented at the 14th International Conference on Human-Computer Interaction, Orlando, Florida, USA, July 2011.

[59] Tang, L., Fourney, D., Huang, F., and Carter, J.A. 2008. Secondary Encoding. In *Proceedings of the Human Factors and Ergonomics Society 52nd Annual Meeting*. New York, NY.

[60] Thatcher, J., Waddell, C., Henry, S., Swierenga, S., Urban, M., Burks, M., Regan, B., and Bohman, P. 2002. *Constructing Accessible Web Sites*. Glasshouse, UK.

[61] Vorburger, M. 1999. *Altifier: Web accessibility enhancement tool*. Retrieved 4 January 2010, from http://www.vorburger.ch/projects/alt/

[62] Web Accessibility for All. n.d. *How to Create Descriptive Text for Graphs, Charts & other Diagrams*. Retrieved 19 January 2010, from http://www.cew.wisc.edu/accessibility/tutorials/descriptionTutorial.htm

[63] Web Accessibility in Mind (WebAIM). 2010. *Appropriate Use of Alternative Text*. Retrieved 6 January 2010, from http://webaim.org/techniques/alttext/

[64] Web Accessibility in Mind (WebAIM). 2010. *Creating Accessible Images: Creating Effective Alternative (alt) Text*. Retrieved 6 January 2010, from http://webaim.org/techniques/images/alt_text

[65] World Wide Web Consortium (W3C). 2004. *Resource Description Framework (RDF)*. Retrieved 18 October 2011, from http://www.w3.org/RDF/

[66] World Wide Web Consortium (W3C). 2008. *Web Content Accessibility Guidelines (WCAG) 2.0*. Retrieved 19 September 2009, from http://www.w3.org/TR/WCAG20/Overview.html

[67] World Wide Web Consortium (W3C). 2010. *Understanding Conformance*. Retrieved 4 January 2010, from http://www.w3.org/TR/UNDERSTANDING-WCAG20/conformance.html

[68] World Wide Web Consortium (W3C). 2011. *HTML5: Techniques for providing useful text alternatives*. Retrieved 1 February 2011, from http://dev.w3.org/html5/alt-techniques/

Glossary

`alt` attribute	A property of the HTML `img` tag that shows a text alternative. It is a short description of an image.
Alternative Text	See “text alternative”.
Caption	The text that may appear above or below the image in question.
Container (for text alternatives)	A field or property through which text alternatives are provided to the user. It may be different for each technology. Also see “primary alternative text” and “secondary alternative text”.
Document	The file or article in which an image may reside. This includes Web pages, Word documents, PDF files, and presentation files. Also see “main document content” and “main document text”
`longdesc` attribute	A property of the HTML `img` tag that leads the user to another document that contains a text alternative. It is a long description of an image.
Main document content	The graphical and textual content of the file or article in which the image in question resides.
Main document text	Textual content of a document that is always presented to the users [31]
Primary alternative text	Main text alternative provided to users of screen readers [31]
Secondary alternative text	Additional text alternative provided to users of screen readers beyond primary alternative text. [31] Also see “caption”, “main document text”, “`alt` attribute”, and “`longdesc` attribute”.
Text Alternative	A textual description or representation of an image [31]

Appendix A Research Studies Images

The images presented here were used in all of the research studies. The Group A participants from each research study were presented with the same set of images and in the same order. The same was done for all Group B participants in each research study.

A.1 Set A Images

A.1.1 Image 1 (A1)

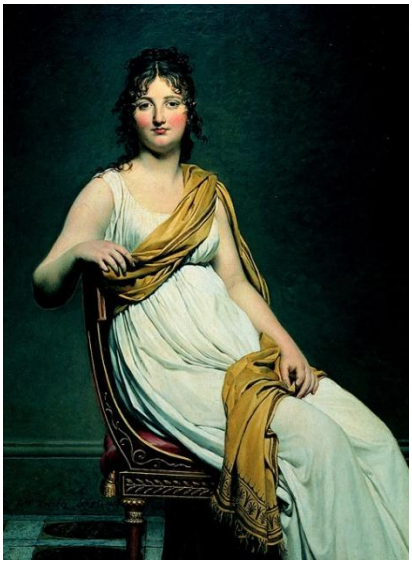


Figure A.1 Portrait Image for Research Studies

This image can be found at the following Web address: http://www.louvre.fr/llv/exposition/detail_exposition.jsp?CONTENT%3C%3Ecnt_id=10134198673493016&CURRENT_LLV_EXPO%3C%3Ecnt_id=10134198673493016&pageId=2&bmLocale=en

A.1.2 Image 2 (A2)

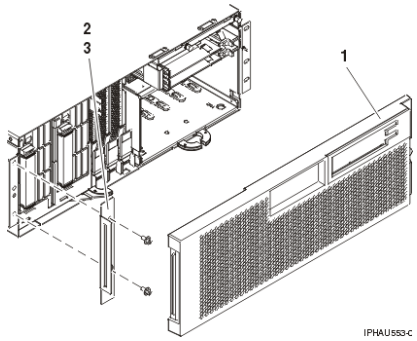


Figure A.2 Assembly Image for Research Study

This image can be found at the following Web address: http://publib.boulder.ibm.com/infocenter/powersys/v3r1m5/index.jsp?topic=/iphau_p5/pa570.htm

A.1.3 Image 3 (A3)

Figure A.3 Chart Image for Research Study

This image can be found at the following Web address: http://www.nyse.com/about/listed/lcddata.html?ticker=SPY&fq=D&ezd=1M&index=3. The image on the Web page may not be the same as the image presented here since the data and the image are updated every day.

A.1.4 Image 4 (A4)

Figure A.4 Landscape Image for Research Study

This image can be found at the following Web address: http://www.flickr.com/photos/newage/2128211277/

A.1.5 Image 5 (A5)

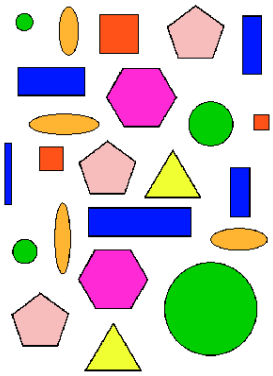


Figure A.5 Non-Representational Image for Research Study

This image can be found at the following Web address: http://www.teachingideas.co.uk/maths/shapes.htm

A.2 Set B Images

A.2.1 Image 1 (B1)

Figure A.6 Logo Image for Research Study

This image can be found at the following Web address: http://www.iqmetrix.com/careers. At the time of publication, the Web page and the company logo has changed from what is shown here.

A.2.2 Image 2 (B2)

Figure A.7 Image with a Person for Research Study

This image can be found at the following Web address: http://www.flickr.com/photos/yewenyi/2256618233/

A.2.3 Image 3 (B3)

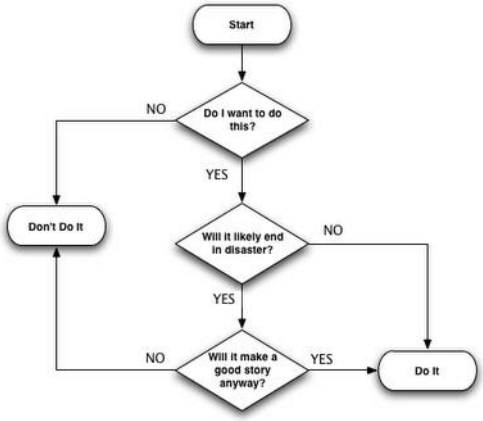


Figure A.8 Flow Chart Image for Research Study

This image can be found at the following Web address:
http://laughgraph.blogspot.com/2008/09/only-flowchart-youll-ever-need-ever.html

A.2.4 Image 4 (B4)

Figure A.9 Abstract Image for Research Study

This image can be found at the following Web address:
http://www.art.com/products/p10022263-sa-i789996/alfred-gockel-dancing.htm?rfid=859943

A.2.5 Image 5 (B5)

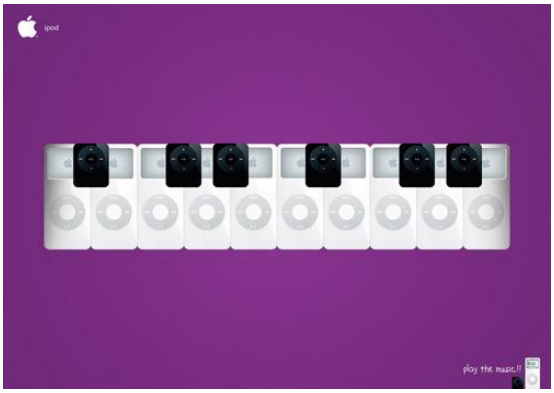


Figure A.10 Advertisement Image for Research Study

This image can be found at the following Web address: http://www.hongkiat.com/blog/60-creative-and-clever-advertisements/

Appendix B Research Studies Materials

The materials presented in this appendix were approved on ethical grounds by the University of Saskatchewan Behavioural Research Ethics Board. Since the materials and content are repeated for each image and for Group A and Group B images, a shortened version of the documents are presented here consisting of the instructions for a single image.

B.1 Questionnaire for Research Study Without Procedure

Participant #:

Age:

Part 1

Many websites make use of visual graphics and content (such as pictures and diagrams) to provide information and create visual appeal. For the visually impaired users, that content is unavailable to them. They do not know what important information they may be missing. Alternative textual descriptions (often referred to as "alt-text") can be used to provide information on an image to those people who cannot see the image.

Current web accessibility guidelines require that a short description (alt-text) be provided for all images in web pages and other documents. In fact, automatic checkers generally make sure that the alt-text exists.

Below are five (5) web addresses and images. Each web page contains an image that might not be able to be seen by all users of the page. Please write what you think would be suitable alt-text. Also, please record the length of time it took to write the alt-text. Provide as much alt-text as is needed to describe each image.

Image 1

[The image in question is presented here.]

Web page where image is used and where alt-text will be used: [The Web address for the image is presented here.]

Time spent describing Image 1: ___ minutes

Part 2

To beat the automatic checkers, it is possible to use the empty text (i.e., “”) as the alt-text. Knowing this, would you have used the empty text instead of what you wrote in Part 1 for any or all of the images? Why or why not?

Image 1:

B.2 Questionnaire for Research Study with Procedure in Document Format

Participant #:

User Category:

Age:

Task

Below are five (5) web addresses and images. Each web page contains an image that you are expected to describe while using the attached document. For each image, please answer/describe the image and its components using the table format provided. Please expand

each cell to contain as much information as you feel is important and expand each table to as many rows as is needed to describe the image.

Image 1

[The image in question is presented here.]

Website: [The Web address for the image is presented here.]

Whole image or component name	Information about the Image	Question(s) Answered	Significance level

Time spent describing Image 1

Feedback

1. What difficulties (if any) did you experience in reading, understanding, or using the procedure / tool provided to you?

2. What suggestions do you have on improving the procedure / tool or parts thereof?

3. Can the researcher contact you regarding your answers? (Yes/No) The purpose of this is to clarify problems or suggestions you described.

B.3 Instructions for Research Study with Procedure in Tool Format

Getting Started

You will be providing information about the image using the Image/Component Description Information page for that image. To see the web page that the image is used, click the "View Website" link beneath the image.

First, **identify the components** in the image. The Whole Image is automatically added as a component in the List of Components box on the left. You can add other components to the list by typing a unique component name in the Component Name field below the List of Components box and then clicking the Add Component button. To delete a component, select it in the List of Components box and then click the Delete Component button.

For help regarding identifying components, see the Description section of the Help Page.

To view questions and enter information for a particular component, select it from the List of Components box.

Each component has a specific set of questions to be answered. You can hide sub-questions (showing only the category question) by clicking the "Hide All Sub-Questions" button or the [-] in front of category questions. Similarly, you can show sub-questions by clicking the "Show All Sub-Questions" button or the [+] in front of category questions.

It is important to answer all questions and significance levels for each component you identify. "Not Applicable" can be a valid answer if a question is truly not applicable. A "Not Applicable" checkbox is provided for you to use in these circumstances. Checking the "Not Applicable" box for a category question will result in the sub-questions automatically being checked as not applicable.

If you have trouble with a question or do not understand the question, placing the mouse over a question will result in a pop up with an explanation and examples. If further explanation is needed, please try your best and let the research team know after you have completed using the tool.

Please note that **you must click the Save button in order to save the information you entered.** If you do not click the save button and navigate to another page or another component, **your progress will be lost.**

Provide as much information about the image as possible. Although you might not have the expert knowledge about the image, you can still describe what you see in the image. When rating the significance of the information, consider how much you would want/need to know that piece of information if you could not see the image.

After you have answered all questions and provided significance levels, click the "Review" button to look over all information entered. You will be prompted if there are information is still missing.

Additional help can be found on the Help page, which can be found on the top right corner of every page. Sample image descriptions can be found on the Examples page.

B.4 Help Page for Research Study with Procedure in Tool Format

About the Project

What is the purpose of this project?

Alternative text (or alt-text) helps people who cannot see the images understand what the image is of by providing the same information in textual form. Alt-text can be useful to those with visual impairments, those who turned images off in order to improve webpage loading speeds, and those who cannot understand the image being displayed.

This project intends to identify the information that should be known about an image and then collect that information. In the future, the information is then used to create the textual description about an image.

What improvements or changes are still to come?

This tool is constantly changing based on your comments and input. Make sure that we know of any issues you encounter or improvements you think of. We will consider your input and make changes accordingly. The changes we plan to make in the future include (but not limited to):

- Modify the Review page to direct the user to the missing information.
- Rephrase questions to quickly eliminate the questions that do not apply.
- Redesign the way questions are being presented. This may be to separate the questions across pages or to use a wizard.
- Have the tool remember the page location you were at before, as well as the collapse/expand states of the question groupings.
- Add more help tooltips to questions.
- Provide more guidance on the significance levels.
- Improve the loading speed of the tool.

If you have any suggestions that are not included in the list, we encourage you to contact the research team and let us know how we can make it better.

Contact Information

Please send any questions, concerns, or comments that you wish to share with the research team to userlab dot alttexttool at usask dot ca.

Login

I lost my login information. What should I do?

Contact the research team and they will send you the login information.

How do I change my username and/or password?

At this time, it is not possible to change your username or password. It is something we plan to include in the tool in the near future.

I've logged in but nothing appears on the page.

If you have successfully entered your login information and nothing appears on the page, contact the research team immediately. A technical problem likely occurred, which they will fix as soon as possible.

I've logged in but it says that I need to log in again.

If you are returning to the web application after leaving it for some time or closing your web browser window, then your session may have ended and you need to log in again. If the message appears after you have logged in again, contact the research team immediately. A technical problem may have occurred, which they will fix as soon as possible.

User Home Page Help

I don't see any images that I am supposed to describe.

There should be a list of 3 images on the home page. If 3 images do not appear on the page, please contact the research team. A technical problem likely occurred, which they will fix as soon as possible.

I've completed describing an image but the status still says "incomplete".

Ensure that you have completed the review process (by clicking the Review button on the Image / Component Description Information page). Once you have reviewed the entered information, the status will be updated to "Complete".

If you have fully completed the review process and the status remains as "Incomplete", please contact the research team. A technical problem likely occurred, which they will fix as soon as possible.

Describing Images

Why do I need to describe the images before writing the alternative text?

A sighted person receives a lot of information about an image at a glance. The same information should also be made available to people who cannot see the image. Therefore, it is important to know what information is being presented. Knowing what is in the image and the information being provided by the image can then help create more informative alternative text.

Why do I have to answer so many questions?

The set of questions have been developed to ensure that nothing important is missed. If you believe that the question is not applicable to the image or the image component, you can mark the question as "Not Applicable" by clicking on the check box. You will not need to provide a significance level in this case.

In the future, we plan to tailor the set of questions to the type of image, resulting in fewer questions.

I don't know what the question means.

Placing the mouse over a question will display an explanation of the question as well as examples. If the explanation or examples are not clear enough, please try your best and let us know about the issue in your feedback.

What do the significance levels mean?

There are three (3) significance levels.

- Significant — the information is needed by most users most of the time. This information is critical in understanding the image and the Web page.
- Useful — the information is needed by some users most of the time or most of the users some of the time. This information can be important in understanding the image and/or the Web page.
- Helpful — the information is needed by some users some of the time. This information can assist in understanding the image and/or the Web page.

The significance levels are subjective and depends on the context of the web page. It is recommended to view the web page while determining the significance level.

What are image components and how do I identify them?

Image components are parts of the image that may be important for the viewer to know about and understand in order to understand the entire image.

Examples include: persons, things, shapes, and areas in the image.

In the image below, the components could include the Whole Image, the plate of desserts, and the cup of coffee. The plate of desserts can be further broken down such that each dessert is a component.

The information that I previously entered is no longer there. What happened?

In order to save the information you entered, you must click the save button at the bottom of the Image Description page. If you close the page or load another component's information before saving, the information you entered will be lost. It is highly recommended that you save frequently.

I tried to edit information for an image but the Save button disappeared. What happened?

After you completed the review process for an image, the image description status is updated to "Complete" and it is not possible to make modifications. However, you may look at the information you entered and use it as a reference for other image descriptions.

If you have not completed the review process (meaning the status is "Incomplete" and the Save button does not appear, please contact the research team. A technical problem likely occurred, which they will fix as soon as possible.

Review Information

Why do I need to review the information I entered?

Reviewing allows you to make sure that the information you provided is accurate, consistent, and complete. When you entered the information, each component is presented separately. You might not be aware that information is missing or conflicting with other components. On the Review screen, all of the information you entered is on the same page so you can make sure that the information is what you had intended to enter.

What does the review process involve?

To review the information you entered, click the Review button on the Image/Component Description Information page. If you have not provided an answer and significance level for every question (including marking a question as not applicable), you will be provided with a list of questions that are missing information. Provide the missing information and click the Review button again.

If all answers and significance levels (including marking a question as not applicable) has been provided, your answers for all components will be presented together for you to look over. You may make modifications by clicking the Edit Information button and the Image/Component Description Information page will appear as a new window.

If no changes need to be made, enter the approximate amount of time spent describing that image and then click the “The Information is Complete” button.

Why is the order of components different from the Image/Components Description Information page?

While the components on the Image/Component Description Information page are in the order that they were added, the components on the Review page are in alphabetical order. The research team is aware of the issue and hope to resolve the problem in the near future.

Why do the pages take so long to load?

We ask that you be patient when using the prototype tool. Sometimes, the university servers may be slow and it could take longer to load the pages. Improvements will also be made to increase the loading speed in the future.

B.5 Instructions for Research Study on Importance of Identified Information

Introduction

ISO 20071-11 defines alternative text as "textual alternative for and/or description of an image". It is meant to communicate the same information to persons who cannot see the image (or other graphical content) as to those who can. Alternative text can be both long and short in length. While the term "alternative text" is often associated with a short description, this research study uses the term to mean both long and short descriptions.

ISO 20071-11 describes a procedure for creating informative alternative text. The procedure was applied on a set of ten (10) images to identify individual pieces of information about those images. In previous research studies, we gathered a collection of information about a

set of ten (10) images. In this research study, you will be rating the importance of the information for a set of five (5) images.

Task

For each image, you will be presented with a set of information about the image. You will be asked to rate the importance of each piece of information. The importance level of a piece of information helps determine if that information should be provided in alternative text.

There are four (4) importance levels you can assign to each piece of information: Essential, Significant, Helpful, and Not Important.

- **Essential** - Essential information is necessary to understand the image within the document. Essential information may have some or all of the following properties. As more of these properties apply, it is more likely that the information is essential.
  - It is aimed at the target audience of the document.
  - It must be known in order to comprehend the document.
  - Most people want / need it most of the time.
  - The person would be confused as to what the document is talking about without this information.
  - Without it, the person has no idea why the image is there or what the image is for.
  - It provides a good first impression of the image.
  - Based on this information, the user will determine if they need/want to know more about it.
  - For the content provider, this is the information that the content provider absolutely wants to tell people about.
  - It provides the essence, purpose, function, or intent of the image.
  - It identifies that the image conflicts with the main document text and that this conflict is intentional.
- **Significant** – Significant information is useful for getting a comprehensive understanding of the image, when such information is desired. Significant information satisfies the more detailed interests of most users most of the time. Significant information may have some or

all of the following properties. As more of these properties apply, it is more likely that the information is significant.

- It is aimed at the target audience.
- It gives a more detailed and thorough understanding of the image and/or document.
- It is information that could be obtained by more than a quick glance.
- The person should know about it as they are reading the document in order to understand the document.
- The person decided to know more based on the essential information. This information goes into more details about the essential information.
- Without this information, the person has an idea of what the image is about and the reason the image is there, but does not have a detailed understanding about it.
- For the content provider, this is information that further explains and gives more details on what the content provider wants to tell people.

- **Helpful** – Helpful information provides a more thorough understanding of the image within the document in which it appear for those users who wish a more detailed description of the image. Helpful information may have some or all of the following properties. As more of these properties apply, it is more likely that the information is helpful.
  - It is specific details that might be of interest to some who are the target audience of the document.
  - It is targeted towards very specific audiences (other than the target audience) or a subset of the target audience.
  - It provides the person with a better understanding of the image when the person is not an expert in the topic area or not the target audience of the document.
  - It might reassure the person that they have not missed something of greater importance.
  - Without this information, the person has a fairly complete understanding of what the document is about but have some things that the person still wants to know.
  - It includes different or other possible interpretations of the information being expressed by the image.
  - For the content provider, this is information that could clarify some things for some people.

- It includes optional extra information that is seldom wanted or needed, but elaborates on what is already there.

- **Not Important** – Information is not important if it does not help to provide much additional understanding of the image for any users in order to understand the image or document in which it occurs. This can include information that is not appropriate to consider given the context of the image within the document. Information that is not important may have some or all of the following properties. As more of these properties apply, it is more likely that the information is not important.
  - Very few to no persons will want to know or care to know this information.
  - It is rarely helpful.
  - It is not important enough to mention.
  - Without this information, the person knows everything they want or need to know in order to understand the document and/or image.
  - This is information that might result in unintended confusion or boredom and does not help people understand what the content provider is saying.
  - The importance level descriptions will be available on all subsequent pages.

The information about the images are not well-formed alternative text. Instead, they are individual pieces of information that could be used to create well-formed alternative text. Therefore, when you are evaluating each piece of information, consider the information itself rather than the way it is written.

Page Layout

The remainder of this research study makes use of frames to show you two web pages at the same time. In the top frame, you will be shown a copy of the web page containing the image that you are being asked about. All (or majority) of the images have been removed to simulate how the page would look when the images cannot be seen. At the top of the page is a link to skip to the location of the image being considered. The image can also be identified by the alternative text "USERLab is evaluating this image."

When you are considering the importance level for a piece of information, please keep in mind the context of the image. Use this top frame as your reference.

The bottom frame contains the questions regarding the image being considered. At the bottom of this frame are three (3) buttons for navigating through this research study. Please refrain from using the web browser navigation buttons because your progress may be lost. If you choose to take a break from the research study, make sure to log out so that your answers are properly saved. The next time you log in, you will return to the same page. You are free to return as many times as necessary to complete the research study.

The pages may be slow to load. We appreciate your patience if this occurs.

Note to Screen Reader Users

The top frame contains a duplicate of the original web page containing the image being evaluated. It may not have been designed to be accessible to all users. It may also make use of frames and ARIA landmarks. To help you identify the landmarks being used for the purpose of this research study, the landmarks have the roles "complimentary" (for the top frame) and "main" (for the bottom frame).

Begin Research Study

When you are ready to proceed, please click the "Begin Evaluation" button below.

Appendix C Research Studies Result Values

This appendix presents details values from the results of the research studies, as well as additional graphs and tables to illustrate the results.

C.1 Number of Questions Answered Across Studies

Table C.1 summarizes the average, minimum, and the maximum number of questions that were answered within each research study. The minimum and maximum values represent the single instance within each research study where one participant answered the least or most questions amongst all of the images.

Table C.1 Number of Questions Answered in Each Research Study

	Without Procedure	With Document	With Tool
Average	3.4	7.5	32.2
Minimum	1.0	1.0	8.0
Maximum	15.0	44.0	89.0

Table C.2, Table C.3, and Table C.4 specify the frequency that the participants answered the specified range of questions about an image without the procedure, with the procedure in document format, and with the procedure in tool format respectively.

Table C.2 Frequency of Number of Questions Answered Without the Procedure

Questions	1	2	3	4	5	6	7	8	9	10+
Frequency	7	15	11	6	5	2	0	0	0	3

Table C.3 Frequency of Number of Questions Answered with the Procedure in Document Format

Questions	0 – 2	2 – 4	4 – 6	6 – 8	8 – 10	10 – 12	12 – 14
Frequency	9	24	26	21	6	12	5

Questions	14 – 16	16 – 18	18 – 20	20 – 22	22 – 24	24+
Frequency	1	2	0	1	2	1

Table C.4 Frequency of Number of Questions Answered with the Procedure in Tool Format

Questions	0 – 5	5 – 10	10 – 15	15 – 20	20 – 25	25 – 30	30 – 35	35 – 40	40 - 45
Frequency	0	6	8	12	9	13	5	4	1

Questions	45 – 50	50 – 55	55 – 60	60 – 65	65 – 70	70 -75	75 – 80	80+
Frequency	3	4	2	1	2	2	2	2

Table C.5 lists the average number of questions answered by each user group for each image in Set A and Set B images, as well as for all images, when the procedure was provided in document format. Table C.6 similarly lists the averages for when the procedure was provided in tool format.

Table C.5 Average Number of Questions Answered by User Group with the Procedure in Document Format

	A1	A2	A3	A4	A5	B1	B2	B3	B4	B5	Overall
Specialists	4.5	3.0	3.5	3.5	4.0	9.7	9.3	18.7	8.0	7.7	6.8
Developers	11.0	8.3	7.0	13.0	13.3	5.0	10.5	3.5	6.0	5.5	7.0
Content Providers	7.0	6.0	5.0	4.5	5.0	6.3	6.3	10.7	8.3	7.7	5.1
Internet Users	7.3	6.0	6.0	4.7	4.3	7.3	7.3	4.0	5.0	7.7	5.3
Overall	15.1	6.3	5.7	7.5	7.6	7.3	8.2	9.7	6.9	7.3	

Table C.6 Average Number of Questions Answered by User Group with the Procedure in Tool Format

	A1	A2	A3	A4	A5	B1	B2	B3	B4	B5	Overall
Specialists	16.5	72.5	50.5	65.5	66.0	17.0	31.0	45.0	36.0	29.0	45.5
Developers	19.0	13.5	16.0	17.0	10.0	22.3	39.0	56.0	35.3	24.0	25.0
Content Providers	31.5	17.0	34.0	29.5	11.0	N/A	N/A	N/A	N/A	N/A	24.0
Internet Users	21.3	33.7	30.7	35.3	34.0	9.0	27.0	19.0	12.0	10.0	23.0
Overall	28.7	36.3	32.4	36.7	33.1	18.9	34.3	46.2	32.6	23.4	

Figure C.7 illustrates and compares the range of number of questions answered by each user group when using the procedure in document format. It shows the lowest, average, and highest values for each user group. Figure C.8 similarly presents the results for when using the procedure in tool format.

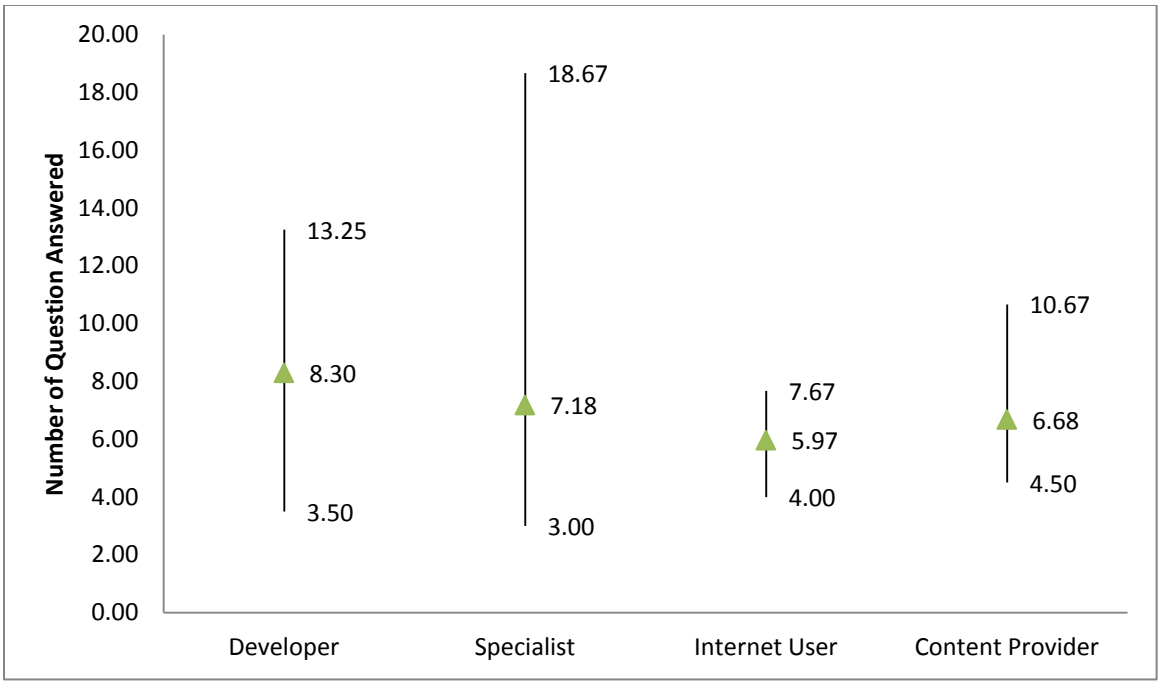


Figure C.7 Range of Number of Questions Answered by User Group with the Procedure in Document Format

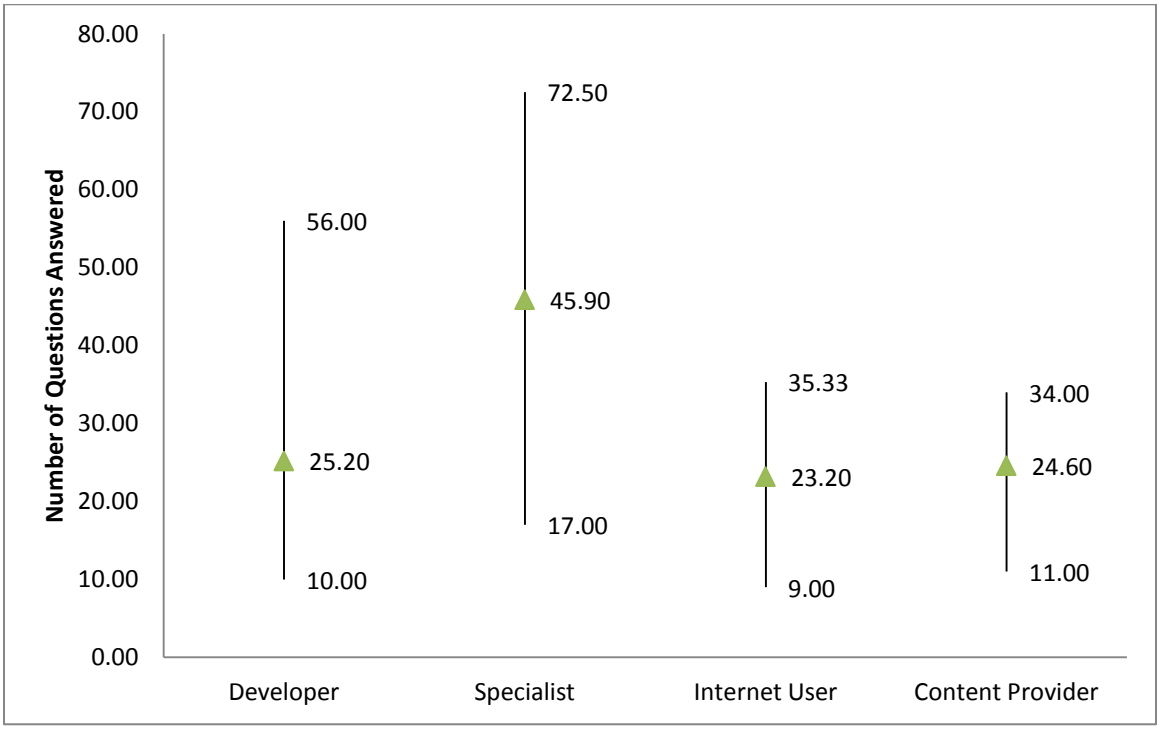


Figure C.8 Range of Number of Questions Answered by User Group with the Procedure in Tool Format

C.2 Amount of Time Spent Across Studies

Table C.9 summarizes the average, minimum, and the maximum number of questions that were answered within each research study. The minimum and maximum values represent the single instance within each research study where one participant spent the least or most time amongst all of the images.

Table C.9 Amount of Time Spent in Each Research Study (in Minutes)

	Without Procedure	With Document	With Tool
Average	2.9	15.9	31.2
Minimum	0.5	2	3
Maximum	10	60	160

Table C.10, Table C.11, and Table C.12 specify the frequency that the participants spent the specified range of time on an image without the procedure, with the procedure in document format, and with the procedure in tool format respectively.

Table C.10 Frequency of Time Spent Without the Procedure

Minutes	0 – 1	2	3	4	5	6	7	8	9	10
Frequency	13	15	8	2	3	0	1	1	1	2

Table C.11 Frequency of Time Spent with the Procedure in Document Format

Minutes	0 – 5	5 – 10	10 – 15	15 – 20	20 – 25	25 – 30
Frequency	30	23	15	14	4	8

Minutes	30 – 35	35 – 40	40 – 45	45 – 50	50 – 55	55 – 60
Frequency	1	5	3	2	0	2

Table C.12 Frequency of Time Spent with the Procedure in Tool Format

Minutes	0 – 5	5 – 10	10 – 15	15 – 20	20 – 25	25 – 30	30 – 35
Frequency	2	6	8	15	8	12	3

Minutes	35 – 40	40 – 45	45 – 50	50 – 55	55 – 60	60+
Frequency	3	6	2	0	6	5

Table C.13 lists the average time spent by each user group for each image in Set A and Set B images, as well as for all images, when the procedure was provided in document format. Table C.14 similarly lists the averages for when the procedure was provided in tool format.

Table C.13 Average Amount of Time Spent by User Group with the Procedure in Document Format

	A1	A2	A3	A4	A5	B1	B2	B3	B4	B5	Overall
Specialists	11.5	5.0	8.0	4.5	5.5	30.0	21.7	35.0	16.7	16.0	15.4
Developers	20.5	14.0	18.3	25.5	27.3	19.5	20.5	6.5	14.0	14.5	18.1
Content Providers	9.0	7.0	15.0	7.0	6.0	19.3	14.3	21.0	18.0	13.7	13.0
Internet Users	14.3	11.3	11.5	15.3	8.7	16.7	19.0	17.0	12.6	11.3	13.8
Overall	15.1	10.4	14.1	15.5	14.4	21.5	18.7	19.7	15.5	13.8	

Table C.14 Average Amount of Time Spent by User Group with the Procedure in Tool Format

	A1	A2	A3	A4	A5	B1	B2	B3	B4	B5	Overall
Specialists	42.5	65.0	50.0	50.0	45.0	11.5	12.5	25.0	20.0	17.5	33.9
Developers	28.0	20.0	20.0	19.5	16.5	27.5	24.3	40.0	62.5	23.8	28.2
Content Providers	37.5	60.0	45.0	30.0	30.0	N/A	N/A	N/A	N/A	N/A	40.5
Internet Users	33.7	40.0	36.7	31.7	28.3	17.0	18.0	19.0	16.0	18.0	25.8
Overall	35.2	43.8	36.9	32.7	29.8	21.4	19.3	31.5	40.8	21.1	

Figure C.15 illustrates and compares the range of time spent on the procedure by each user group when using the procedure in document format. It shows the lowest, average, and highest values for each user group. Figure C.16 similarly presents the results for using the procedure in tool format.

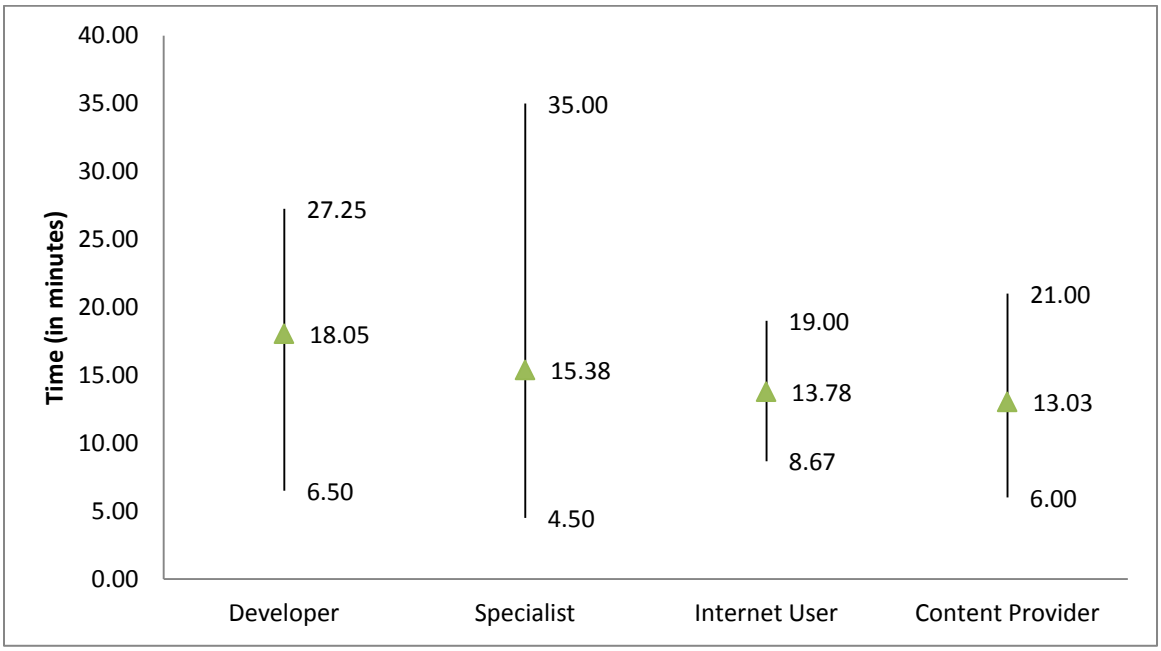


Figure C.15 Range of Time Spent by User Group with the Procedure in Document Format

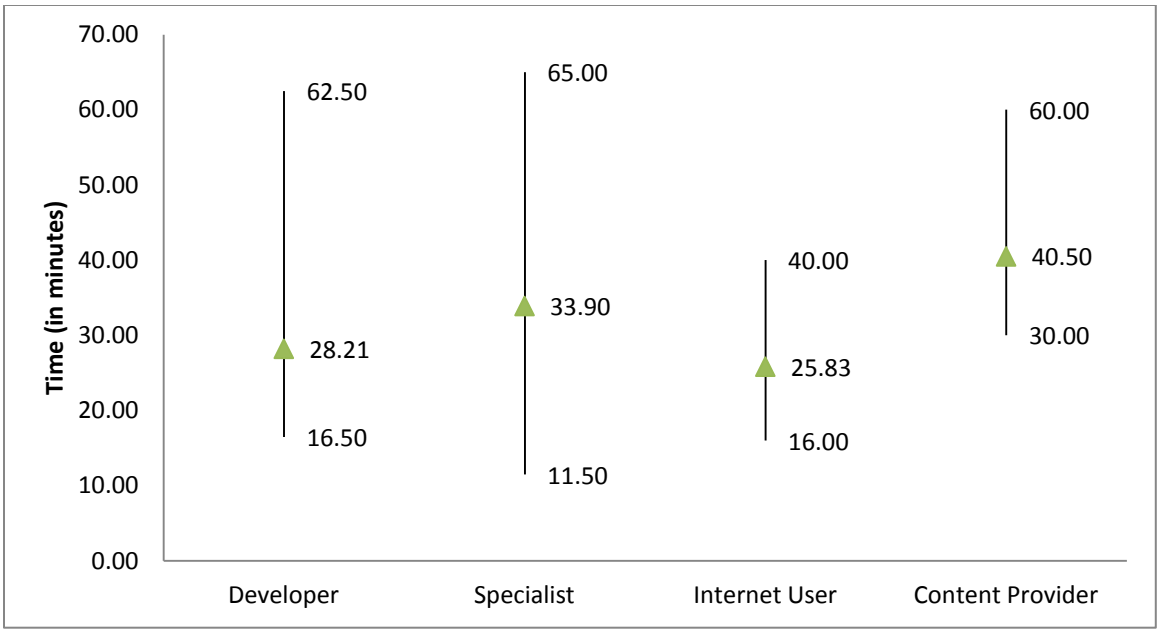


Figure C.16 Range of Time Spent by User Group with the Procedure in Tool Format

C.3 Number of Words Written Across Studies

Table C.17 summarizes the average, minimum, and the maximum number of questions that were answered within each research study. The minimum and maximum values represent the single instance within each research study where one participant wrote the least or most words amongst all of the images.

Table C.17 Number of Words Written in Each Research Study

	Without Procedure	With Document	With Tool
Average	23.2	73.4	154.8
Minimum	3	5	11
Maximum	116	323	654

Table C.18, Table C.19, and Table C.20 specify the frequency that the participants wrote the specified range of words for an image without the procedure, with the procedure in document format, and with the procedure in tool format respectively.

Table C.18 Frequency of Number of Words Written Without the Procedure

Words	0 – 10	10 – 20	20 – 30	30 – 40	40 – 50	50 - 60
Frequency	16	11	10	5	2	2

Words	60 – 70	70 – 80	80 – 90	90 – 100	100+
Frequency	1	1	0	0	1

Table C.19 Frequency of Number of Words Written with the Procedure in Document Format

Words	0 – 20	20 – 40	40 – 60	60 – 80	80 – 100	100 – 120
Frequency	26	19	18	12	9	3

Words	120 – 140	140 – 160	160 – 180	180 – 200	200 – 220	220 – 240
Frequency	5	5	5	2	0	1

Words	240 – 260	260 – 280	280 – 300	300+
Frequency	2	0	1	2

Table C.20 Frequency of Number of Words Written with Procedure in Tool Format

Words	0 – 25	25 – 50	50 – 75	75 – 100	100 – 125	125 – 150	150 - 175
Frequency	7	8	12	5	7	10	4

Words	175 – 200	200 – 225	225 – 250	250 – 275	275 – 300	300 – 325	325 – 350
Frequency	2	3	1	3	2	3	2

Words	350 – 375	375 – 400	400 – 425	425 – 450	450 – 475	475 – 500	500+
Frequency	2	0	1	2	0	0	2

Table C.21 lists the average number of words written by each user group for each image in Set A and Set B images, as well as for all images, when the procedure was provided in document format. Table C.22 similarly lists the averages for when the procedure was provided in tool format.

Table C.21 Average Number of Words Written by User Group with the Procedure in Document Format

	A1	A2	A3	A4	A5	B1	B2	B3	B4	B5	Overall
Specialists	36.5	35.5	32.0	20.5	17.0	118.0	93.3	167.7	89.3	90.3	70.0
Developers	81.3	84.8	75.3	79.8	51.3	134.0	133.5	53.0	159.5	85.5	93.8
Content Providers	74.0	61.0	93.5	33.5	53.0	60.0	85.0	186.3	68.3	102.7	81.7
Internet Users	35.0	35.3	40.0	27.3	30.3	64.0	79.0	53.0	36.0	52.7	45.3
Overall	59.2	58.0	61.1	46.3	39.6	90.4	94.5	120.6	81.8	82.5	

Table C.22 Average Number of Words Written by User Group with the Procedure in Tool Format

	A1	A2	A3	A4	A5	B1	B2	B3	B4	B5	Overall
Specialists	216.5	483.0	452.0	287.0	329.5	47.0	118.5	256.0	118.7	116.0	242.4
Developers	100.0	81.5	67.5	99.5	50.5	71.0	161.3	250.7	100.3	90.8	107.3
Content Providers	354.0	129.0	339.0	235.5	163.0	N/A	N/A	N/A	N/A	N/A	244.1
Internet Users	80.0	143.0	170.0	104.3	137.7	11.0	51.0	45.0	15.0	13.0	77.0
Overall	175.7	210.9	236.0	173.0	167.0	55.6	128.7	218.2	96.5	86.9	

Figure C.23 illustrates and compares the range of number of words written by each user group when using the procedure in document format. It shows the lowest, average, and highest values for each user group. Figure C.24 similarly presents the results for when using the procedure in tool format.

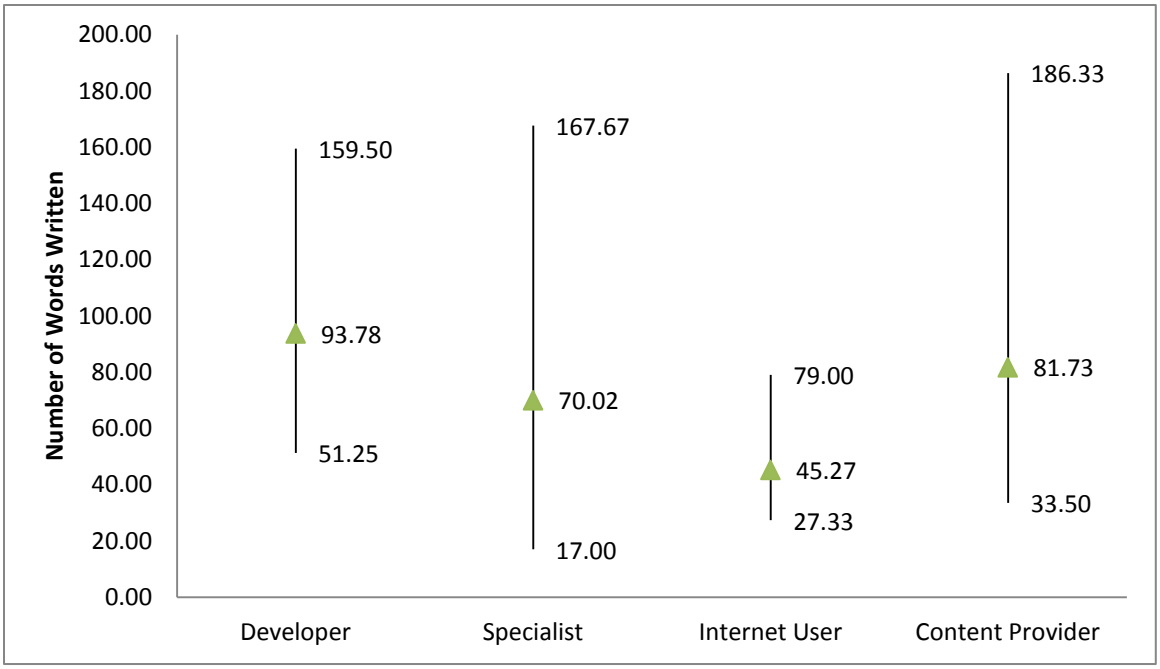


Figure C.23 Range of Number of Words Written by User Group with the Procedure in Document Format

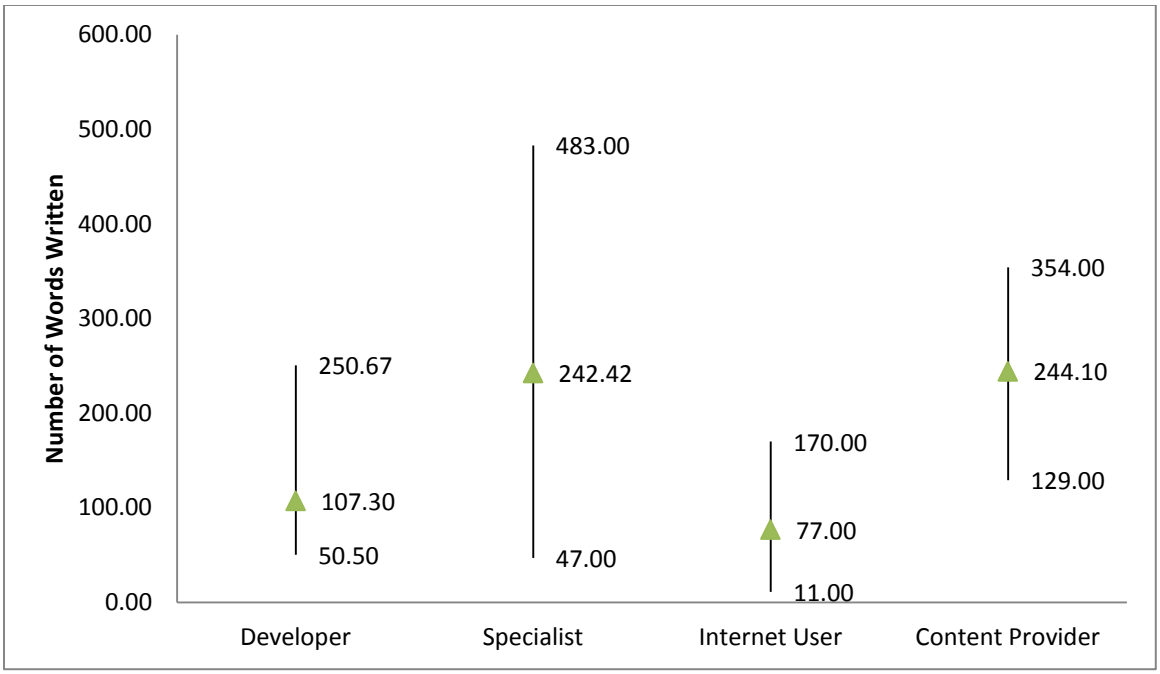


Figure C.24 Range of Number of Words Written by User Group with the Procedure in Tool Format

C.4 Quality of Information Identified Across Studies

In Research Study of the Identified Information, numerous pieces of information were rated by users as to the importance of each piece of information. There was a total of 395 pieces of information, with 94 of them rated as Essential, 84 rated as Significant, 167 rated as Helpful, and 50 rated as Not Important.

Table C.25 lists the percentage of times that question within the set of questions derived a piece of information that had an overall importance rating of Essential (denoted as E), Significant (denoted as S), Helpful (denoted as H), and Not Important (denoted as NI). Every question that derived information that was presented in the Research Study of the Identified Information is listed in the table. Questions that were not used have been removed.

Table C.25 Importance Level for Each Question

Question	E	S	H	NI
Why: What is the image being used for?	77.8%	0.0%	11.1%	11.1%
Why: What is the visual appeal of the decorative image?	33.3%	0.0%	33.3%	33.3%
What – Textual: What textual information is in the image/component?	40.0%	33.3%	20.0%	6.7%
What – Textual: What does the caption of the image say?	0.0%	0.0%	0.0%	100.0%
What – Classification: What kind of image is it?	90.0%	10.0%	0.0%	0.0%

What – Physical Object: What is the object in the image/component?	35.1%	35.1%	24.3%	5.4%
What – Physical Object: What is the brand / model / part name (number)?	44.4%	33.3%	22.2%	0.0%
What – Perceptual: What is the shape of the image/component?	27.3%	6.1%	63.6%	3.0%
What – Perceptual: What is the size or dimensions of the image/component?	0.0%	0.0%	66.7%	33.3%
What – Perceptual: What is (are) the colour(s) of the image/component?	26.3%	25.0%	21.1%	27.6%
What – Perceptual: What is the texture of the component?	0.0%	11.1%	88.9%	0.0%
What – Perceptual: What other perceptual information is important to know?	9.1%	9.1%	72.7%	9.1%
What – Subjective: What themes are represented in the image/component?	25.0%	25.0%	50.0%	0.0%
What – Subjective: What emotions are being expressed in the image/component?	25.0%	0.0%	75.0%	0.0%
What – Subjective: What emotions are meant to be felt by the viewer?	0.0%	0.0%	100.0%	0.0%
What – Subjective: What concepts are associated with the image/component?	14.3%	42.9%	42.9%	0.0%
What – Subjective: If the colour(s) of the image/component is important, what is the colour(s) representing?	16.7%	0.0%	83.3%	0.0%
What – Subjective: What is the image/component representing or symbolizing?	12.5%	18.8%	68.8%	0.0%
What – Subjective: What other subjective information is represented in the image/component?	10.0%	40.0%	50.0%	0.0%
What – Logical Relationships: What interaction is taking place?	0.0%	100.0%	0.0%	0.0%
What – Logical Relationships: What or who is the subject of the interaction?	0.0%	100.0%	0.0%	0.0%
What – Logical Relationships: What or who is the object of the interaction?	0.0%	100.0%	0.0%	0.0%
What – Logical Relationships: What is the intended result of the interaction?	100.0%	0.0%	0.0%	0.0%
What – Logical Relationships: How is the interaction being performed?	50.0%	0.0%	50.0%	0.0%
Who: Who is the image/component of?	100.0%	0.0%	0.0%	0.0%
Who: What is the person doing?	0.0%	100.0%	0.0%	0.0%
Who: What is the facial expression of the person?	0.0%	33.3%	66.7%	0.0%
Who: What does the person look like?	12.5%	18.8%	68.8%	0.0%
Who: What position is the person in?	0.0%	0.0%	100.0%	0.0%
Where – Location/Place: Where is the setting of the image/component?	50.0%	0.0%	50.0%	0.0%
Where – Location/Place: What specific landmarks are visible?	0.0%	100.0%	0.0%	0.0%
Where – Spatial Relationships: Where is the image/component spatially located?	0.0%	3.1%	56.3%	40.6%
Where – Spatial Relationships: Where is the image with regards to the content that contains it?	0.0%	0.0%	30.0%	70.0%
Where – Spatial Relationships: Where is the component relative to other entities or components?	6.9%	27.6%	44.8%	20.7%
When: When is the image taking place?	25.0%	50.0%	25.0%	0.0%

When: What season is represented?	0.0%	100.0%	0.0%	0.0%
When – Sequential Relationships: What is the basis of the sequential relationship?	100.0%	0.0%	0.0%	0.0%
When – Sequential Relationships: What types of sequential relationships are involved (linear, branching, cyclical, network, one directional, bi-directional)?	100.0%	0.0%	0.0%	0.0%
When – Sequential Relationships: What are the individual steps or components of the relationship?	100.0%	0.0%	0.0%	0.0%
When – Sequential Relationships: What is a suitable basis for logically ordering of the individual steps or components?	12.5%	12.5%	75.0%	0.0%
When – Sequential Relationships: What are the start and/or end point(s) to the set of relationships?	100.0%	0.0%	0.0%	0.0%
When – Sequential Relationships: How is each step or component related temporally to other steps or components?	87.5%	0.0%	12.5%	0.0%
How Much: What is the quantity associated with?	0.0%	100.0%	0.0%	0.0%
How Much: What is the quantity?	0.0%	91.7%	8.3%	0.0%
How Much: What is the unit of the quantity?	100.0%	0.0%	0.0%	0.0%
How Much: What is the precision (or statistical significance) of the quantity?	0.0%	0.0%	100.0%	0.0%
How Much: Is the quantity fixed or dynamic?	0.0%	100.0%	0.0%	0.0%
How Much: What does each axis represent?	0.0%	0.0%	100.0%	0.0%
How: How is the viewer supposed to interact or use the image/component?	16.7%	16.7%	66.7%	0.0%
How: What is intended to result from interacting with the image/component?	16.7%	33.3%	50.0%	0.0%

Table C.26 summarizes the average percentage of information from each importance level that was identified within each of the research studies. The percent of Essential information identified by each participant in each research study was calculated and then averaged. The same was done for Significant, Helpful, and Not Important information.

Table C.26 Averages of the Quality of Information Identified in Each Research Study

	Without Procedure	**With Document**	**With Tool**
Essential	31.8%	35.4%	65.1%
Significant	14.9%	24.8%	39.9%
Helpful	4.7%	14.1%	24.6%
Not Important	2.2%	11.0%	43.4%

Table C.27 summarizes the average percentage of information from each important level that was identified by each user group as well as for all user groups when the procedure was provided in document format. Table C.28 similarly summarizes the averages for when the procedure was provided in tool format.

Table C.27 Average Percentage of Information Identified by Each User Group when Using Procedure in Document Format

	Essential	Significant	Helpful	Not Important
Overall	35.4%	24.8%	14.1%	11.0%
Specialists	35.0%	25.4%	17.0%	20.1%
Developers	38.6%	29.8%	16.0%	5.7%
Content Providers	36.8%	22.0%	16.2%	15.3%
Internet Users	31.3%	21.6%	8.0%	5.0%

Table C.28 Average Percentage of Information Identified by Each User Group when Using Procedure in Tool Format

	Essential	Significant	Helpful	Not Important
Overall	65.1%	39.9%	24.7%	43.4%
Specialists	69.8%	54.2%	38.9%	64.1%
Developers	63.0%	34.6%	19.8%	33.3%
Content Providers	83.9%	40.4%	28.4%	54.4%
Internet Users	56.6%	32.2%	15.4%	31.9%

C.5 Ordering of Images by User Group

This section presents tables arranges the Set A and Set B images in decreasing order based on the number of questions answered, the amount of time spent, and the number of words written. The overall average values for each image reported in Sections C.1, C.2, and C.3 were used for the ordering.

C.5.1 Overall

Table C.29 lists each of the images in decreasing order of questions answered, time spent, and words written by all participants when the procedure was provided in document format. Table C.30 presents similar information for when the procedure was provided in tool format.

Table C.29 Ordering of Image in Decreasing Order for Specialists with the Use of the Procedure in Document Format

	Most									**Least**
Questions	B3	B2	A1	A5	A4	B1	B5	B4	A2	A3
Time	B1	B3	B2	A4	B4	A1	A5	A3	B5	A2
Words	B3	B2	B1	B5	B4	A3	A1	A2	A4	A5

Table C.30 Ordering of Image in Decreasing Order for Specialists with the Use of the Procedure in Tool Format

	Most									Least
Questions	B3	A4	A2	B2	A5	B4	A3	A1	B5	B1
Time	A2	B4	A3	A1	A4	B3	A5	B1	B5	B2
Words	A3	B3	A2	A1	A4	A5	B2	B4	B5	B1

C.5.2 Specialists

Table C.31 lists each of the images in decreasing order of questions answered, time spent, and words written by the Specialists when the procedure was provided in document format. Table C.32 presents similar information for when the procedure was provided in tool format.

Table C.31 Ordering of Image in Decreasing Order for Specialists with the Use of the Procedure in Document Format

	Most									Least
Questions	B3	B1	B2	B4	B5	A1	A5	A3	A4	A2
Time	B3	B1	B2	B4	B5	A1	A3	A5	A2	A4
Words	B3	B1	B2	B5	B4	A1	A2	A3	A4	A5

Table C.32 Ordering of Image in Decreasing Order for Specialists with the Use of the Procedure in Tool Format

	Most									Least
Questions	A2	A5	A4	A3	A1	B3	B4	B2	B5	B1
Time	A2	A3	A4	A5	A1	B3	B4	B5	B2	B1
Words	A2	A3	A5	A4	B3	A1	B4	B2	B5	B1

C.5.3 Developers

Table C.33 lists each of the images in decreasing order of questions answered, time spent, and words written by the Developers when the procedure was provided in document format. Table C.34 presents similar information for when the procedure was provided in tool format.

Table C.33 Ordering of Image in Decreasing Order for Developers with the Use of the Procedure in Document Format

	Most									Least
Questions	A5	A4	A1	B2	A2	A3	B4	B5	B1	B3
Time	A5	A4	A1	B2	B1	A3	B5	A2	B4	B3
Words	B4	B1	B2	B5	A2	A1	A4	A3	B3	A5

Table C.34 Ordering of Image in Decreasing Order for Developers with the Use of the Procedure in Tool Format

	Most									**Least**
Questions	B3	B2	B4	B5	B1	A1	A4	A3	A2	A5
Time	B4	B3	A1	B1	B2	B5	A2	A3	A4	A5
Words	B3	B2	B4	A1	A4	B5	A2	B1	A3	A5

C.5.4 Content Providers

Table C.35 lists each of the images in decreasing order of questions answered, time spent, and words written by the Content Providers when the procedure was provided in document format. Table C.36 presents similar information for when the procedure was provided in tool format. The images in Set B were not evaluated with the prototype tool; therefore, those images were not listed.

Table C.35 Ordering of Image in Decreasing Order for Content Providers with the Use of the Procedure in Document Format

	Most									**Least**
Questions	B3	B4	B5	A1	B1	B2	A2	A3	A5	A4
Time	B3	B1	B4	A3	B2	B5	A1	A2	A4	A5
Words	B3	B5	A3	B2	A1	B4	A2	B1	A5	A4

Table C.36 Ordering of Image in Decreasing Order for Content Providers with the Use of the Procedure in Tool Format

	Most				**Least**
Questions	A3	A1	A4	A2	A5
Time	A2	A3	A1	A4	A5
Words	A1	A3	A4	A5	A2

C.5.5 Internet Users

Table C.37 lists each of the images in decreasing order of questions answered, time spent, and words written by the Internet Users when the procedure was provided in document format. Table C.38 presents similar information for when the procedure was provided in tool format.

Table C.37 Ordering of Image in Decreasing Order for Internet Users with the Use of the Procedure in Document Format

	Most									**Least**
Questions	B5	A1	B1	B2	A2	A3	B4	A4	A5	B3
Time	B2	B3	B1	A4	A1	B4	A3	A2	B5	A5
Words	B2	B1	B3	B5	A3	B4	A2	A1	A5	A4

Table C.38 Ordering of Image in Decreasing Order for Internet Users with the Use of the Procedure in Tool Format

	Most									**Least**
Questions	A4	A5	A2	A3	B2	A1	B3	B4	B5	B1
Time	A2	A3	A1	A4	A5	B3	B2	B5	B1	B4
Words	A3	A2	A5	A4	A1	B2	B3	B4	B5	B1

C.6 Questions Improvement Between Research Studies

The results presented in this section were calculated by calculating the number and percentage of participants in each research study who identified a piece of information derived from a particular question. These values were then compared between the research studies to determine the number of instances and the particular questions where a particular research study had a higher value than another. A percentage was calculated for each question in the comparison to determine how much improvement was made for a particular question.

The tables in this section consist of three columns. The first column states the category, sub-category (if applicable), and the question concerned. The second column states the percentage that a piece of information derived from that question was identified (or missed) more often. The third column states the number of instances that a piece of information derived from that question was derived (or missed) more often, along with the total number of pieces of information that was derived by that question.

C.6.1 Comparison Between Without a Procedure and Procedure in Document Format

Table C.39 shows the results of the comparison between Research Study of Text Alternatives Without the Procedure and Research Study of the Procedure in Document Format. It lists the instances where participants identified a piece of information more frequently while using the procedure in the document format than without a procedure.

Table C.39 Questions where the Procedure in Document Format Improved (Sorted by Percentage of Instances Improved)

Question	Percentage	Instances
How Much: What is the quantity?	100.0%	12/12
What – Subjective: What other subjective information is represented in the image/component?	100.0%	5/5
What – Subjective: What themes are represented in the image/component?	100.0%	4/4
Who: What is the facial expression of the person?	100.0%	3/3
When – Sequential Relationship: What are the individual steps or components of the relationship?	100.0%	2/2
What – Logical Relationship: What interaction is taking place?	100.0%	1/1
What – Logical Relationship: What or who is the object of the interaction?	100.0%	1/1
Who: What position is the person in?	100.0%	1/1
When – Sequential Relationship: What is the basis of the sequential relationship?	100.0%	1/1
When – Sequential Relationship: What types of sequential relationships are involved (linear, branching, cyclical, network, one directional, bi-directional)?	100.0%	1/1
How Much: What does each axis represent?	100.0%	1/1
What – Subjective: What is the image/component representing or symbolizing?	81.3%	13/16
What – Textual: What textual information is in the image/component?	80.0%	12/15
When – Sequential Relationship: What is a suitable basis for logically ordering of the individual steps or components?	75.0%	6/
When: When is the image taking place?	75.0%	3/4
Why: What is the image being used for?	66.7%	6/9
What – Perceptual: What is the texture of the component?	66.7%	6/9
What – Subjective: If the colour(s) of the image/component is important, what is the colour(s) representing?	66.7%	4/6
What – Perceptual: What is (are) the colour(s) of the image/component?	65.8%	25/38
When – Sequential Relationship: How is each step or component related temporally to other steps or components?	62.5%	5/8
What – Physical Object: What is the object in the image/component?	62.2%	23/37
Where – Spatial Relationship: Where is the image with regards to the content that contains it?	60.0%	6/10
Where – Spatial Relationship: Where is the component relative to other entities or components?	58.6%	17/29
What – Subjective: What concepts are associated with the image/component?	57.1%	4/7
What – Classification: What kind of image is it?	50.0%	5/10
Who: What does the person look like?	50.0%	4/8
How: What is intended to result from interacting with the image/component?	50.0%	3/6
What – Subjective: What emotions are being expressed in the image/component?	50.0%	2/4
Who: Who is the image/component of?	50.0%	1/2
Who: What is the person doing?	50.0%	1/2
When – Sequential Relationship: What are the start and/or end point(s) to the set of relationships?	50.0%	1/2

What – Perceptual: What is the shape of the image/component?	48.5%	16/33
What – Physical Object: What is the brand / model / part name (number)?	44.4%	4/9
Where – Spatial Relationship: Where is the image/component spatially located?	37.5%	12/32
Why: What is the visual appeal of the decorative image?	33.3%	1/3
What – Subjective: What emotions are meant to be felt by the viewer?	25.0%	1/4
Where: Where is the setting of the image/component?	25.0%	1/4
What – Perceptual: What other perceptual information is important to know?	9.1%	1/11
What – Perceptual: What is the size or dimensions of the image/component?	5.6%	1/18

Table C.40 shows the results of the comparison between Research Study of Text Alternatives Without the Procedure and Research Study of the Procedure in Document Format. It lists the instances where participants identified a piece of information less frequently while using the procedure in the document format than without a procedure.

Table C.40 Questions where the Procedure in Document Format Did Not Improve (Sorted by Percentage of Instances Missed)

Question	Percentage	Instances
Who: Who is the image/component of?	100.0%	2/2
How Much: What is the quantity associated with?	100.0%	1/1
How Much: What is the unit of the quantity?	100.0%	1/1
Classification: What kind of image is it?	50.0%	5/10
Where – Location/Place: Where is the setting of the image/component?	50.0%	2/4
When: What season is represented?	50.0%	1/2
Who: What does the person look like?	37.5%	3/8
What – Perceptual: What is the size or dimensions of the image/component?	33.3%	6/18
Why: What is the image being used for?	33.3%	3/9
Why: What is the visual appeal of the decorative image?	33.3%	1/3
What – Subjective: What emotions are meant to be felt by the viewer?	25.0%	1/4
When: When is the image taking place?	25.0%	1/4
What – Physical Object: What is the object in the image/component?	24.3%	9/37
What – Physical Object: What is the brand / model / part name (number)?	22.2%	2/9
What – Subjective: What is the image/component representing or symbolizing?	18.8%	3/16
What – Subjective: If the colour(s) of the image/component is important, what is the colour(s) representing?	16.7%	1/6
What – Perceptual: What is (are) the colour(s) of the image/component?	15.8%	6/38
What – Subjective: What concepts are associated with the image/component?	14.3%	1/7
What – Perceptual: What is the texture of the component?	11.1%	1/9
Where – Spatial Relationships: Where is the image with regards to the content that contains it?	10.0%	1/10

What – Perceptual: What other perceptual information is important to know?	9.1%	1/11
Where – Spatial Relationships: Where is the component relative to other entities or components?	6.9%	2/29
Where – Spatial Relationships: Where is the image/component spatially located?	3.1%	1/32

C.6.2 Comparison Between Without a Procedure and Procedure in Tool Format

Table C.41 shows the results of the comparison between Research Study of Text Alternatives Without the Procedure and Research Study of the Procedure as a Prototype Tool. It lists the instances where participants identified a piece of information more frequently while using the procedure in the tool format than without a procedure.

Table C.41 Questions where the Procedure in Tool Format Improved Over Without a Procedure (Sorted by Percentage of Instances Improved)

Question	Percentage	Instances
What – Subjective: What themes are represented in the image/component?	100.0%	4/4
What – Subjective: What emotions are being expressed in the image/component?	100.0%	4/4
When: When is the image taking place?	100.0%	4/4
Why: What is the visual appeal of the decorative image?	100.0%	3/3
Who: What is the facial expression of the person?	100.0%	3/3
What – Logical Relationships: How is the interaction being performed?	100.0%	2/2
Where – Location/Place: What specific landmarks are visible?	100.0%	2/2
When: What season is represented?	100.0%	2/2
When – Sequential Relationships: What are the start and/or end point(s) to the set of relationships?	100.0%	2/2
What – Logical Relationships: What interaction is taking place?	100.0%	1/1
What – Logical Relationships: What or who is the subject of the interaction?	100.0%	1/1
What – Logical Relationships: What or who is the object of the interaction?	100.0%	1/1
What – Logical Relationships: What is the intended result of the interaction?	100.0%	1/1
Who: What position is the person in?	100.0%	1/1
When – Sequential Relationships: What is the basis of the sequential relationship?	100.0%	1/1
When – Sequential Relationships: What types of sequential relationships are involved (linear, branching, cyclical, network, one directional, bi-directional)?	100.0%	1/1
How Much: What is the quantity associated with?	100.0%	1/1
How Much: What is the unit of the quantity?	100.0%	1/1
How Much: What is the precision (or statistical significance) of the	100.0%	1/1

quantity?		
How Much: Is the quantity fixed or dynamic?	100.0%	1/1
Textual: What textual information is in the image/component?	93.3%	14/
Where – Spatial Relationships: Where is the image with regards to the content that contains it?	90.0%	9/10
Why: What is the image being used for?	88.9%	8/9
What – Perceptual: What is the texture of the component?	88.9%	8/9
What – Subjective: What concepts are associated with the image/component?	85.7%	6/7
How Much: What is the quantity?	83.3%	10/12
How: What is intended to result from interacting with the image/component?	83.3%	5/6
Classification: What kind of image is it?	80.0%	8/10
What – Subjective: What other subjective information is represented in the image/component?	80.0%	4/5
Where – Spatial Relationships: Where is the component relative to other entities or components?	79.3%	23/29
Where – Spatial Relationships: Where is the image/component spatially located?	78.1%	25/32
What – Perceptual: What is the size or dimensions of the image/component?	77.8%	14/18
What – Physical Objects: What is the brand / model / part name (number)?	77.8%	7/9
What – Subjective: What is the image/component representing or symbolizing?	75.0%	12/16
When – Sequential Relationships: How is each step or component related temporally to other steps or components?	75.0%	6/8
What – Perceptual: What is the shape of the image/component?	72.7%	24/33
What – Subjective: If the colour(s) of the image/component is important, what is the colour(s) representing?	66.7%	4/6
How: How is the viewer supposed to interact or use the image/component?	66.7%	4/6
What – Perceptual: What is (are) the colour(s) of the image/component?	65.8%	25/38
What – Physical Objects: What is the object in the image/component?	59.5%	22/37
Who: What does the person look like?	50.0%	4/8
What – Subjective: What emotions are meant to be felt by the viewer?	50.0%	2/4
Where – Location/Place: Where is the setting of the image/component?	50.0%	2/4
Who: Who is the image/component of?	50.0%	1/2
Who: What is the person doing?	50.0%	1/2
When – Sequential Relationships: What are the individual steps or components of the relationship?	50.0%	1/2
When – Sequential Relationships: What is a suitable basis for logically ordering of the individual steps or components?	37.5%	3/8
What – Perceptual: What other perceptual information is important to know?	27.3%	3/11

Table C.42 shows the results of the comparison between Research Study of Text Alternatives Without the Procedure and Research Study of the Procedure as a Prototype Tool. It

lists the instances where participants identified a piece of information less frequently while using the procedure in the tool format than without a procedure.

Table C.42 Questions where the Procedure in Tool Format Did Not Improve Over Without a Procedure (Sorted by Percentage of Instances Missed)

Question	Percentage	Instances
How Much: What does each axis represent?	100.0%	1/1
Who: What does the person look like?	37.5%	3/8
What – Subjective: What emotions are meant to be felt by the viewer?	25.0%	1/4
Where – Location/Place: Where is the setting of the image/component?	25.0%	1/4
What – Physical Objects: What is the object in the image/component?	18.9%	7/37
How Much: What is the quantity?	16.7%	2/12
What – Subjective: If the colour(s) of the image/component is important, what is the colour(s) representing?	16.7%	1/6
When – Sequential Relationships: How is each step or component related temporally to other steps or components?	12.5%	1/8
What – Physical Objects: What is the brand / model / part name (number)?	11.1%	1/9
What – Perceptual: What is the texture of the component?	11.1%	1/9
What – Perceptual: What is (are) the colour(s) of the image/component?	10.5%	4/38
What – Perceptual: What other perceptual information is important to know?	9.1%	1/11
Where – Spatial Relationships: Where is the component relative to other entities or components?	6.9%	2/29
What – Subjective: What is the image/component representing or symbolizing?	6.3%	1/16
What – Perceptual: What is the size or dimensions of the image/component?	5.6%	1/18
Where – Spatial Relationships: Where is the image/component spatially located?	3.1%	1/32

C.6.3 Comparison Between Procedure in Document Format and Procedure in Tool Format

Table C.43 shows the results of the comparison between Research Study of the Procedure in Document Format and Research Study of the Procedure as a Prototype Tool. It lists the instances where participants identified a piece of information more frequently while using the procedure in the tool format than the document format.

Table C.43 Questions where the Procedure in Tool Format Improved Over the Procedure in Document Format (Sorted by Percentage of Instances Improved)

Question	Percentage	Instances
Classification: What kind of image is it?	100.0%	10/10
Why: What is the image being used for?	100.0%	9/9
What – Subjective: What concepts are associated with the image/component?	100.0%	7/7
How: What is intended to result from interacting with the image/component?	100.0%	6/6
What – Subjective: What themes are represented in the image/component?	100.0%	4/4
What – Subjective: What emotions are being expressed in the image/component?	100.0%	4/4
Why: What is the visual appeal of the decorative image?	100.0%	3/3
What – Logical Relationships: How is the interaction being performed?	100.0%	2/2
Who: Who is the image/component of?	100.0%	2/2
Where – Location/Place: What specific landmarks are visible?	100.0%	2/2
When: What season is represented?	100.0%	2/2
Textual: What does the caption of the image say?	100.0%	1/1
What – Logical Relationships: What interaction is taking place?	100.0%	1/1
What – Logical Relationships: What or who is the subject of the interaction?	100.0%	1/1
What – Logical Relationships: What or who is the object of the interaction?	100.0%	1/1
What – Logical Relationships: What is the intended result of the interaction?	100.0%	1/1
Who: What position is the person in?	100.0%	1/1
When – Sequential Relationships: What is the basis of the sequential relationship?	100.0%	1/1
When – Sequential Relationships: What types of sequential relationships are involved (linear, branching, cyclical, network, one directional, bi-directional)?	100.0%	1/1
How Much: What is the quantity associated with?	100.0%	1/1
How Much: What is the unit of the quantity?	100.0%	1/1
How Much: What is the precision (or statistical significance) of the quantity?	100.0%	1/1
How Much: Is the quantity fixed or dynamic?	100.0%	1/1
Who: What does the person look like?	87.5%	7/8
How: How is the viewer supposed to interact or use the image/component?	83.3%	5/6
What – Subjective: What is the image/component representing or symbolizing?	81.3%	13/16
Where – Spatial Relationships: Where is the image with regards to the content that contains it?	80.0%	8/10
What – Perceptual: What is the size or dimensions of the image/component?	77.8%	14/18
What – Physical Objects: What is the brand / model / part name (number)?	77.8%	7/9
Where – Spatial Relationships: Where is the image/component spatially located?	75.0%	24/32
When – Sequential Relationships: How is each step or component related temporally to other steps or components?	75.0%	6/8
What – Subjective: What emotions are meant to be felt by the viewer?	75.0%	3/4

Where – Location/Place: Where is the setting of the image/component?	75.0%	3/4
When: When is the image taking place?	75.0%	3/4
Textual: What textual information is in the image/component?	73.3%	11/15
What – Perceptual: What is the shape of the image/component?	66.7%	22/33
What – Subjective: If the colour(s) of the image/component is important, what is the colour(s) representing?	66.7%	4/6
Who: What is the facial expression of the person?	66.7%	2/3
Where – Spatial Relationships: Where is the component relative to other entities or components?	62.1%	18/29
What – Perceptual: What is (are) the colour(s) of the image/component?	60.5%	23/38
What – Physical Objects: What is the object in the image/component?	56.8%	21/37
Who: What is the person doing?	50.0%	1/2
When – Sequential Relationships: What are the individual steps or components of the relationship?	50.0%	1/2
When – Sequential Relationships: What are the start and/or end point(s) to the set of relationships?	50.0%	1/2
What – Perceptual: What is the texture of the component?	44.4%	4/9
What – Perceptual: What other perceptual information is important to know?	27.3%	3/11
What – Subjective: What other subjective information is represented in the image/component?	20.0%	1/5
How Much: What is the quantity?	8.3%	1/12

Table C.44 shows the results of the comparison between Research Study of the Procedure in Document Format and Research Study of the Procedure as a Prototype Tool. It lists the instances where participants identified a piece of information less frequently while using the procedure in the tool format than the document format.

Table C.44 Questions where the Procedure in Tool Format Did Not Improve Over the Procedure in Document Format (Sorted by Percentage of Instances Missed)

Question	Percentage	Instances
How Much: What does each axis represent?	100.0%	1/1
How Much: What is the quantity?	91.7%	11/12
What – Subjective: What other subjective information is represented in the image/component?	80.0%	4/5
When – Sequential Relationships: What is a suitable basis for logically ordering of the individual steps or components?	75.0%	6/8
What – Perceptual: What is the texture of the component?	55.6%	5/9
When – Sequential Relationships: What are the individual steps or components of the relationship?	50.0%	1/2
When – Sequential Relationships: What are the start and/or end point(s) to the set of relationships?	50.0%	1/2
What – Physical Objects: What is the object in the image/component?	40.5%	15/37
What – Perceptual: What is (are) the colour(s) of the image/component?	36.8%	14/38
What – Subjective: If the colour(s) of the image/component is important,	33.3%	2/6

what is the colour(s) representing?		
Who: What is the facial expression of the person?	33.3%	1/3
Where – Spatial Relationships: Where is the component relative to other entities or components?	31.0%	9/29
Textual: What textual information is in the image/component?	26.7%	4/15
When – Sequential Relationships: How is each step or component related temporally to other steps or components?	25.0%	2/8
What – Subjective: What emotions are meant to be felt by the viewer?	25.0%	1/4
Where – Location/Place: Where is the setting of the image/component?	25.0%	1/4
When: When is the image taking place?	25.0%	1/4
What – Physical Object: What is the brand / model / part name (number)?	22.2%	2/9
What – Perceptual: What is the shape of the image/component?	21.2%	7/33
What – Subjective: What is the image/component representing or symbolizing?	18.8%	3/16
What – Perceptual: What other perceptual information is important to know?	18.2%	2/11
Where – Spatial Relationships: Where is the image/component spatially located?	12.5%	4/32
Who: What does the person look like?	12.5%	1/8
Where – Spatial Relationships: Where is the image with regards to the content that contains it?	10.0%	1/10
What – Perceptual: What is the size or dimensions of the image/component?	5.6%	1/18

Appendix D ISO/IEC TS 20071-11: Guidance on Text Alternatives for Images [31]

ISO/IEC JTC 1/SC 35 N

DATE: 2012-08-30

ISO/IEC JTC 1/SC 35
User Interfaces
Secretariat: AFNOR

DOC TYPE:	**TS text**
TITLE:	**TS 20071-11** **Information technology — User interface component accessibility — Part 11: Guidance for alternative text for images**
SOURCE:	**WG6**
STATUS:	
ACTION ID:	**Text for TS publication**
DUE DATE:	**2012-**
DISTRIBUTION:	**Open**
MEDIUM:	**E**
NO. OF PAGES:	**37**

Secretariat of ISO/IEC JTC 1/SC 35 AFNOR — Philippe Magnabosco
11 rue Francis de Pressensé 93571 - La Plaine Cedex Saint-Denis - France
Telephone: +33 1 41 62 85 02; Facsimile: 33 1 49 17 90 00;
e-mail: philippe.magnabosco@afnor

ISO/IEC TC JTC1/SC 35 N

Date: 2012-08-30

ISO/IEC PDTR 20071-11

ISO/IEC TC JTC1/SC 35/WG 6

Secretariat: AFNOR

Information technology — User interface component accessibility — Part 11: Guidance for alternative text for images

Warning

This document is not an ISO International Standard. It is distributed for review and comment. It is subject to change without notice and may not be referred to as an International Standard.

Recipients of this draft are invited to submit, with their comments, notification of any relevant patent rights of which they are aware and to provide supporting documentation.

Contents

Page

Foreword

- ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

- International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

- The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

- In other circumstances, particularly when there is an urgent market requirement for such documents, the joint technical committee may decide to publish an ISO/IEC Technical Specification (ISO/IEC TS), which represents an agreement between the members of the joint technical committee and is accepted for publication if it is approved by 2/3 of the members of the committee casting a vote.

- An ISO/IEC TS is reviewed after three years in order to decide whether it will be confirmed for a further three years, revised to become an International Standard, or withdrawn. If the ISO/IEC TS is confirmed, it is reviewed again after a further three years, at which time it must either be transformed into an International Standard or be withdrawn.

- Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

- ISO/IEC TS 20071-11 was prepared by Technical Committee ISO/TC JTC1, *Information technology*, Subcommittee SC 35, *User interfaces*.

- This second/third/... edition cancels and replaces the first/second/... edition (), [clause(s) / subclause(s) / table(s) / figure(s) / annex(es)] of which [has / have] been technically revised.

- ISO/IEC TS 20071 consists of the following parts, under the general title *Information technology — User interface component accessibility*:

Part 11: Guidance for text alternatives for images

Part [n]:

Part [n+1]:

Introduction

The saying that "A picture is worth a thousand words" recognizes that images can present a wealth of information. It is important that alternative textual descriptions or representations be used to present the user with a comprehensive account of the purpose and content of images to people unable to see or interpret them.

Text alternatives help people who cannot see the images to understand what the image is of or the purpose it serves by providing the same information in textual form. Text alternatives can be useful to those with visual impairments, those who turned images off in order to improve webpage loading speeds, and those who cannot understand the image being displayed. This document provides guidance for web and document developers to help them create informative descriptions for various types of illustrations.

Information technology — User interface component accessibility — Part 11: Guidance for alternative text for images

1 Scope

This part of ISO/IEC 20071 applies to all static images that are used in any type of electronic document. It also applies to individual images within a slide show of electronic images.

NOTE While text alternatives can be implemented via various mechanisms in various types of electronic documents, the contents of this technical report are not dependent on the choice of implementation mechanism or of electronic document type.

This part of ISO/IEC 20071 does not apply to moving images (e.g. movies).

The guidance contained in this part of ISO/IEC 20071 is intended to be used by the person who creates content to be placed in an electronic document. There is no expectation that this person will have any additional expertise beyond understanding the contents of the document and why an image was chosen to be placed within the document.

While the main intent of the guidance within this part of ISO/IEC 20071 is the creation of text alternatives, the information identified in this guidance could be placed in the main document text, reducing the length of the resulting text alternatives. However, placing information in the main document text does not fully replace the function of having some text alternatives for each image.

2 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

2.1 Text alternatives and related definitions

2.1.1
text alternative
a textual description or representation of an image

NOTE 1 By storing this description or representation in text format, it is able to be rendered in any available modality.

NOTE 2 The main audience of text alternatives is the users of screen reading features.

NOTE 3 Text alternatives are often provided to screen reader users in the form of primary and secondary alternative texts of an image.

2.1.2
primary alternative text
main text alternative provided to users of screen readers

NOTE: Different technologies and platforms provide various mechanisms for containing and presenting primary alternative text.

EXAMPLE In XHTML, HTML4, and HTML5, primary alternative text is provided in the “alt” attribute of the img tag.

EXAMPLE In Flash™, primary alternative text is provided through the “Name” field.

EXAMPLE In PDF, primary alternative text is provided through the /Alt entry in a structure element’s dictionary.

2.1.3
secondary alternative text
additional text alternative provided to users of screen readers beyond primary alternative text

NOTE: Different technologies and platforms provide various mechanisms for containing and presenting secondary alternative text.

EXAMPLE In XHTML and HTML4, secondary alternative text is provided in the “longdesc” attribute of the img tag.

EXAMPLE In Flash™, secondary alternative text is provided through the “Description” field.

EXAMPLE In EPUB, secondary alternative text is provided through the “describedAt” attribute.

2.1.4
main document text
textual content of a document that is always presented to the users

2.2 Image and related definitions

2.2.1
image
<digital> graphical content intended to be presented visually

NOTE This includes graphics that are encoded in any electronic format, including, but not limited to) formats that are comprised of individual pixels (e.g. those produced by paint programs or by photographic means) and formats that comprised of formulas (e.g. those produced as scalable vector drawings).

2.2.2
static image
image where the set of image components and their relationships to one another do not change over time

NOTE 1 This includes images where the content / representation of individual image components might change over time, e.g. indicators where the value they are indicating changes in real time.

NOTE 2 The concept of static image is used for all images that are not slide shows or moving images.

NOTE 3 This use of static image is similar to the ISO/IEC 13249-5 use of "still image". However, it differs in that a static image might have moving components. ISO/IEC 13249-5 states "A still image user-defined type is generic to image handling. It addresses the need to store, manage, and retrieve information based on aspects of inherent image characteristics such as height, width, and format and based on image features such as average color, color histogram, positional color, and texture. It also addresses the need to employ manipulation such as rotation, scaling as well as similarity assessment."

2.2.3
slide show
set of images that replace one another periodically

NOTE 1 The replacement of one static image by another static image can be controlled automatically by the system (in which case the timing for each image is usually predetermined) or manually by the user (where the timing for each image is determined on a case by case basis).

NOTE 2 Slide shows are usually composed of static images, but might include short movies. The interval between static images in a slide show are considered longer than in a movie, such that the motion being portrayed by the slide show would appear staggered instead of smooth like in a movie.

2.2.4
moving image
image where the contents are dynamically changing

NOTE This includes realistic moving images (often referred to as movies), abstract moving images (often referred to as cartoons), and even non-representational moving images (often referred to as light shows).

2.2.5
component
<image> logical part of an image that provides important content that the user should be aware of

NOTE 1 Types of image component include (but not limited to) shapes, objects, persons, areas, and text.

NOTE 2 Text components can include natural and/or formal languages (such as mathematical equations).

2.3 Importance and related definitions

2.3.1
importance
<information about an image> level of need for users to know information about an image (or image component)

2.3.2
essential
<information about an image> information that is necessary to understand the image and/or its function within the document

2.3.3
significant
<information about an image> information that is needed for a more detailed understanding of the image within the document to most users most of the time

2.3.4
helpful
<information about an image> information that provides a thorough understanding to target audiences regarding the image within the document

2.3.5
not important
<information about an image> information that does not help provide additional understanding about the image within the document to users

2.4 Information relationship definitions

2.4.1
relationship type
information about an association between entities

2.4.2
logical
<relationship> information about what entities are interacting and how they interact

2.4.3
temporal
<relationship> information about when some action or entity occurs

2.4.4
physical
spatial
<relationship> information about where one entity is in relation to another entity

2.5 Image content and related definitions

2.5.1
content
<image> data, information, objects, relationships, and/or concepts to be communicated from the originator to the user according to certain communication goals

NOTE 1 Adapted from ISO 14915-1 definition 3.1.

NOTE 2 Content can be presented in realistic, abstract, or even non-representational manners. The distinction between these types of presentation is how closely they represent the natural world.

2.5.2
realistic
<type of image> image perceived by the user to faithfully represent data, information, objects, relationships, and/or concepts in the natural world

NOTE Adapted from ISO 14915-3 definition 3.7.5.

EXAMPLE Photographic images, pictures intended to be true-to-life, diagrams used to illustrate how to assemble a set of parts

2.5.3
abstract
<type of image> image intended to present important major data, information, object, relationship, and/or conceptual components, without faithfully representing them as they occur in the natural world

EXAMPLE Cartoons, abstract art (where the basis for abstraction can be recognized), graphs and charts

2.5.4
non-representational
<type of image> image intended for decorative purposes without the intent to represent any particular natural world data, information, objects, relationships, and/or concepts

EXAMPLE Art presenting colors and textures (without any recognizable objective contents)

2.6 Image sources and related definitions

2.6.1
source
<image> means of obtaining an image (or image component)

2.6.2
photograph
<image> electronic copy of an image of something that has its own independent existence in the real world

NOTE While the photograph is a true rendering of the object(s) that it illustrates, those objects can themselves be realistic, abstract, or even non-representational.

2.6.3
electronic drawing
<image> image created as an original work to be rendered on the computer

NOTE Electronic drawings can be realistic, abstract / non-representational and can contain realistic, abstract / non-representational components. They can even contain embedded photographs as components.

2.7 Information type definitions

2.7.1
physical
<information within an image> information about phenomena which have a concrete existence; objects, agents, or scenes that have a physical existence

[14915-3 definition 3.6.7]

NOTE This can include states and histories of objects.

2.7.2
value
<information within an image> quantitative information describing properties of an object

[14915-3 definition 3.6.12]

2.7.3
quantitative
<information within an image> statistical information or numerical data and the relationships between the numbers

NOTE 1 Quantitative information is often presented in a graphical manner

NOTE 2 Quantitative images are often used for comparison between related sets of data, such as comparing net profit over a period of time.

NOTE 3 Examples of quantitative images include charts and graphs.

2.7.4
control
<information within an image> information that can be used to take some action which manipulates data, other objects or their attributes

NOTE Adapted from ISO 14915-2 definition 3.8.

2.7.5
event
<information within an image> information about a state change, message indicating the occurrence of an action, or conveying a significant change in the world

[14915-3 definition 3.6.6]

2.7.6
state
<information within an image> properties of the environment, objects or agents that remain constant during a period of time

[14915-3 definition 3.6.11]

3 Framework for describing images and creating text alternatives

3.1 Uses of text alternatives

Images are often used to convey a large amount of information quickly, whether it is a diagram for constructing a desk or a photo of what happened at a birthday party. In one glance, a person can retrieve a large amount of information and have a general understanding about the remaining content in the document in which the image resides.

Images are sometimes used to supplement or complement the document content or can be another representation of the same content. However, sometimes the image stands alone or adds information that is not part of the other document content. The information that is present in the image but not the other document content does not get conveyed to those who are unable to see the image. Text alternatives are needed to convey that information.

There are many reasons why a person might need text alternatives, including (but not limited to)

a) The person has a visual impairment,
b) The person is using a program that aurally reads the document content while doing something else (such as driving or cooking),
c) The device being used to view the image is unable to properly display the image or the image is difficult to see, (such as on a mobile device),
d) The person turned off images on their Web browser to increase loading speed, and
e) The person cannot understand and/or interpret the image.

Tools (such as screen readers) exist that can read aloud text that appears on a document to those who cannot or are not looking at the screen. If an image can be described and represented textually, then the tools can also read the text alternatives aloud.

Text alternatives might include a description of what the image looks like and/or an interpretation of what the image represents or its function. Different text alternatives might exist for the same image, differing in length and (as a result) information. Technology often allows for a primary alternative text as well as a secondary alternative text. Providing both primary alternative text and secondary alternative text can give the user a choice in the amount of detail they wish to receive about an image.

3.2 Structure of descriptions

In order to write informative text alternatives for an image, it is important to first know the information being represented in the image. It is difficult to share knowledge about an image with others when the writer does not have knowledge of what the image is. Therefore, it is important to gather or identify as much information as possible about an image.

To gather as much information as possible, the following image description structure is used. Initially, the structure identifies information about the whole image. It focuses on information that applies to the entire image. Because an image can have a vast amount of information, the structure breaks down the image into several parts (called image components) and focuses on identifying information about each part (or image component). Breaking down the image into several parts allows for focus on the details of those particular parts, resulting in more information about the image. An image component could be broken down further into additional image components.

The structure of image information (illustrated in Figure 1) for the Whole Image and each image component includes a name, the purpose, the basic image information or content, and the elaboration of the basic image information. The name identifies the part of the image that the information applies. The purpose identifies why the image (or image component) exists on the document. The basic image information or content identifies what is in the image (or image component). The elaboration of the basic information looks at more specific details about the image (or image component).

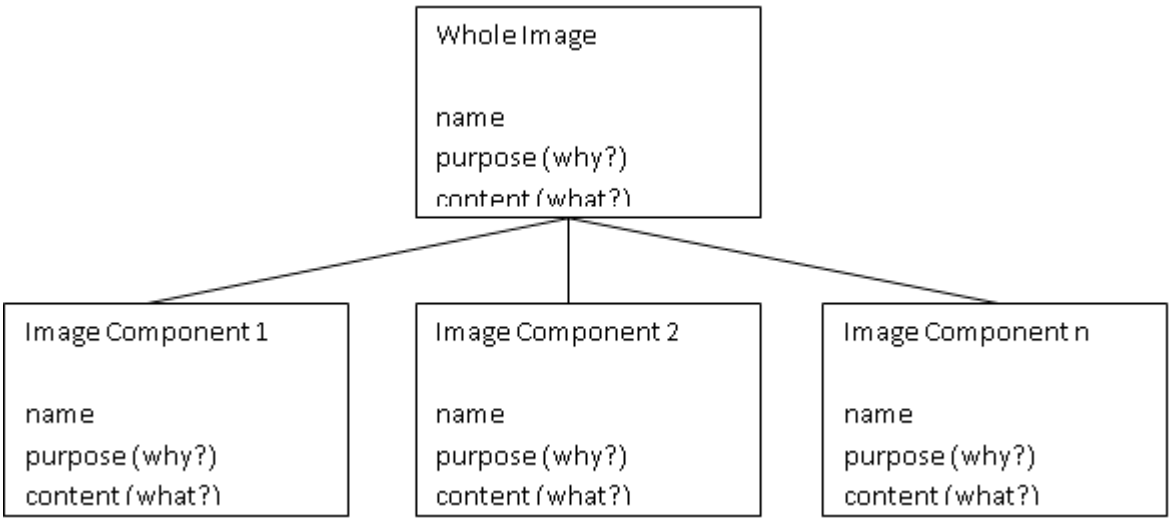


Figure 1 — Structure of image information

The intent of the structure is to identify as much information about an image as possible. The same piece of information might appear multiple times. A piece of information identified at the Whole Image level might be identified again at the image component level or it may be identified as part of multiple image components. It is not necessary for the information to be unique.

3.3 Structure of text alternatives

From the identified image information, the text alternatives are composed. The structure presented in this document strives to generate informative text alternatives for images. The two main kinds of text alternatives are the primary alternative text and the secondary alternative text.

The primary alternative text is sometimes automatically displayed (in a hidden manner) to tools such as screen readers. There is generally a limitation on the length of the primary alternative text. Different tools suggest or impose different limits. For example, for Web pages, it is sometimes suggested that it be less than 125 characters. Given the limited number of characters, the amount of information that primary alternative text can contain is also limited. Therefore, the primary alternative text is often viewed as an overview of the image content, what the image is about, or the function of the image.

The secondary alternative text, on the other hand, generally does not have limits on the length of the description and therefore can contain a larger amount of information about the image. It can consist of details about the image that could not be part of the primary alternative text. Since there is no limit on the amount of information, it can include information that some users might not need.

The structure of text alternatives should be based on the importance of each piece of image information. This importance can help to determine if the information is presented in the primary alternative text, secondary alternative text, both primary and secondary alternative texts, or not at all.

The context in which the image is used is a central concept in this framework. The same image can be used for different reasons or purposes. Depending on the purpose and context, different pieces of information become important or unimportant. Therefore, different text alternatives (both primary and secondary alternative texts) might be created for a single image.

4 Requirements and recommendations for describing images

4.1 A procedure for creating text alternatives

Text alternatives should be created by the person creating the main content of the document and/or selecting the image(s) for use in the document. While various experts might analyze an image to a

greater extent (e.g. for cataloguing purposes in a library of images), the important expertise involved in this procedure is understanding the information the image is to convey within its document.

The creation of suitable text alternatives (both the primary and secondary alternative texts) should be based on a thorough understanding of the image, its components, and its purpose in the document within which it is contained. This can be done by applying the following procedure:

a) Identify the purpose of the image

Identify and document the purpose of each image in accordance with clause 4.2. This step influences which image components and information are important for the user to know.

This involves answering the question "Why?"

b) Identify of the image components

Depending on the purpose of the image, identify the image components in accordance with clause 6.4. This step is necessary to properly identify information about the image that might be important to the user.

This can be done in a two stage process:

Identify the image as a whole

Identify the image components of the image

NOTE Identifying image components is an iterative process. Individual image components can be further separated into a number of (lower level) image components until all components that are important to describe have been identified.

c) Identify the image (or image component) content

Depending on the purpose of the image and the importance of the image (or image components), identify the content about an image and its components in accordance with clause 4.5.

This involves answering the question "What?"

NOTE 1 “Who” is a specific instance of “What” that involves recognizable people.

NOTE 2 Importance is based on the need of users to know information about the image and is related to the purpose of the image (or image component) and context in which the image is presented. Different information becomes important depending on the purpose of the image and the environment or context of the image.

NOTE 3 While identified content is complex, it might become useful to separate that image (or image component) into several (lower level) image components, in order to better be able to identify simpler content components.

d) Elaborate on the image (or image component) content

Where elaboration aids understanding of the image, elaborate on the description of the image (or image component) in accordance with clause 4.6 and clause 5.

This involves answering applicable questions relating to:

1) "What?"
2) "When?"
3) "Where?"
4) "How much?"

5) "How?"

NOTE 1 This step helps to identify information that might be important to the user.

NOTE 2 While initial identification of image (or image component) content might capture the most obvious information, elaboration focuses more on specific details that might be missed at first glance but that might also be important for the user to know.

e) Organize information into text alternatives

Organize the information obtained in steps a) to d) to improve its readability and allocate to the primary or secondary alternative texts of an image. This involves:

1) Removing redundancies within the identified information.

2) Allocating each piece of information to primary or secondary alternative text or other applicable location (e.g. caption or main document text) based on the importance of the information. Where the resulting description is too long or too detailed, moving less significant information from the primary and secondary alternative texts of an image.

3) Organizing the information in each location in a logical, readable order.

4) Ensuring compatibility between the descriptive information with the environment and context of the image (e.g. descriptions outside the image that are redundant or conflicting with this information such as captions and document content).

f) Evaluate text alternatives

Evaluate potential text alternatives (both primary and secondary alternative texts) by someone other than the person who created it to check that it suitably describes the image within the context of the document within which it is contained.

NOTE 1 While it is ideal for this to involve actual user testing, it is important that this step not be omitted due to lack of available users or other resources.

NOTE 2 Evaluation by a colleague or friend is better than no external evaluation at all.

4.2 Purposes

4.2.1 Introduction to purposes

A purpose identifies "why" an image is being presented as part of a document. It often identifies the function that the image serves within the document.

NOTE 1 While it is important to identify the purpose or purposes of an image, it is usually not productive to try and identify the purposes of individual components of an image.

NOTE 2 A "purpose" is similar to but different from a statement of the abstract content of an image. A purpose identifies why the image was presented, while the subjective content identifies the meaning of the object. [see clause 4.5.2]

Purposes can be classified in terms of:

— informative purposes

— control purposes

— decorative purposes

— formatting purposes

NOTE 3 While the purpose of the image can be considered in terms of the four classifications, it is more useful to describe the purpose instead. For example, “The image is used to illustrate this process” rather than “informative”.

NOTE 4 It is recognized that images (and image components) can have more than one reason for being presented.

a) The image was chosen to be in the document for a reason (purpose). This purpose should be alternatively achieved by the text alternatives.

b) Depending on the purpose and the context of use of an image, a statement of purpose may be needed as the first part of the primary alternative text for the image.

4.2.2 Informative purposes

Most images are intended to provide information to the user that duplicates, supplements, and/or provides an example of content that is also presented textually, usually within the same document.

Images used primarily for informative purposes generally contain information that is important for the user to receive.

a) Where an image is presented for informative purposes, the text alternatives shall include an identification of the content of the image. [see clause 4.5]

NOTE This often also involves elaboration of the content. [see clause 4.6]

Some images are accompanied with a caption that provides a brief statement of its purpose. In such cases, repeating the statement of purpose (and especially repeating the caption) in text alternatives is not helpful.

b) If a caption is associated with an image, and it provides a suitable statement of the purpose of the image, text alternatives should not restate the purpose.

NOTE Within WAI-ARIA it is possible to use the caption as part of the text alternative.

c) If a caption is not associated with an image, either a caption should be added to the document containing the image or the text alternatives for the image should start with a statement of purpose that would make a suitable caption.

d) If the main purpose of an image (or image component) is to present information that the user can interpret from an objective description of the image (and its components), then its purpose in the text alternatives may be stated at the level of a caption.

e) If the purpose of the image (or image component) includes presenting subjective/emotional/motivational information, then this aspect of the purpose should be described in text alternatives.

4.2.3 Control purposes

Images are often used as the basis for developing controls.

NOTE 1 Images can be used with/as controls such as buttons, sliders, knobs, icons, links.

EXAMPLE An image is used as the target for a hyperlink.

The ability to associate text alternatives with these images depends on how they are implemented. In cases where the images are separately addressable from the controls, it is both possible and useful to provide suitable text alternatives.

Images used with/as controls are significant images.

Where an image is used in conjunction with a control, text alternatives should provide both information about the image and its relationship to its associated control.

NOTE 2 This can be fully accomplished by following the guidance within the other parts of this document. [See clause 5.6]

4.2.4 Decorative purposes

Decorative images are often ignored by creators of text alternatives. However, this can result in failure to provide screen reader users with important information.

The use of images for decorative purposes might be intended to add visual appeal to a document. This visual appeal might be important in attracting and retaining the attention of users to a document. If text alternatives are not used for such images, screen reader users are deprived of getting the same emotional information that is provided visually. However, in this case the image is actually used to present emotional and subjective information [see clause 4.2.2].

The information present in some decorative images is of minimal importance. Providing text alternatives for such images might create unnecessary work for users of text alternatives.

Some situations of where text alternatives might not be needed or appropriate include:

a) an image that is used only to fill space that otherwise would be empty;

EXAMPLE 1 Background images composed of colors / textures are often used for the sole role of making Web pages appear attractive, without adding any particular meaning to the Web page.

b) an image that is excessively used where redundant complete text alternatives for each usage would provide a hindrance to the user.

EXAMPLE 2 A corporate logo is used instead of standard bullets to precede items in a list. The items in the list can also be recognized as items in the list from their formatting, and thus knowledge of the existence of the bullets (or images used to replace standard bullets) does not provide any additional information on formatting. Furthermore, repeatedly providing the same text alternatives for each of these bullets could become annoying to screen reader users.

NOTE It is best to analyze all images according to the procedure in clause 6.1 and to allow the importance level of descriptive information to determine whether or not text alternatives are needed.

4.2.5 Formatting purposes

Formatting organizes, separates, and/or highlights some information to distinguish it from other information.

Markup languages, such as HTML and SGML, provide explicit means for formatting text entries. If they are properly used, the addition of images, such as horizontal line separators, only provides additional redundant formatting.

NOTE 1 While developers often fail to provide text alternatives for images used for formatting, the use of such images instead of or in addition to standard formatting methods often is also done for decorating the document [see clause 6.2.4].

NOTE 2 Even if an image used for formatting has multiple image components, these components are not significant or useful in describing the purpose of the image for formatting.

NOTE 3 The existence of components used for formatting within an image might appear to have some importance. However, this can be described by describing the relationship between the components, without needing to describe the means of formatting used, unless they also fulfill some other purpose in the image.

a) It is not necessary to provide text alternatives to describe the use of images solely for formatting purposes, provided that the formatting information is otherwise provided to the user textually or through markup.

EXAMPLE 1 Fancy borders are used instead of spaces to visually highlight the separation of different chunks of content on a web page. While these borders do not provide any significant content, the distinction between the different chunks of content is primarily provided by the proper use of headings. However, these fancy borders have a second purpose in that they are also used for decorating the web page.

b) Text alternatives may be provided to describe images used solely for formatting, where such descriptions provide the user with important guidance regarding the use of the document where the image is used.

c) Where text alternatives are provided to describe formatting purposes, the purpose of the image should be described as simply as possible.

EXAMPLE 2 Text alternatives for a border between two sections of a form states, "Separation between personal information and product information".

d) If blank images are used to space content out, they should not have text alternatives. However, when the space separates words visually that would otherwise be near each other or otherwise conveys information by the space, this information should be conveyed to the user in some fashion.

4.2.6 Brevity of statements of purposes

Where statements of purpose are included in text alternatives, they should be brief.

NOTE 1 Further details about the purpose (within text alternatives) are provided by identifying the content and identifying qualifications and relationships to the content.

4.3 Context of an image

4.3.1 Text alternatives relate an image to its context within a document

The text alternatives should communicate whatever information that the content provider intended to communicate by choosing to use the image in the document.

NOTE 1 Text alternatives are context dependent and thus are different from information used for cataloguing of images within an image library.

NOTE 2 There is a difference between understanding of the document content and the understanding of the image. Understanding of the image might or might not have an effect on the understanding of the document. It depends on the purpose and context of the image. An image can have an effect on the understanding of the document content in two manners:

1) Objective information can influence intellectual or knowledge-based understanding. Objective information is factual and/or logical. For example, in a bar chart, the statistical data and axis information are objective information.

2) Subjective information can influence affective or emotional-based understanding. Subjective information consists of emotions, concepts, opinions, and judgments that are not necessarily universally shared. For example, different cultural interpretations of symbolisms in a painting are subjective information.

4.3.2 Context of images within panels within a document

Each document might be composed of multiple presentation panels. Each presentation panel has its own context.

EXAMPLE A web page (document) for a newspaper article uses presentation panels for: its navigation menu; the article itself, various advertisements; and a footer.

An image within a presentation panel should be described based primarily on the context of the presentation panel

NOTE 1 It is important to focus on the context of the presentation panel in which the image occurs, since the context of the overall document can change unpredictably based on changes to other presentation panels.

NOTE 2 Within this technical report, guidance relating to the document in which an image appears applies to the panel in which the image appears in situations where a document is composed of multiple panels.

4.4 Levels of importance

4.4.1 Importance is context dependent

The importance level of a piece of information relating to an image is context dependent.

NOTE It might change when the use, purpose, and/or context of the image changes.

NOTE 2 In situations where the image presents unique information, this information adds to what is presented in the main document text. Additional objective and subjective information presented in an image is important to the understanding of the document.

NOTE 3 In situations where the image complements (restates, modifies, elaborates, supplements) the main document text, the image might also contain information that is not provided in the main document text. Complimentary objective and subjective information might be important to the understanding of the document. A full understanding of the image can provide a better understanding of the document.

NOTE 4 In situations where the image adds visual appeal, the image creates or modifies the mood of the document, the information might be focused on the subjective rather than the objective understanding of the document. Images used for visual appeal might have little or no objective information relevant to understanding the document. Subjective information might be important to the understanding of the document.

NOTE 5 Text alternatives for an image are intended to inform users of information that the image is communicating. There might be times when the image presents information that conflicts with what is presented in the main document text. The conflicting information could be intentionally or unintentionally presented.

a) If the conflicting information is intentionally there, then the information should be pointed out and explained as part of the text alternatives or the main document text. Since the conflicting information was intentionally made available visually, it should also be made visible in text alternatives or the main document text.

b) If the conflicting information is unintentionally there and recognizing this, the decision is made to retain the image despite the conflicting information, then the conflicting information can be ignored in the text alternatives. This is not information that the content provider intended to provide and so should not be highlighted. It can also be ignored in hopes that sighted users will not notice the conflicting information.

4.4.2 Importance

4.4.2.1 Levels of importance

There are three levels of importance (essential, significant, helpful) and one level of being not important.

4.4.2.2 Essential information

Essential information is necessary to understand the image within the document in which it appears.

Essential information shall be presented either in the main document text (when referring to the image) or in the primary alternative text.

NOTE By placing essential information in the main document text (when appropriate) it ensures that all users (not just users with screen readers) will have access to this information.

NOTE 2 The purpose of an image is generally essential, especially for images that are used as user interface elements such as icons.

Essential information may have some or all of the following properties:

— It is aimed at the target audience.

— It must be known in order to comprehend the document.

— Most people want / need it most of the time.

— The user would be confused as to what the document is talking about without this information.

— Without it, the user has no idea why the image is there or what the image is for.

— It provides a good first impression of the image.

— Based on this information, the user will determine if they need/want to know more about it.

— For the content provider, this is the information that the content provider absolutely wants to tell people about.

— It provides the essence, purpose, function, or intent of the image.

— It identifies that the image conflicts with the main document text and that this conflict is intentional.

NOTE 3 As more of these properties apply, it is more likely that the information is essential.

4.4.2.3 Significant information

Significant information is useful for getting a comprehensive understanding of the image within the document in which it appears, when such information is desired by the user (based on the user's understanding of the essential information). Significant information satisfies the more detailed interests of most users most of the time.

Significant information shall be presented either in the primary alternative text or in secondary alternative text.

NOTE 1 The placement of significant information depends on the amount of essential information that is already contained in primary alternative text.

a) Where all the essential information has been placed in the main document text, then significant information can usually be placed in the primary alternative text.

b) Where a large amount of essential information has been placed in the primary alternative text, then significant information is better placed in the secondary alternative text to avoid overloading the primary alternative text.

Significant information may have some or all of the following properties:

— It is aimed at the target audience.

— It gives a more detailed and thorough understanding of the image and/or document.

— It is information that could be obtained by more than a quick glance.

— The user should know about it as they are reading the document in order to understand the document.

— The user decided to know more based on the essential information. This information goes into more details about the essential information.

— Without this information, the user has an idea of what the image is about and the reason the image is there, but does not have a detailed understanding about it.

— For the content provider, this is information that further explains and gives more details on what the content provider wants to tell the users.

NOTE 2 As more of these properties apply, it is more likely that the information is significant.

4.4.2.4 Helpful information

Helpful information provides a more thorough understanding of the image within the document in which it appear for those users who wish a more detailed description of the image.

Because helpful information is only of interest to some of the users some of the time, it should not be placed in the primary alternative text. It may be placed in secondary alternative text or in a separate document that is linked from either the main document text or the primary alternative text.

Helpful information may have some or all of the following properties:

— It is specific details that might be of interest to some who are the target audience of the document.

— It is targeted towards very specific audiences (other than the target audience) or a subset of the target audience.

— It provides the user with a better understanding of the image when the user is not an expert in the topic area or not the target audience of the document.

— It might reassure the user that they have not missed something of greater importance.

— Without this information, the users have a fairly complete understanding of what the document is about but have some things that the users still want to know.

— It includes different or other possible interpretations of the information being expressed by the image.

— For the content provider, this is information that could clarify some things for some people.

— It includes optional extra information that is seldom wanted or needed, but elaborates on what is already there.

NOTE As more of these properties apply, it is more likely that the information is helpful.

4.4.2.5 Not important information

Information is not important if it does not help to provide much additional understanding of the image for any users in order to understand the image or document in which it occurs. This can include information that is not appropriate to consider given the context of the image within the document.

Information that is not important should not be presented to users either in the main document text or in text alternatives.

Information that is not important may have some or all of the following properties:

— Very few to no users will want to know or care to know this information.

— It is rarely helpful.

— It is not important enough to mention.

— Without this information, the user knows everything they want or need to know in order to understand the document and/or image.

— This is information that might result in unintended confusion or boredom and does not help users understand what the content provider is saying.

NOTE As more of these properties apply, it is more likely that the information is not important.

4.5 Images and image components

An image can be considered first as a whole entity. Some information might be present only when the image is viewed as a whole.

Many images can also be broken down into image components where a number of image components might present important information to the user.

Whether or not a component is identified and elaborated upon depends on how important information about the component is with respect to the purpose and context of the image.

Images can consist of one or more image components. Images can focus on a single component, such as a single shape or icon. Images can contain multiple components that are important for the user to know and understand in order to comprehend the entire image. In an image with multiple components, the components can be considered individually or as a set of components.

Each person, object, shape, text, landmark, or step in a process can be considered an image component.

EXAMPLE 1 In the image in Figure 2, possible components include whole image, the plate of desserts and the cup of coffee. Depending on the purpose and context of the image, it might or might not be necessary to break the image down into smaller components. The plate of desserts could be broken down such that each dessert is its own component.

Figure 2 —

EXAMPLE 2 In the image in Figure 3, possible components include the whole image, the furnace, the kettle, and the brick. Depending on the purpose and context, it might be important to identify additional components. If the purpose is to illustrate the ways homes were in the 1800s, then it may be important to identify the bench in the background, the rocking chair, the walls, and the flooring as components.

Figure 3 —

Although image components can be described as part of the whole image (without breaking down into image components), less information about the component is identified as a result. When image components are identified, there will be additional focus and attention on providing detailed information about that image component. It is important to have as much information as possible about an image in order to create appropriate text alternatives.

Images should be broken down into components whenever this helps to identify important information.

4.6 Identification of content

4.6.1 Introduction to identification of content

The question "What" can be asked about both subjective content and objective content.

NOTE The question "Who" can be substituted for "What" when the answer involves people.

The identification of this content is intended to provide the main information elements out of which text alternatives can be constructed. While it can include some amount of elaboration, it is intended to focus on these basic elements. A more extensive focus on the elaboration of these elements is discussed in clause 4.6 and in clause 5.

4.6.2 Subjective content

Subjective content answers the question "What is the meaning of this image (or image component)?"

NOTE 1 Clauses 5.1.2 and 5.2.1 provide guidance on potential elaborations of subjective content, for use where such content and/or elaborations are important.

Subjective content can include concepts, theories, symbolic meanings, intended emotions, opinions, judgments, and other explanations that go beyond identifying individual components and obvious relationships.

NOTE 2 A description of the subjective content of an image is similar to but different from a statement of purpose of an image. The subjective content identifies the meaning of the object while a purpose identifies why the image was presented. [see clause 4.2]

a) Where the meaning of an image (or image component) is essential, this meaning shall be included within the primary alternative text of an image.

b) Where the meaning of an image (or image component) is significant, this meaning should be included within the primary alternative text of an image.

c) Where the meaning of an image (or image component) is helpful, this meaning may be included within the secondary alternative text of an image.

4.6.3 Objective content

4.6.3.1 Identifying objective content

Objective content answers the question "What objects / actions can be identified from the image (or image component)?"

NOTE 1 Actions generally do not occur on their own, they generally involve a subject and/or an object. Thus actions are typically combined with objects in identifying objective content.

EXAMPLE The combination of subject, action, and object can be expressed in various formats, such as those used in: "A boy playing", "Playing baseball" and "A boy playing baseball".

NOTE 2 Clause 6 provides a variety of guidance on potential elaborations of objective content, for use where such content and/or elaborations are significant.

Identification can occur over a range from highly specific to highly general. The appropriate level of specificity

NOTE 3 Sub-clauses 4.5.3.2 and 4.5.3.3 provide further discussion about specific and generic objective content. However, they are not definitive. This technical report intentionally avoids making any firm distinction between these two concepts.

Identification of objective content is best done briefly, leaving elaboration to be dealt with later (see clause 4.6). While it is usually preferable to identify images (or image components) as specifically as possible, it is important not to allow this specificity to impair the understandability of the resulting text alternatives. The object can be described in general or specific terms. [See clauses 4.5.3.2 and 4.5.3.3]

a) When it is essential to identify an image (or image component) in terms of objective content, this identification shall be included within the primary alternative text of an image.

b) When it is significant to identify an image (or image component) in terms of objective content, this identification should be included within the primary or secondary alternative text of an image.

c) When it is helpful to identify an image (or image component) in terms of objective content, this identification may be included within the secondary alternative text of an image.

4.6.3.2 Generic content

Generic content identifies an object and/or action in a general manner that involves only general everyday knowledge available to a user. It answers the question "What is this?", and/or "What is happening?"

Generic content does not deal with specific branding or naming of an object.

NOTE 1 In English, these general identifications of objects would typically use the indefinite article "a".

EXAMPLE 1 (objects) a man, a baseball player, a car, a building, a table, a tower

NOTE 2 In English, generic actions tend to specify "What is happening" without identifying "How" the action is performed.

EXAMPLE 2 (actions) skipping, cutting, chopping.

NOTE 3 Generic content can include elaborations that identify properties or attributes of the object or action. Further elaborations of content are discussed in clause6.6 and in clause 7.

EXAMPLE 3 a green car

4.6.3.3 Specific content

Specific content answers the questions "Who is this?", "What is this?", and/or "What is happening?" It uniquely identifies a particular instance of an object and/or action that can be named in a manner that distinguishes them from other instances of similar objects.

NOTE 1 In addition to graphical objects appearing in an image (or image component), textual objects sometimes also appear. Where the text is readable, it can be considered to be specific content that is to be identified in the primary alternative text of an image.

EXAMPLE 1 A picture of a highway includes a road sign that indicates that Saskatoon is 30 km further down the highway.

Specific content does not interpret the meaning of these objects beyond identifying them. Meaning belongs within subjective content. If an interpretation of the subjective meaning of an image (or image component) is significant, the image (or image component) can be identified at both the specific and subjective levels.

NOTE 2 In English specific names of objects would either be used without an article, or would typically use the definite article "the".

EXAMPLE 2 (objects) Alan Touring, Babe Ruth, the Batmobile

NOTE 3 In some cases, a specific name is created by adding specific qualifications to what would otherwise be a generic identification. These qualifications are usually made because of their being readily identifiable. Further elaborations of content are discussed in 6.6.

EXAMPLE 3 (objects) King Arthur's round table, the Eiffel tower

NOTE 4 In English specific actions tend to specify "How" the action is performed in addition to identifying "What is happening".

EXAMPLE 4 (actions) jumping over hurdles at 8.5m/s.

4.6.4 Relationships of images and their components

4.6.4.1 Importance of relationships

In addition to identifying images (and their components), it is important to identify relationships between images (and their components), both with each other and with the document in which they are displayed. For images with multiple image components, the relationships between those components might be essential to the understanding of the image.

Relationships can provide important information about the content of an image (or image component). However, they are more complex and more diverse than merely answering "What relationships occur?"

EXAMPLE Images of the sides or backs of a computer may be intended to show the possible inputs and outputs to the computer. Knowing the relative locations between a headphones output port and a microphone input port may be important to help users properly insert the correct cables.

Relationships can be categorized into

a) logical relationships,

b) temporal relationships,

c) physical (spatial) relationships.

4.6.4.2 Logical relationships

Logical relationships explain what elements are interacting and how some element (e.g. an image or image component) interacts with some other element (e.g. the Web page where an image is displayed, the image that contains the image component, or some other image component).

NOTE The "purpose" of an image provides a semantic relationship of the image to the document where the image is used.

4.6.4.3 Temporal relationships

Temporal relationships explain when some action occurs or the time in which an entity occurs. This can include when apparent (or intended) changes to an image (or image component) occurs. It can include providing information on the sequencing of events.

4.6.4.4 Physical (spatial) relationships

Physical (spatial) relationships explain where an image (or image component) is located in relation to other physically occurring elements. The positioning of an image provides a physical (spatial) relationship of the image to the document where the image is used.

5 Expanding on the identification and elaboration of content

5.1 Introduction

This clause presents detailed questions that can help in the preparation of text alternatives in certain circumstances.

There are a number of ways that the basic content (identified according to clause 4.6) can be elaborated. The appropriate types of elaboration depend both on the image (or image component) and on the document within which it is contained.

This clause is organized based on the general questions identified in clause 4 that these detailed questions are related to. This involves answering applicable questions relating to:

a) "What?" (going beyond basic identification)

b) "When?" (including temporal relationships)

c) "Where?" (including spatial relationships)

d) "How much?" (including quantitative stuff)

e) "How?"

It is recognized that many of the detailed questions may overlap with one another. It is better to answer as many questions as apply and then organize the information from the answers once all the answers have been obtained and evaluated for their importance.

These elaborations can be included within the primary or secondary alternative texts of an image depending on their importance.

NOTE While individual pieces of elaboration are often identified individually, they are best presented in a manner that reads smoothly in combination with the content that they elaborate.

5.2 Detailed questions relating to identification ("what")

5.2.1 Classifying the image (or image component)

There are many different ways an image (or image component) can be categorized. This clause identifies some possible ways of classifying an image (or image component). The ways discussed in this clause are neither exhaustive nor definitive.

It is possible to distinguish between:

a) a realistic image (or image component)

b) an abstract image (or image component)

c) a non-realistic image (or image component)

For works of art, it is often important to distinguish the school to which the image belongs (e.g. impressionist, surrealist, etc.).

It is sometimes possible to distinguish the source of the image. For images (or image components) that are known to be a photograph of another image with its own independent existence, it may be important to include this information in the text alternatives.

EXAMPLE 1 A photograph (which can be of various types of images, including paintings, objects, scenes, etc.) which has its own independent existence,

EXAMPLE 2 An electronic drawing (which can be of various types of images, including pictures, charts, graphs, diagrams) rendered for computer display that is created as an original work (and that is not a photograph)

Images (or image components) should be classified in whatever categories have importance within the document within which it is contained.

NOTE It is not necessary to classify images using the terms described here (e.g. realistic, abstract, non-realistic). However, it might be important to classify the image in some appropriate manner, if this classification can quickly let the user know what types of information to expect e.g. if an image is classified as a chart, the user can expect quantitative information.

5.2.2 Elaborating on textual content

It is not appropriate to use an image to present only text that could be otherwise presented as a text element within the document within which it is contained. However, there are many instances of images that contain textual components.

Each separate instance of text within an image can be considered either as a separate component or as an elaboration of a separate component of the image.

a) All text with some importance in an image should be made available as text in some manner to the user. The manner of making text available will depend on its importance level. This can include:

    1) Placing large transcriptions (of over 30 continuous characters) within the main document text within which the image is contained or if they are of lesser importance, within the secondary alternative text of the image.

    2) Placing short transcriptions within the primary alternative text of an image or if they are of lesser importance, within the secondary alternative text of an image.

b) Text that is blurred or partially obscured and that is not at least helpful may be ignored.

5.2.3 Elaborating on physical objects

Each object in the image can be considered as an image component. The following questions should be considered in creating text alternatives for images (or image components) involving physical objects:

a) What does the object represent?

b) What is the brand / model / part name (number) of the object?

5.2.4 Elaborating on people

Each person in the image can be considered as an image component. The following questions should be considered in creating text alternatives for images (or image components) involving people:

a) Who is the image or image component of?

b) What does the person look like? (i.e. age, sex, nationality, hair color, eye color, hair style, etc.)

c) What is the facial expression of the person?

d) What is the person doing?

e) What position is the person in? (ex. Standing with hands across the chest)

f) What other information about the person is important for users to know?

5.2.5 Elaborating on perceptual objects and perceptual properties of other objects

Perceptual content answers the question "What can I know about this?" It describes the physical appearance of the object.

NOTE Perceptual content describes low-level perceptual features of an image (or image component), in a manner similar to that possible by basic vision detection systems.

Typical perceptual features include: color, texture, shape, and pattern.

EXAMPLE Some perceptual features can be expressed as: blue, yellow, blue-green, aqua, smooth, rough, round, square

Perceptual content does not assign any meanings to any of these features. Meaning belongs within subjective content.

The following questions should be considered in creating text alternatives for the perceptual properties of images (or image components):

a) What (are) the color(s) of the image (or image component)?

b) What is the shape the image (or image component)?

c) What is / are the size / dimensions of the image (or image component)?

d) What is the texture of the image (or image component)?

e) How is the image (or image component) positioned? (e.g., sideways, angled, and facing left)

f) What other perceptual information is important for users to know?

5.2.6 Elaborating subjective descriptions

The following questions should be considered in creating text alternatives for images (or image components) with subjective information:

a) What concepts are associated with the image or the image component?

b) What is the image or image component representing or symbolizing?

c) If the colour(s) of the image or image component is symbolic, what is the colour(s) representing?

d) What themes are represented?

e) What emotions are being expressed?

f) How is the user expected to respond emotionally (with feelings, judgments, and opinions) to the image?

g) What other subjective information is important for users to know?

5.2.7 Elaborating on logical relationships / interactions and actions

Elaboration on logical relationships / interactions and actions involve both “What” and “How”. (Because “What” comes first, they are discussed under it.)

Logical relationships describe how entities interact with each other to achieve some purpose or goal. They include the various types of relationships that are the foundation of typical entity-relationship diagrams. In Software Engineering terms, these are often referred to as associations.

Actions are seldom depicted without some entity performing them. In this case the activity modifies the subject. Actions can usually be expressed in terms of a verb or verb phrase (e.g. walks carefully, jumps high, falls down). Where an object of the action is involved, an action can be considered to be a logical relationship.

Logical relationships can usually be expressed in terms of a verb (e.g. purchases, builds) or verb phrase that unites a subject (entity) with an object (entity).

EXAMPLE 1 A customer purchases a book.

EXAMPLE 2 The monkey is climbing the banana tree.

EXAMPLE 3 Part A (the seat) is connected to the bicycle at location F.

While it is usual that both the subject and the object will be a part of the image being described, it is possible that either the subject or the object will only be implied. Where the image makes such an implication, it is important that the text alternatives make the user aware of this being implied.

EXAMPLE 4 The sailor is scanning the horizon. It is implied that he is looking for land.

The following questions should be considered in creating text alternatives for images (or image components) involving interactions:

a) What interaction or action is taking place?

b) What or who is the subject of the interaction or action?

c) What or who is the object of the interaction?

d) How is the interaction or action being performed?

e) What is the intended result of the interaction or action?

f) What other information about the interaction or action is important for users to know?

5.2.8 Elaborating on locations/places

While locations are also dealt with under "Where", there are situations in which a location is a significant component of an image. In these situations, a location also becomes a "What".

NOTE It is better to identify redundant information than to miss identifying important information. Redundancies can be eliminated when the actual text alternatives are written.

The following questions should be considered in creating text alternatives for images (image components) involving locations or places:

a) What is the setting, place, or location depicted in the image (or image component)?

b) What specific landmarks are visible in the image (or image component)?

c) What other information about the location is important for users to know?

5.3 Detailed questions relating to temporal elaboration ("when")

5.3.1 Introduction to temporal properties

There are two types of temporal properties to consider

a) Temporal properties that are part of the information that is presented by an image (or image component), including:

   1) Time periods

   2) Events

   3) Sequential relationships

b) Temporal properties relating to changes in an image (or image component) including:

   1) Image components which change in order to represent state information

   2) images within slide shows

NOTE 1 While moving images are outside the scope of this document, it is recognized that some information relating to change can apply to the images (or image components) that are within the scope of this document.

NOTE 2 While the focus of various temporal properties is "when" something occurs, the wording in the following sub clauses does not always use the word "when".

5.3.2 Elaborating about time periods

An image (or image component) might be readily identifiable with some particular time period(s), such as:

— Time of day (e.g. morning, midday, afternoon, evening, night)

— Special days (e.g. holiday celebrations)

— Time of year (e.g. spring, summer, fall, winter, January)

— Historic period (e.g. the 1960's, the Victorian era, the stone age)

If time period information is readily identifiable regarding an image (or image component), then it should be included within the primary or secondary alternative text of an image (or image component) according to its importance level.

5.3.3 Elaborating about events / activities

Events in people's lives are often synonymous to activities and might be related to actions (see 4.6.3 and 5.2.7). An image (or image component) might be readily identifiable with some particular event/activity (e.g. a birthday party, a wedding, a picnic, a baseball game, baking a pizza).

In more scientific terms, events are usually considered to occur instantaneously when something changes. An image (or image component) might be readily identifiable with some particular instantaneous event (e.g. the start or end of a race, the loss of power to some device, an accident).

If information about an event is readily identifiable regarding an image (or image component), then it should be included within the primary or secondary alternative text of an image (or image component) according to its importance level.

5.3.4 Elaborating sequential relationships

Some images illustrate processes or other sets of sequential relationships (e.g. assembly lines, assembly instructions, flow diagrams, structure charts). Each step or item in the sequence can be considered as an image component.

Sequential relationships can be:

— linear (where one element leads to a single other element)

— branching / hierarchical (where one element leads to multiple other elements)

— cyclical / networked (any element in a set of elements eventually leads to all elements in the set, including itself)

— sets of one or more of the above

Individual sequential relationships can be:

— one directional

— bi directional

The following questions should be considered in creating text alternatives for images (or image components) involving sequential relationships:

a) What is the basis (or purpose) of the sequential relationship?

b) What types of sequential relationships are involved (linear, branching, cyclical, one directional, bi directional)?

c) If there is a start and/or end point(s) to the set of relationships, what are the start and/or end point(s)?

d) What are the individual steps (components) of the relationship?

e) What is a suitable basis (or method) for logically ordering the individual steps (components)?

f) How is each step (component) related temporally to the other steps (components)? This can include:

   1) connections to previous and following steps (components)

   2) time involved in individual steps (components) or for the transition between steps (components)

   3) logic (decision or event) involved in moving from one step (component) to another

g) What other information about the sequential relationship is important for users to know?

5.3.5 Elaborating on states

Individual image components might change within an image to provide various types of state related information.

EXAMPLE 1 Connectivity and power icons are used to demonstrate the signal strength and amount of charge of a system.

EXAMPLE 2 Images used for links are provided a distinctive border to indicate that the link has already been visited.

EXAMPLE 3 Images are displayed lighter or grayed in order to indicate that the control associated with the image is not available.

The following questions should be considered in creating text alternatives for images (or image components) involving changes in individual image components:

a) What different states (values) can occur for the image component?

b) How are different states (values) visually represented?

5.3.6 Elaborating on slide shows

In addition to providing appropriate text alternatives for individual images within a slide show, the following questions should be considered in creating text alternatives for images within a slide show:

a) Questions relating to the entire slide show

   1) How many separate images does the slide show consist of?

   2) How can the user control (e.g. interrupt, go back) the playing of the slide show?

b) Questions relating to an individual image

   1) What is the position of each image within the slide show?

   2) What time interval will the image be presented for?

   3) What action or motion is being portrayed in transitioning to the next image?

Each individual image within a slide show can be considered as an image component.

5.4 Detailed questions relating to physical (spatial) relationship elaboration ("where")

The question "where" can encompass both the physical (spatial) location of an image (or image component) and the context in which it occurs.

Spatial relationships can range from general statements to very precise ones. They can relate the image to other items of document content or relate an image component with respect to other image components.

The following questions should be considered in creating text alternatives for images (or image components) with spatial relationship:

a) Where is the setting of the image (or image component)?

NOTE 1 This can describe information on the setting of the image (or image component).

EXAMPLE 1 The image is at Siesta Beach.

EXAMPLE 2 The chair is sitting on the beach, near the water

b) Where is the image (or image component) spatially located within the document (or image)?

NOTE 2 This can be answered in terms of the coordinates of the image (or image component). There are various ways in which coordinates can be expressed based on:

1) The coordinates of the top left corner of the image component

2) The coordinates of the center of the image component

3) The coordinates of the top, bottom, left, and right of the image component

c) Where (in general) is the image (or image component) located with regards to the object (e.g. document, image, image component) that contains it? (e.g., upper left, lower right.)

d) Where is the image (or image component) relative to other entities (surrounding content, or internal components)?

NOTE 3 Relative positioning can be express various spatial relationships including: above, below, to the right of, to the left of, in front of, behind, touches, crosses, overlaps, contains, within)

5.5 Detailed questions relating to quantitative elaboration ("how much")

The following questions should be considered in creating text alternatives for images (or image components) with quantitative information (such as charts and graphs):

a) What is the quantity?

b) What is the quantity associated with? / What does the quantity represent?

c) What is the unit of the quantity?

d) Is the quantity fixed or dynamic?

e) What is the precision (or statistical significance) of the quantity?

f) What is the trend that can be interpreted from the data?

g) What do the axis or axes represent?

NOTE To present the specific quantitative data, the data can be presented in a table while a summary or analysis of the data can be presented in the form of text alternatives.

5.6 Detailed questions relating to procedural elaboration ("how")

5.6.1 Elaborating on the intended use of the image

The following questions should be considered in creating text alternatives for the control properties of images (or image components):

a) How is the user supposed to interact with the image (or image component)?

b) What action is the user supposed to perform in order to interact with the image (or image component)?

c) What is intended to result from interacting with or performing an action on the image (or image component)?

d) What can go wrong in the interaction with the image (or image component)?

5.6.2 Elaborating on control

Where users are expected to interact with an image or a control that the image is used with, it is important for the user to be made aware of this expectation.

6 Guidance on writing text alternatives

6.1 Importance and purpose

Some technologies present the primary alternative text automatically by default, while the user needs to request the secondary alternative text. Some technologies have a limitation or restriction on the length of the primary alternative text. It can be as short as 125 characters.

The most important information (not already present in the main document) should be presented in the primary alternative text in order to let the user know what the image represents.

6.2 Elaborating on the context of an image

An understanding of the context of an image can help determine the amount and content of text alternatives that is provided to help a user to understand the image.

Where a document contains an explanation (or caption) of an image, this information should not be repeated within text alternatives.

Elaborations of the information may be made to the existing explanation (or caption) or included within the primary or secondary alternative texts of an image.

6.3 Flow with the document content

Depending on the placement of the image, sometimes, the text alternatives being read by screen readers interrupt the document content and disrupt the flow of the content. This often occurs when an image is placed directly above or below the content that references it. This can result in confusion if the text alternatives discuss a different topic than the surrounding content. For example, the screen reader might begin reading the Web page content about Napoleon’s adventures, then reads the text alternatives for the image of Napoleon, then continues reading the Web page content about Napoleon’s adventures.

It is important that the text alternatives be written in a way that when read it flows with the rest of the document content. Depending on how and where the image is positioned in the document, the text alternatives might be read at different times.

The text alternatives should be tested to ensure the flow of the document content.

6.4 Story telling

One method of writing text alternatives is as if it is a story. Describing an image as if it is a story can captivate the user while providing important information. Audio guides at some art museums describe paintings in this manner. The same method can be used to describe images.

6.5 Independence of importance from order

Importance is used for selecting the information to present, but not necessarily the order of presentation

NOTE Discussions about wanting specific information first or wanting it later is not appropriate here because whether or not the user gets it first does not affect the importance of the information. The information that users may want first might be a summary, which is could be part of text alternatives or might be placed in the main document text. The information within the summary may be either essential and/or significant. Instead this issue should be discussed in the guidance on organizing the information.

7 Guidance on evaluating text alternatives

Text alternatives should be evaluated to ensure that it describes the image suitably given the context of the web page. It should be evaluated by both sighted and visually impaired users.

Three different evaluations can be done. It is preferable that more than one evaluation be done.

a) It is preferable that evaluation be done by visually impaired users. Since the text alternatives will more frequently be used by visually impaired users, they are the most appropriate people to perform the evaluation. They can determine whether or not the text alternatives flow with the surrounding content and if the text alternatives make sense. However, visually impaired users are not always available to evaluate each image's text alternatives.

b) While some images may be hidden from screen readers (and hence the screen reader users), a sighted person can still see all image placeholders even when an image is not loaded. A second evaluation method would be for a sighted user to view the document with the images being not visible.

   EXAMPLE A web page, using a text-only web browser or a web browser with images not presented to view the web page can help a sighted individual experience a web page where text alternatives are needed.

c) A third evaluation can be done by people who can see the image. They can help determine if the text alternatives truly represent the content of the image and communicate the important information (or identify important information that is missing).

Although it may be difficult to perform all three of the evaluations, performing at least one of the evaluations is better than not evaluating at all. Evaluation of the text alternatives is a crucial part of the approach. Based on the results of the evaluation, the approach can be repeated to identify additional information or to rewrite the text alternatives.

Bibliography

[1] Defense Visual Information Directorate, Captioning Basics, http://www.defenseimagery.mil/learning/captionstyle/basics.html

[2] Eakins, J. P. 2001. Retrieval of still images by content. In Lectures on information Retrieval, M. Agosti, F. Crestani, and G. Pasi, Eds. Springer Lecture Notes In Computer Science Series, vol. 1980. Springer-Verlag New York, New York, NY, 111-138.

[3] Eriksson, Y. 1999. How to make tactile pictures understandable to the blind reader. http://www.dinf.ne.jp/doc/english/Us_Eu/conf/z19/z19001/z1900116.html

[4] Heller, R.S.; Martin, C.D., "A media taxonomy," Multimedia, IEEE , vol.2, no.4, pp.36-45, Winter 1995.

[5] Hunter, B., Crismore, A., and Pearson, P.D. “Visual Displays in Basal Readers and Social Studies Textbooks,” in *The Psychology of Illustration, Instructional Issues*, Vol. 2, D.M. Willows and H.A. Houghton, eds., Springer, New York, 1987, pp. 116-135.

[6] ISO 14915-3:2003 Software ergonomics for multimedia user interfaces — Part 3: Media selection and combination

[7] ISO/IEC13249-5:2003 Information technology — Database languages — SQL Multimedia and Application Packages — Part 5: Still Image

[8] JISC Digital Media, Approaches to Describing Images, http://www.jiscdigitalmedia.ac.uk/stillimages/advice/approaches-to-describing-images/

[9] Mikovec, Z. and Slavik, P., 1999. "Perception of pictures without graphical interface," in *Proceedings of the 5th ERCIM Workshop on User Interfaces for All (UI4ALL '99),* Dagstuhl, Germany, November-December 1999.

[10] Shatford, S. “Analyzing the Subject of a Picture: A Theoretical Approach” Cataloging & Classification Quarterly, v. 6(3), Spring 1986.

[11] Shen, H. T., Ooi, B. C., and Tan, K. 2000. Giving meanings to WWW images. In Proceedings of the Eighth ACM international Conference on Multimedia (Marina del Rey, California, United States). MULTIMEDIA '00. ACM, New York, NY, 39-47. DOI= http://doi.acm.org/10.1145/354384.376098

[12] Tang, L. and Carter, J., 2011, "Verbalizing Images", in *HCII International 2001 – Posters Extended Abstracts*, C. Stephanidis (ed.), Communications in Computer and Information Science, Vol. 173, 394-398, Springer, Heidelberg.

[13] Tripod Project, The Automatic Captioning of Photographs, http://grupoweb.upf.es/tfws08/slides/tripod-1x2.pdf

[14] WebAIM. Appropriate Use of Alternative Text, http://www.webaim.org/techniques/alttext/