Abstract: The “Content-Identifying Methodology”, or CIM, is an approach developed by environmental psychologists Rachel and Stephen Kaplan to understand the landscape characteristics that people find visually attractive. The Kaplans did this by surveying people’s landscape preferences and then analyzing the preferences to develop sets of landscape scenes to which people reacted in a similar pattern. The underlying assumption is that a common stimulus or content exists in the photographs of a set responsible for the preference. However, identifying the common stimulus or content in each set or grouping of scenes and how it affects preference can still be challenging. Eye-tracking is a tool that can identify what the survey participants were looking at when indicating their preference for a landscape. This paper demonstrates how eye-tracking was used in two different landscape preference studies to identify the content important to people’s preferences and provide insights into how the content affected preference. Eye-tracking can help identify a common stimulus, help determine if the stimulus is a physical or spatial characteristic of the landscape, and show how the stimulus varies in different landscape contexts.

Keywords: eye-tracking; preference; visual; landscape; streetscape; battlefield

1. Introduction

Is scenic beauty in the eye of the beholder? Many involved in the visual or scenic assessment of landscapes get irritated whenever we hear that old idiom because it implies no agreement among people about what is scenic or beautiful. But maybe that is because we do not know how to find out what is in the eye of the beholder. The objective of this paper is to demonstrate a tool for carrying out that, “eye-tracking”.

In order to answer the question “Is beauty in the eye of the beholder?”, we must first review a process developed by Rachel and Stephen Kaplan, environmental psychologists at the University of Michigan. The Kaplans use people’s preferences for landscapes to identify the physical features and spatial attributes of the landscape that are important to people’s experience of the landscape. This methodology is the Category or Content Identification Methodology [1] or CIM (Stephen Kaplan originally referred to this method as a Category-Identifying Methodology, but many now refer to it as a Content-Identifying Methodology). Next, the research of two graduate students at Virginia Tech who used CIM and eye-tracking will demonstrate how eye-tracking can be a helpful tool when undertaking a CIM analysis. Finally, this paper concludes with a summary of how eye-tracking can enhance the process of identifying physical and spatial characteristics that are important to people’s visual experience of the landscape.

2. Literature Review

2.1. The Public Role in Scenic Assessment

In the past, there has been tension between those involved in scenic assessment of public lands, often using some objective rating framework, and those who want scenic assessment to represent the public’s reaction to the landscape, most often through some survey. Lothian [2] refers to these two paradigms as the objectivist and subjectivist paradigms.
Lothian advocates for using the subjectivist paradigm, saying, “It is more scientific and statistically rigorous” [2] (p. 93). A subjectivist approach based on people’s preferences can be scientific and rigorous. The public does not have the technical expertise to assess the scenic value of the landscape, but they have no difficulty expressing how much they like different landscapes. A survey can collect people’s preferences for different landscapes. These surveyed preferences can then be analyzed to identify the landscape’s physical and spatial characteristics contributing to the preference for each landscape.

While each individual is unique, there is also tremendous agreement among individuals regarding landscape preferences. One reason for this is evolution [3]. Evolutionary survival requires the ability to read and understand a landscape to determine where food, water, and shelter can be found. Survival depends on it. So, part of the reaction to the landscape is innate. It is instinctively within us.

In addition, humans often share a tremendous amount of learned knowledge about the environment. Human landscape preferences are like many other human attributes, such as appearance. We have little difficulty telling people apart because they are each different. However, if someone walked into a room with three eyes or an ear on their forehead, we would all gasp because they are different. So, people are also the same, but within bounds.

However, knowing how much someone likes a landscape does not tell us why they like it. Landscapes are complex, with many combinations of physical and spatial attributes that can influence people’s reactions to them. How can the results of a preference survey be turned into something helpful for managing the physical landscape?

2.2. Content-Identifying Methodology

Rachel and Stephen Kaplan, environmental psychologists at the University of Michigan, developed a method of analysis (CIM) [1] that uses the results of a preference survey to identify categories of landscape content based on the physical and spatial characteristics, to which people are reacting. Thus, people’s preferences can provide knowledge on how to manage landscapes consistent with public preferences. This analysis method is well documented in Kaplan’s book, *The Experience of Nature* [4].

The CIM procedure categorizes landscapes or groups of landscapes according to a common pattern of preferences across a group of survey respondents. Photographs are used as a surrogate for landscapes in the preference surveys. Survey participants are asked to provide their preference for each scene on a 1 to 5 Likert scale, with 1 being not preferred at all and 5 being preferred very much [5]. The CIM procedure then uses a factor analysis to group or categorize scenes with a similar pattern of preferences. The groupings of scenes are called dimensions.

Note that the scenes are not grouped based on the similarities in the magnitude of the response. Instead, each category or dimension is based on a similar response pattern across the group of survey participants. The underlying assumption is that if people are reacting consistently to a group of scenes, then there must be a common stimulus in each group of scenes.

Each cluster or grouping of scenes must be visually examined to determine the common stimulus. This involves judgment. Often, there are very obvious common physical or spatial elements in each group of scenes. The common stimuli may not be as obvious in other dimensions or groupings of scenes. The purpose of this paper is to demonstrate how eye-tracking can be used to help understand the common physical or spatial content of the scenes in a dimension. Two additional aspects of this procedure worth noting are dimension preferences and how scenes are sampled for use in the survey. An overall dimension average can be determined by averaging the preferences of all survey participants for the scenes in each dimension. This provides a sound overall sense of how the content or spatial organization influences people’s preferences. In addition, eye-tracking can help determine why scenes with a common stimulus or content may vary in preference or what other landscape characteristics influence preferences.
Since the procedure involves collecting a sample of landscape scenes for a survey, some assume it should be a random sample. However, this type of sample will not work with a CIM analysis. Different types of landscapes vary in their frequency of occurrence. It is possible that a potentially highly preferred landscape is relatively rare and may not show up in a random sample. The survey sample should be selected to represent the range of different landscape types in the study area, with at least three of each possible type [4]. Determining the landscape types is carried out by systematically photographing the landscape and taking many more photos than are needed for the survey. Copies of the photos are then laid out on a table, and a final sample is selected by eliminating extras and making sure there are at least three of each landscape type.

2.3. Eye-Tracking

The use of eye-tracking in research has grown tremendously over the past 30 years, both in terms of types of applications and in computer technology. There is a plethora of literature on eye-tracking. Horsley’s book [6] documents 50 applications of eye-tracking research. The breadth of applications speaks to the perceived power of eye-tracking. It can be used for everything from tracking the movement of the eyes of clinicians viewing electrocardiograms [7] to eye-tracking to provide information for marketing purposes. Most eye-tracking studies have not involved landscapes or scenic assessment. There are a few studies, however, that involve eye-tracking while looking at the landscape. The eye-tracking results from these studies have been engaging in terms of the landscape content people focus on. However, it has been difficult for researchers to determine how the eye-tracking content relates to peoples’ preference for different landscapes, which is ultimately a landscape management goal. A study in 2022 examined the potential of eye-tracking software for analyzing landscape preferences [8] and found, “Thus, our findings indicate that the analysis of eye-tracking hotspots can support the identification of important elements and areas of a landscape, but it is limited in explaining preferences across different landscape types. Future research should therefore focus on specific landscape characteristics such as complexity, structure or visual appearance of specific elements to increase the depth of information obtained from eye-tracking simulation software” [8] (p. 4). This is precisely what the author of this paper is carrying out. This study demonstrates how eye-tracking can identify landscape characteristics or content related to the landscape based on a CIM preference-based analysis.

3. Method

3.1. Study Objectives—Eye-Tracking as a Tool

CIM has been around for quite a while, but there are still many people who do not understand CIM. One objective of this paper is to demonstrate what CIM does and how it can be better understood with the use of eye-tracking. Eye-tracking is identifying what a survey participant is looking at when they rate their preference for a scene. This can be recorded as a hot-point on the scene being viewed. Multiple viewpoints from different survey participants can also be recorded, creating a heat map, and can reveal something about the physical content and spatial organization of the scene that is being viewed. Eye-tracking can be used to examine reactions to multiple scenes in a dimension, thus helping to understand the common stimulus of CIM but also how the content can vary within a dimension and how it influences people’s landscape preferences.

3.2. Eye-Tracking Examples

Two Virginia Tech graduate students used eye-tracking to assist in interpreting CIM dimensions as part of their research. Both students were strong and thorough researchers, committed to understanding how people react to different environments. One student, Seth Estep, was a master’s student (MLA) [9] interested in urban pedestrian environments. The other student, Shamsul Abu Bakar [10], was a Ph.D. student interested in Civil War battlefield sites that were being preserved and interpreted at National Parks and National
Monuments. There is not enough room to present each student’s entire project. Instead, each student’s use of eye-tracking will be described, and their results will be summarized. Two or three examples from each study will demonstrate how eye-tracking can assist in identifying spatial and physical content and other factors influencing the content of CIM dimensions.

Some readers might ask, what do pedestrian environments and battlefields have to do with visual assessment? It is essentially the same process. Landscape preferences are an easy way for people to express their reactions, drawing on what they know about different types of landscapes. It is very natural. People make many preference decisions every day; examples are what to eat for breakfast, what to watch on television, and what tasks to complete that day, to mention a few. So, the context and mental factors of a preference decision may differ. However, the process of expressing preference is the same, whether it is a preference for a battlefield interpretation or a reaction to a natural environment. The mental process of making these is essentially the same. CMI identifies landscape content underlying the preferences, and eye-tracking assists in understanding the nature of the content and how it affects the experience of a landscape.

4. Results

4.1. Pedestrian Streetscapes

Seth’s research project intended to identify the physical and spatial elements of urban streetscapes that make the environment more attractive to pedestrians, thus encouraging them to walk to work. He obtained permission from several offices in the Washington, D.C. area to email their employees, asking if they would volunteer to complete an online preference survey. Seventy-five office employees completed the survey and provided their preferences on a 5-point Likert scale for 38 scenes of urban street landscapes. The mean scene preferences ranged from 1.85 to 4.52. The CMI analysis of these scenes generated five dimensions. Since the survey was online and on different computers, the computer could not be used to track eye movement. Instead, the survey participants were told to use their computer mouse and click anywhere within the image on the element most influential to their preference rating for that scene. These are called hot-points. The responses were combined to create heat maps for each image. If the hot-points on the heat map are close together, the colors on the map are yellow and red, appearing to be hot. The heat maps can assist in interpreting the preference dimensions in three ways:

1. They can help identify the common stimuli of each dimension.
2. They can help determine whether the stimulus is a physical element or a spatial arrangement.
3. Heat maps can help determine factors causing variation in the mean preference of the scenes within a dimension.

All of the scenes in Dimension 1, the dimension with the highest mean rating ($\bar{x} = 3.72$), depict a content of broad ground planes with vegetation providing a sense of spatial containment. The heat map for Scene A2 in this dimension (Figure 1) has the second highest mean preference ($\bar{x} = 4.2$), which supports this as the common content for this dimension.

The hot-points are tightly clustered on or under the vegetation that forms a canopy over the broad pedestrian space. Hot-points clustered in this manner suggest a spatial aspect to the content. Note that the bollards in the scene do not have hotspots on them, suggesting that they do not influence visual preference. The role of bollards as an influential content will be important in another image shown later in this paper.

The scenes in the second dimension, Dimension B, have many of the characteristics of the Dimension A. However, the scenes also contain human-made objects that influence visual preference. Scene B7 ($\bar{x} = 3.75$) (Figure 2) is in this dimension and has hot-points clustered around different small human-made objects. The fact that human-made content detracts from preference is easily confirmed in the heat map. Most of the hot-points are on the bike rack, sign, and small structure, confirming that human-made structures are influencing preference for the scene. By comparing the means for the cluster of points
around the bike rack and those around the covered structure, it is also possible to determine if one of these human-made features had more of a negative influence on preference for this scene than the others.

Figure 1. Scene A2 from Dimension A is on the left with a heat map of the scene on the right. Dimension A has the highest mean preference among all the dimensions. Scene A2 is the second most preferred scene in Dimension A. The common content of Dimension A is a wide pedestrian area spatially defined by trees or vegetation. Most of the points on the heat map are located on or under the trees, defining a space. Thus, confirming spatial content is important to people’s preference for this dimension.

Figure 2. Scene B7 from Dimension B is on the left and the heat map for this scene is on the right. Dimension B has many characteristics similar to the scenes in Dimension 1, with a wide pedestrian area. However, the scenes in Dimension B also contain human-made content that detracts from the preference. There are hotspots on the bike rack, a small structure, and a sign.

Another scene, B3, is in the same dimension (Figure 3) but has an even lower mean preference ($\bar{x} = 2.75$). A large cluster of hot-points on the human-made grates and metal covers on the sidewalk indicates a stronger negative influence of this physical content, accounting for the lower mean preference. Eye-tracking helps us see what the content or stimuli of a dimension are and how the similar content can affect preference differently. The human-made content of this scene is visually very evident and has a more substantial negative impact than the human-made content of Scene B7.
Figure 3. Scene B3 from Dimension B is on the left and a heat map of this scene on the right. It is the third most preferred scene in this Dimension B. The common content of is a wide pedestrian area with human-made content that detracts from the preference. The heat map shows that almost all of the hot-points are located on the grates and steel covers on the sidewalk, causing a negative influence to preference.

Eye-tracking can show the complex relationships between elements affecting human preference. Lastly, Scene B5 (Figure 4) is also in the same dimension but with a much lower mean preference ($\overline{X} = 2.16$). The landscape in the scene is flat and open but lacks vegetation. The critical observation is that many hot-points are clustered on the bollards. As noted previously, there are no hot-points on the bollards in Scene A2 (Figure 1). Scene A2 has much more visual content in the image. This demonstrates the dynamic role that some content can play in people’s preferences. When a scene lacks other positive content, like vegetation, attention quickly falls on negative content that is not noticed in a more visually complex landscape. Negative content can have a stronger effect if there is no positive content to distract attention from potentially negative content. Eye-tracking can help to understand how and why a viewer reacts to a scene.

Figure 4. Scene B5 from dimension B is on the left and the heat map for this scene is on the right. It is the least preferred scene in dimension B. The common content of dimension B is a wide pedestrian area with human-made content that detracts from the preference. The effect of the human-made content is confirmed in the heat map.
4.2. Civil War Battlefields

Shamsul’s research was on Civil War battlefields, which are an unusual type of landscape for visual assessment. This type of content was very unusual for landscape assessment. It is a completely different context for landscape preference. He was not interested in scenic beauty or visual landscape assessment. He was interested in how a battlefield site could be preserved and interpreted. What do people want to see when they visit a historic battlefield site? Most battlefield sites are green and peaceful today. So, he used photographs of battlefield landscapes taken during and immediately following Civil War battles. Historic photos from many different battlefield sites were used in his study. Shamsul obtained permission from the U.S. Park service to administer an on-site preference survey to visitors at five different American Civil War battlefields:

1. Chickamauga and Chattanooga National Military Park;
2. Shiloh National Military Park;
3. Manassas National Battlefield Park;
4. Antietam National Battlefield;
5. Gettysburg National Military Park.

Participation was voluntary. Shamsul surveyed 242 battlefield visitors. A five-point Likert scale was used to determine visitor preference for thirty-eight battlefield scenes. A CIM analysis produced six dimensions.

- Dimension 1—Civilian ruins/outcomes of war.
- Dimension 2—Large artillery in battle positions.
- Dimension 3—Soldiers in encampments and defensive posts.
- Dimension 4—Civilian structures on battlefields.
- Dimension 5—Battlefield vantage points and viewsheds.
- Dimension 6—Soldiers and civilians in posed positions.

Eye-tracking confirmed the common content of each dimension. For example, the scenes in Dimension 2 (Figure 5) are depicted with their scene number and mean preference. All of the scenes contain large artillery pieces and soldiers. The mean preference for each scene is quite high. This demonstrates that the participants could express their preferences in different contexts. The participants understand that they are providing their preferences for what they want to see in historical exhibits about the Civil War, which would be different than what they would prefer in a scenic landscape assessment. To avoid shocking the survey participants, gruesome pictures of dead and injured soldiers on the battlefield were not included in the survey. However, as part of the survey, the participants were asked if they felt it would be appropriate to include gruesome battlefield images in interpretive exhibits if the viewers were given adequate warning ahead of time. Most people responded to this in the affirmative. Some felt it was important to include such images to understand that war is gruesome.

Computer eye-tracking could not be carried out in the field because of the equipment required. Instead, eye-tracking was carried out with a subset of participants in a laboratory on the Virginia Tech campus. Two connected computer monitors were set up for eye-tracking (Figure 6). One was set up to record the participants’ preferences and eye movements (eye-tracking). The second computer monitor was used by the researcher to monitor the survey participants and ensure the eye-tracking software worked correctly. For a more detailed description of the setup, see the paper by Bakar and Miller [11].

Several types of eye-tracking data were collected as the participants’ eyes moved over the image, and the participants made sense of the image. The heat maps are created by recording the collective locations on the scene where the participant’s eyes finally came to a rest. This produces a heat map similar to those in the first study described above. Figure 7 depicts a heat map for a scene in Dimension 2—large artillery and soldiers in battle position. Dimension 2 has the highest mean preference rating among all the dimensions. Knowing that they would see Civil War battlefield images, the survey participants found these to be interesting and highly preferred.
gruesome pictures of dead and injured soldiers on the battlefield were not included in the survey. However, as part of the survey, the participants were asked if they felt it would be appropriate to include gruesome battlefield images in interpretive exhibits if the viewers were given adequate warning ahead of time. Most people responded to this in the affirmative. Some felt it was important to include such images to understand that war is gruesome.

Computer eye-tracking could not be carried out in the field because of the equipment required. Instead, eye-tracking was carried out with a subset of participants in a laboratory on the Virginia Tech campus. Two connected computer monitors were set up for eye-tracking (Figure 6). One was set up to record the participants’ preferences and eye movements (eye-tracking). The second computer monitor was used by the researcher to monitor the survey participants and ensure the eye-tracking software worked correctly. For a more detailed description of the setup, see the paper by Bakar and Miller [11].

Figure 5. The factor analysis of the survey preferences revealed the scenes above for Dimension 2. The content in each of the scenes in Dimension 2 (\(\bar{x} = 4.01\)) is the presence of large artillery and soldiers in battle positions. Note the relatively high mean preferences. The respondents provided their preference for Civil-War-scene-appropriate interpretive exhibits.

Figure 6. Two shuttles were used for eye-tracking. Shuttle 1 monitors the survey participant as they provide their preferences for each scene. Shuttle 2 has an eye-tracking sensor at the bottom of the screen. The eye-tracking sensor is calibrated for each participant before they begin the survey.
Several types of eye-tracking data were collected as the participants’ eyes moved over the images. One of the important data points collected was the fixation time, or the average amount of time in milliseconds it takes the viewer’s eyes to become fixed on an object.

Many of the heat maps generated by the eye-tracking software demonstrated an interesting phenomenon. There were people in many of the scenes. The people, and particularly their faces, attracted the attention of the survey participants and were often the location of hotspots, as Figure 8 depicts. Most visual preference studies exclude people from the scenes being rated because the researchers want the participants’ reaction to the landscape features, not people.

Figure 7. This is an example of a heat map produced by eye-tracking for Scene 28 ($\bar{x} = 2.86$), including the average fixation time, or the average amount of time in milliseconds it takes the viewer’s eyes to become fixed on an object.

However, battlefield scenes often include people. In spite of the hotspots on the people, it did not seem to affect their reaction to the dimension content. It seemed that this was almost an innate reaction, that while their eyes fixated on the people, they were still able to react to the other content of the scene.

An excellent example of this is Scene 28 ($\bar{x} = 2.86$) in Figure 8, in a dimension whose content was “civilian structures in battlefields”. While all the images in this dimension contained civilian structures, the hotspots on scenes with people, like Scene 28, were on the doors, windows, or next to buildings where people were located. Perhaps the fact that the people are not close and that survey participants expected to see people allowed other content to still influence their preference.

Figure 8. The heat map indicates that the eyes of the survey participants tend to fix on people in Scene 28 ($\bar{x} = 2.86$).
Two additional data sources from eye-tracking are the fixation time (Figure 9) and scan-path. Fixation time is how long it takes a participant’s eye movement to become relatively fixed. The eye scanner records the participant’s eye movement as they comprehend a scene. More complex scenes may require more time to comprehend. This can be valuable information for the researcher who is trying to determine relevant scene content and influential factors. Fixation time may help interpret more complex CIM content.

![Figure 8. The heat map indicates that the eyes of the survey participants tend to fix on people in Scene 28 (\(x = 2.86\)).](image)

![Figure 9. An example of a scan path for Scene 27 (\(x = 2.86\)) that depicts the path that a survey participant’s eyes move across the scene as they make sense of a scene. The line depicts the path the eye followed, and the circles are the places where the eye paused when scanning the scene. The size of the circles depicts the relative amount of time each pause lasted.](image)

The scan-path records a participant’s eye movement as they comprehend a scene (Figure 9). The order of the scan or what attracts viewer attention first may help determine what CIM content is more dominant. The circles indicate where the eye pauses and changes direction. The size of the circle indicates the length of time that the eye is focused on a specific location in the scene. While trying to make sense of a scene, the survey participant’s scan-path could provide essential insights into understanding the content and its impact on people. It can be helpful to know how long it takes someone to understand what they are looking at and the order of what attracts their attention first.

5. Summary and Conclusions

This paper demonstrates how eye-tracking from two studies of very different landscape types can be used to reveal and interpret CIM content. Heat maps were generated by eye-tracking, and we are able to

- Assist in a CIM analysis in very different contexts.
- Identify physical content that is important to people’s preferences.
- Identify spatial content that is important to people’s preferences.
- Identify if there were multiple types of content influencing people’s preferences, both positively and negatively.
- Identify the path or the order of the elements that were looked at in order to reach a preference decision for a scene.
- Identify the complexity of landscape content and combinations of content by how long it takes people to reach a preference for scenes with different content.

This paper is a very modest examination of how eye-tracking can assist in a CIM analysis. Further studies are needed to demonstrate how understanding CIM content can
be turned into knowledge useful for landscape planning and design, thus confirming Lothian’s [2] claim that the subjectivist method “is more scientific and statistically rigorous” [2] (p. 93). Eye-tracking is a rapidly progressing area of research. More research is needed not only to identify preferred landscape content but also to understand how patterns and the density of that content influence preference. Computer technology is developing rapidly. We can now use glasses or goggles to track eye movement as people move through the environment. This new technology will enable eye-tracking research on scenic byways and other scenic designated landscapes.

The phrase “beauty is in the eye of the beholder” is an idiom. An idiom is a phrase that contains a figurative meaning that differs from the phrase’s literal meaning. In this case, the idiom meaning is that beauty is completely subjective with no agreement among people. This is ironic because the literal meaning may be more accurate. The eye is the way most people take visual information into their brain when determining what they like about the landscape. This paper demonstrates how eye-tracking can assist in identifying common content that influences people’s preferences in a landscape. So, maybe beauty is in the eye of the beholder—if you know how to find it.

**Funding:** This research received no external funding.

**Data Availability Statement:** The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

**Acknowledgments:** I am grateful to Seth Estep and Shamsul Abu Bakar for their excellent graduate student research that utilized eye-tracking. I also would like to thank Rachel Kaplan and Stephen Kaplan for their pioneering work developing the Content-Identifying Methodology (CIM).

**Conflicts of Interest:** The author declares no conflicts of interest.

**References**


**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.