SKETCHUP VALIDITY MODELING: A COMPARISON BETWEEN PHOTOGRAPHS AND 3D MODELS

By

Shawn Partin

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF ARTS

Environmental Design

2011

ABSTRACT

SKETCHUP VALIDITY MODELING: A COMPARISON BETWEEN PHOTOGRAPHS AND 3D MODELS

By

Shawn Partin

Landscape architects, engineers, urban planners, and designers are interested in the use of 3D models for presentation of designs. The use of 3D models is a centuries-old practice beginning with physical models and advancing to the digital realm. In this study, participants were asked to evaluate the aesthetic quality of 100 images including 25 images of a 3D Google SketchUp 7 (SU) model with 25 corresponding photographs of the same scene and 50 photographs used in a previous study by Burley (1997). There was found to be significant concordance between SU images and corresponding photographs ($p<0.025$). These findings suggest SU models are valid constructs for photographs, allowing the inference that SU models may be a valid and useful tool for designers when presenting designs to the general public or a typical private client.

Key Words: Q-sort method, 3D modeling, SketchUp, visual quality.

ACKNOWLEDGEMENTS

I would like to thank my committee members, Dr. Crawford and Dr. Schutski, for the help and guidance they provided me during my thesis work and throughout my degree. I would also like to thank my friends for their support and help with data collection. Finally, I would like to especially thank Dr. Burley for all his guidance and countless hours spent with me in meetings, phone calls, and answering emails to help make this research a success.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LITERATURE REVIEW

Visual Quality

The study of visual quality involves the analysis of the visual world and how humans perceive that world (Smardon, Palmer, & Felleman 1986, pp. vii-ix). Its roots lie in the basic anatomy of the human eye and the connection between the eye and brain. In other words, visual quality focuses on how humans see and process their environment. Through the study of visual quality, it is possible to build a guide on the construction of visual simulations such as photographs, drawings and models to name a few. Variations of this guide are available for any medium that is used to represent scenes. Examples for drawing include Lin (1993), Doyle (1999), and Reid (2002).

For a brief history on visual quality it is appropriate to turn largely to Smardon, Palmer, & Felleman (1986) who have compiled a comprehensive collection of background information, experiments, and insights on the subject up to 1986. Sections following *A Brief History of Visual Quality* will describe more recent research in visual quality related to this thesis.

A Brief History of Visual Quality

Visual quality dates back to as early the late 1330s when Ambrogio Lorenzetti graphically represented "the effects of good versus bad administrative and managerial decisions on urban and rural landscapes" (Smardon, Palmer, & Felleman 1986, p. 4). This depiction can be found in Siena in the Sala della Pace of the palazzo Publico (Smardon, Palmer, & Felleman 1986, p. 4).

During the Middle Ages, the concept of landscape paralleled the emergence of landscape painting as nature became no longer a background element but, in many cases, a focus of beauty. During this time, early gardens took the form of "geometric, usually symmetrical, architectural forms" where plants were carefully placed and nourished (Smardon, Palmer, & Felleman 1986, p. 5).

In the early 1900s, Humphrey Repton, "a highly successful eighteenth-century landscape gardener and practitioner" would draw or paint (usually watercolors) the landscape as it existed to present to his clients. In addition, he would illustrate the same space as if his design had been implemented to provide his clients with a better understanding of his concept (Smardon, Palmer, & Felleman 1986, p. 6).

It was around this time that American landscape values were largely affected by the English picturesque landscape and a fascination with wilderness. These influences encouraged the creation of America's landscape as national parks and to protect its aesthetic beauty (Smardon, Palmer, & Felleman 1986, p. 7).

With this in mind, "the concept of landscape as a source of pleasure and satisfaction is, historically, relatively recent. Even more recent is the popular fascination with the beauty of wild nature" (Smardon, Palmer, & Felleman 1986, p. 9). This is supported by the increasing number of nature preserves and national parks in recent history. "Wild and rural landscapes were [and still are] sources of pleasure and emotional satisfaction" (Smardon, Palmer, & Felleman 1986, p. 9).

Visual Mechanics

When representing a landscape, average eye height is assumed to be 5 feet above the ground (Smardon, Palmer, & Felleman 1986, p. 40). Smardon, Palmer, & Felleman (1986) describes the specific mechanics of the eye:

> "Each eye has a field of vision of approximately 166 degrees (head stationary, eye moving). The skull position creates a central area of 124 degrees where the images overlap. This is called the binocular field. It is of particular importance due to the stereo nature of depth perception which occurs in this region. Within the binocular field is a narrow region of highest acuity, the foveal (macular) field… On either side is a monocular field of 42 degrees, containing inputs from only one eye [peripheral vision]… The total resultant cone is 208 degrees…" (Smardon, Palmer, & Felleman 1986, pp. 40-41).

The eyes and head both move independently allowing the viewer to see beyond 208 degrees. For this reason, it is common to create wide-view images and 3D models (either physical or digital). This allows the viewer of the image or model the ability to move their eyes and head to focus on different visual elements similar to being in a real setting.

Environmental Perception

As defined by Smardon, Palmer, & Felleman (1986, p. 65), the environment refers to those features surrounding a perceiving organism.

Theories of Picture Perception

It would be incorrect to assume that an object and a picture of the same object would be perceivably the same. The picture is simply a representation of the object itself as expressed in Belgian artist Rene Magritte's "La trahison des images" (The treachery of images). The artist was making a point that no matter the realism of a representation of an object, it is still only a representation. Although a great deal of information may still be obtained from the image, there are obvious limitations. For instance, no further views may be obtained by changing one's angle of view when looking at a still image. Even in the case of a digital simulation or a physical scale model of a landscape, one cannot feel the wind or walk through the space as if it were a real landscape.

Smardon, Palmer, & Felleman (1986, pp. 71-72) discusses four theories of picture perception: Conventions, Similarity, Elements of Light, and Optic Information. *Conventions* refers to the notion that pictures are graphic representations composed of common symbols arranged in a matter similar to the grammar and syntax in a language (Smardon, Palmer, & Felleman 1986, p. 71). *Similarity* refers to pictures as a "likeness" to the objects they represent in that they display common qualities or characteristics of those objects (Smardon, Palmer, & Felleman 1986, p. 71). The third theory, *Elements of Light*, focuses around projective geometry where an image emits light to a stationary viewer in a specific location in relation to the image (Smardon, Palmer, & Felleman 1986, p. 72). Finally, *Optic Information*, states that pictures can provide the same optical information as the world but cannot recreate sense data (wind, temperature, smell, things of the like) (Smardon, Palmer, & Felleman 1986, p. 72).

Principles of Pictoral Depiction

Visual simulations are commonly used in representing both existing and proposed site conditions in design. Construction of simulation images is based on visual clues. If an image contains the proper elements, then it will be perceived as a valid simulation. Smardon, Palmer, & Felleman (1986, pp. 73-74) discuss principles of pictoral depiction in relation to stationary monocular perspective. These principles include texture perspective, size perspective, linear perspective, aerial perspective, relative upward location, shift of texture density, completeness or continuity of outline, and transitions between light and shade. These principles are represented graphically in Figure 1.

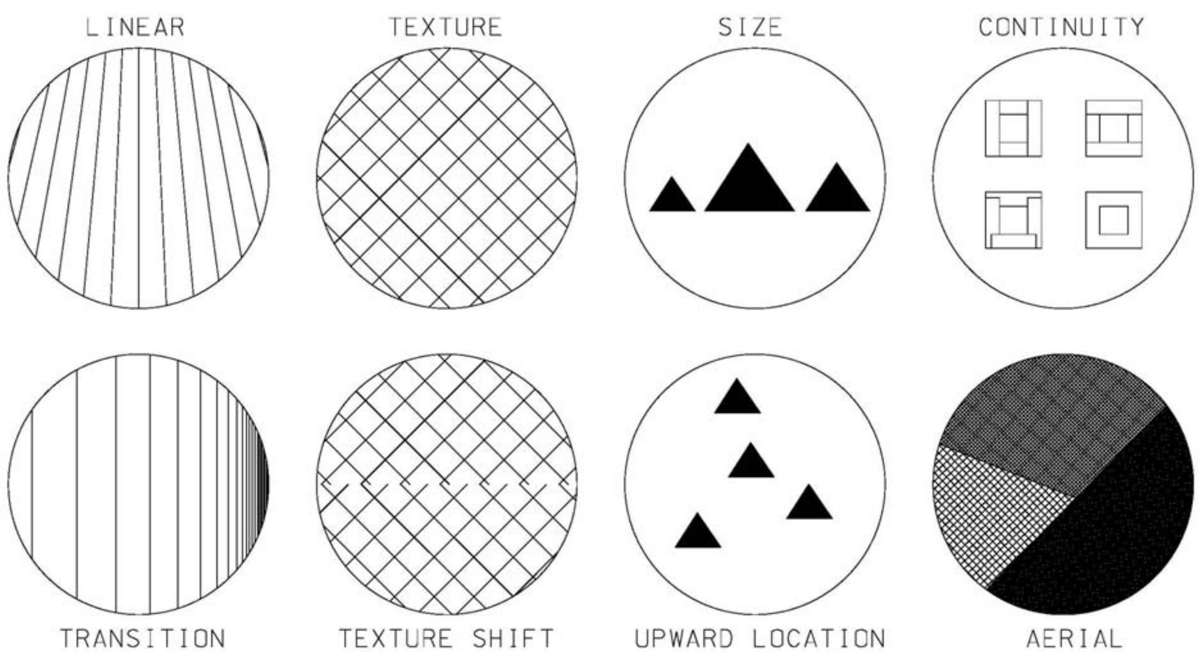

Figure 1: The stationary monocular clues that specify perspective.

Three additional principles of pictoral depiction are made possible by motion; binocular perspective (a concept used in modern 3D films where images are displayed simultaneously of the same scene, but offset slightly), motion perspective (perceiving movement of either the camera or an object on screen), and shift in the rate of motion (objects further away appear to move more quickly than objects up close as a camera turns or pans) (Smardon, Palmer, & Felleman 1986, pp. 75-76). In addition, motion pictures can portray information of which still images are not capable. This includes movement-dependent information such as spatial representation and scenes that are larger than the viewing size due to the ability to move and turn (Smardon, Palmer, & Felleman 1986, p. 76).

Assessment of Landscape Quality

Assessment of landscape quality is important in design, especially when designing public-owned spaces or spaces designed for use by a large population. Assessing visual quality of a landscape is a subjective matter dependent both on the landscape and the observer. Traditional methods were case by case and involved written or verbal subjective opinions of a landscape. In a case such as this, it is hard to quantify the visual quality of a space. However, additional methods have been developed that help to better understand visual quality by quantifying qualitative data. These include surveys and questionnaires and perceptual preference assessment. Examples of these can be seen in Smardon, Palmer, & Felleman (1986, Chap. 10).

Surveys and questionnaires are simple to administer and distribute, however, the questions must be written with care to avoid bias by the author. A downfall of this method is intended respondents do not always respond for various reasons. It is possible to have bias in such a method due to, for example, dissatisfied respondents wanting to voice their opinion whereas a satisfied respondent may not feel the need to do so. Survey questions can be weighted with a number scale such as asking preference on scale of 1-10, allowing the researcher to quantify responses.

Perceptual preference assessment involves participants assigning a numeric score to images or actual landscapes in comparison to other images or landscapes. This allows the researcher to quantify the preference of the group. If each image is to receive a score from, for example, 1-10, it is possible to have inconsistencies between participants. For example, one participant may prefer an image and assign it a 10 whereas another participant may prefer the same image but only assign it an 8. There are corrective, mathematical measures for this situation, none the less, it remains a factor to consider. It is also possible that a participant may assign all images the same score.

Another perceptual preference assessment method is to have participants Q sort the images. Q sorting is described in the *Q Methodology and the Q Sort Method* section of this paper. Burley (1997) used a variation of Q sort in his research. In addition to this method, he was able to apply an equation to the results which makes possible fairly accurate predictions about visual quality in the geographic areas studied.

History of 3D Models

Models have been used to represent environments in three dimensions for centuries. This tool dates back as early as Egyptian and ancient Chinese tombs where they were used for the study of environments (Zube et al. 1987). Additionally, the use of models was reportedly studied by Peruzzi and Sangallo to create a reference for builders to work from during the construction of St. Peter's Basilica in Rome (Zube et al. 1987).

Since then, physical models have been used for design and decision-making. These models have been constructed using a variety of materials including wood, foam-core board, cardboard, plant material, plastic, and other materials to represent features such as ground plane, built elements, vehicles, landscaping, and scale figures. Although they are still used today by many professionals, physical models can be time consuming and costly to create.

Physical models are traditionally built for one of two main purposes, assisting a designer or planner in the design or planning process or presenting to a client, or to the public, an already completed design. Designers utilize completed models by either presenting the model itself or a depiction of the model such as a photograph, video, or graphic representation. Videos and photographs from within the physical model can be obtained through the use of a modelscope (Smardon, Palmer, & Felleman 1986, p.57) or handheld camera.

Physical models can be created at a variety of scales and detail. In creating a model to examine spatial relationships and view obstructions, a simple model is typically adequate. However, in creating a model that is intended to present a finalized design and to invoke similar emotion as the finished product, more detail is required. In large projects that need to show detail at several scales, several physical models must be constructed.

With the advent of personal computers came the ability to create basic digital 3D models for design. Early digital models required a considerable amount of manual input of data and returned simple, abstract graphics such as those created by Bergen et al. (1995) using

Vantage Point software. In early software programs, complex shapes such as those which are required to represent realistic trees were difficult and time consuming to build. With this in mind, trees were often represented using simple shapes such as tall, thin pyramids, cones, or a basic “line-“ or “stick-with-ball-on-top.”

Early digital models were best used for the analysis of spatial relationships since they did not possess the capability to create realistic scenes. This was largely due to the inability to create realistic surfaces or detailed shapes without extensive time and effort. In addition, early software programs could not represent atmospheric conditions such as fog or clouds. Although early digital models had several constraints, they still offered the ability to produce multiple 2D views from many angles once the model was complete. However, this advantage exists with physical models as well with the use of a modelscope or camera. At this point in time, it was likely more cost effective to build a physical model.

As computers and software advanced, digital models became more useful due to faster and easier construction. Software such as AutoCAD, 3D Studio Max, Google SketchUp and others allowed the user to quickly and accurately build and modify 3D spaces with basic geometry. In addition, software such as 3D Studio Max and Google SketchUp began offering the ability to create and modify atmospheric elements such as fog and shadows. It was now possible to create basic graphics from any angle within the digital model at a rapid rate once the model was completed.

3D Models Today

Digital, 3D models today are used frequently in presenting designs. Proof of this can be seen in online portfolios of most medium- to large-scale design firms. In addition, college students emerging from universities with various degrees in design have at least basic knowledge of how to create and use a 3D model in the design process as well as for presentation.

With frequent updates of 3D modeling and rendering software such as Google SketchUp, 3D Studio Max, Autodesk's Revit, Kerkythea, and Podium, designers are able to create more accurate and detailed models.

On Site vs. Photos

Previous studies have been performed by Hull & Stewart (1992), Stewart et al. (1984), and Shafer & Richards (1974) analyzing the effectiveness of photographs as representations of the landscape. Discovering the link between on-site visits and photography is important in determining the effectiveness of SU modeling as a representation of future landscapes. As noted by Bergen et al. (1995), "Photographs have been found to be adequate surrogates for actual scenes." Through these studies, it is possible to infer the relationship between on-site visits and SU modeling by comparing SU images and photographs. Select case studies are explored below to better understand the relationship between on-site visits and photographs.

In the study performed by Stewart et al. (1984), subjects were asked to visit a site while performing two tasks: make judgments about specific scenes based on *visual air quality, clarity, border* (clarity of objects' borders in the landscape), *color,* and *cloud cover* and; take a photograph of each scene. In addition, many of the subjects were asked to make judgments about the same scenes as represented by photographs. All three of these tasks were performed several times. Testing for this study began in December 1981 and was completed in April of the following year. Several weeks elapsed between tasks in an attempt to avoid subject bias caused by recent memory of visiting the site. A strong correlation was found between the judgments made on-site and the judgments made in the photographs in regard to visual air quality, showing a promising ability of photographs to effectively represent a landscape.

Hull and Stewart (1992) assess the relationship of scenic beauty judgments made by subjects both on-site and via photographs. Subjects were asked to make judgments about specific landscapes along a hiking trail. Three and nine months later, all subjects were asked to make judgments of photographs of the same landscapes. Similarities occurred in the group judgments but not in the individual judgments. It was suggested by Hull and Stewart (1992) that differences in the data may be caused by the mood and group dynamic of the subjects on the hiking trail. Some subjects may have been tired or soar from the hike and/or may have been influenced by comments made from others in the group.

In the research performed by Shafer & Richards (1974), three different groups of subjects were asked to assess landscape scenes; Group A performed an on-site visit and assessed the actual landscape, Group B assessed color slides of the landscape, and Group C assessed color photographs of the landscape. Subjects performed the assessment by rating a series of “27 bi-polar, seven-step adjective scales” (Shafer & Richards 1974). It was found that color slides and color photographs received similar reactions as the actual landscape if the slides and photographs displayed most of the variation that was present in the actual landscape. However, if the slides and photographs did not contain most of the variation present in the landscape, then responses were significantly different.

3D Models vs. Photos

A literature search has revealed several studies that analyze the relationship between real-world, or photographs of the real world, and 3D models. Among these studies are Bergen et al. (1995), Rohrmann and Bishop (2002), and Bishop and Rohrmann (2003). These studies directly relate to the research and hypothesis presented in the *Purpose of Study* section of this paper as they compare either photographs or real-world views to computer generated 3D models. Many of the methods discussed below differ from those presented in the *Methods* section. However, fundamentally, they remain the same. In all cases, the researcher is identifying participants' reactions to both 3D models and either the real world, or images of the real world. To provide a more detailed view, these case studies are presented below.

In a study by Bergen et al. (1995) images were created using Vantage Point software to imitate existing photographs. Participants were asked to compare aesthetics of the Vantage Point images against the photographs using a 10-point scale similar the Scenic Beauty Estimation method. It was determined that the Vantage Point images were not effective representations of the photographs.

This study relates to design by pointing out strengths and deficiencies in the Vantage Point images. The biggest downfalls of the generated images were a lack of detail and of man-made objects that were present in the corresponding photographs.

Rohrmann & Bishop (2002) analyze the subjective responses to a computer-simulation of a site walk-through (1) with sound/with shadows, (2) without sound/with shadows, (3) with sound/without shadows, and (4) without sound/without shadows. Subjects were instructed to watch two of the four simulations, as determined by the researchers (videos were grouped in pairs so every participant watched a video with and without sound but both videos either had shadows or did not have shadows), and respond to a series of scale-based and open-ended qualitative questions assessing subjects' impressions, associations, recall ability, and comprehensibility of the site.

The results of the study show favor to the validity of computer-simulated environments (CSE). Rohrmann & Bishop (2002) note that "quality and completeness clearly matters" and "for an environmental simulation to be considered valid it should evoke a similar set of responses as would a direct experience of the same environment". They call for additional studies that can help to generalize findings related to the validity of CSEs.

Bishop & Rohrmann (2003) studied the relationship between a real environment and a simulation of the same environment. In this study, Group 1 watched a daytime simulation of a site walk-through and filled out a response form about stimulus features, appraisal of the content, retention, and comprehension in relation to the simulated site. In addition, the same group performed a real site walk-through during the day and filled out a similar response form. Group 2 performed the same actions with the order of the real and simulated site walk-throughs reversed. Groups 3 and 4 followed the same procedure using a nighttime simulation and site walk-through. Finally, participants were asked to compare the two experiences.

The simulation was determined to be an acceptable representation of the natural environment. Although "appreciation of the site is less positive" using a simulated environment, Bishop & Rohrmann (2003) note that "the value of presenting environments via computer graphics, at least in this urban/park environment, is supported."

While it would be convenient if the reactions from 3D models matched those from photographs, Bishop and Rohrmann (2003) note, "The two sets cannot be expected to be identical because it is impossible to match the richness and complexity of reality. However, the more similar they are the greater the faith we can place in environmental simulation as a tool…" It may be possible to one day quickly generate a digital model as complex and detailed as reality; Hollywood computer-generated-images show potential for this (Bishop and Rohrmann 2003). However, the ability for designers and planners to achieve such complex models seems, for now, unfeasible considering the time and costs associated with this level of detail.

Q Methodology and the Q Sort Method

Q methodology began as a scientific theory based around a technique, known as Q sorting, that had yet to be proven as a viable tool in research of any kind. The theory was strengthened immediately by the use of statistical analysis. Q methodology as a research tool was further strengthened by its continued use and accuracy over the years since its introduction.

The origins of the Q sort method first begin with the introduction of Q methodology sent in a letter to *Nature* by William Stephenson (1902-1989) in 1935 (Brown 1993).

Stephenson, both a psychologist and a physicist, is considered the “father” of Q methodology and of the Q sort method. Although introduced in 1935, Stephenson further discusses and describes Q methodology in publications such as “Correlating Persons Instead of Tests,” Foundations of Psychometry: Four Factor Systems,” and *The Study of Behavior: Q-technique and Its Methodology* (Brown 1993). Additional discussions and descriptions of Q methodology and the Q sort method have been presented by Stephen Brown in a variety of texts, but most inclusive being “A Primer on Q Methodology” (Brown 1993) as well as by Pitt and Sube (1979) and Swaffield and Fairweather (1996) in which they describe how Q methodology is used, in detail, in their particular studies.

Early uses of the Q sort method, a key component of Q methodology, involved personality assessment. This was done by asking a panel of trained professionals to each give their opinion of a respondent’s personality. In order to accomplish this, each panelist was given a deck of cards with personality descriptors on them. The panelists were then asked to organize the deck of descriptors into piles from *least like the respondent* to *most like the respondent*. This process is known as Q sorting. Once all panelists had sorted their decks, the results from each panelist were compared, averaged, and analyzed. From this process, an assessment of the respondent’s personality emerged (Pitt and Sube 1979).

Although the process of Q sorting has basically remained the same over the years, its application has been found useful across many disciplines. A quick search of recent literature has revealed examples of Q methodology’s use in psychology (Westen et. Al. 1997), planning (Swaffield and Fairweather 1996), child psychology (Buckley et. Al. 2002), landscape planning and visual quality (Dearden 1984), education and literature (Boscolo and Cisotto 1999), workforce education (McKnight 2008), sociology (Previte et. Al. 2007), nursing (Akhtar-Danesh et. Al 2008), and mobile communications (Liu 2008) to name a few.

As Q methodology's popularity as a research tool grew and began to spread to other disciplines, it was determined that a journal devoted to Q methodology would be created. Originally started as a newsletter in 1977, *Operant Subjectivity: the Q Methodology Newsletter* later evolved into a journal entitled *Operant subjectivity: The International Journal of Q Methodology* in 1989 (Brown 1993). In addition, several websites have been created and maintained on the subject including www.operantsubjectivity.org, the official website of the *Operant Subjectivity* journal, www.qmeth.org, and www.lrz.de/~schmolck/qmethod maintained by Peter Schmolck, a university professor in Munich, Germany. These are only a few of many websites currently available on the web.

So with all the current uses of Q methodology and the Q sort method, what is it? Simply put, Q methodology is the "…science of subjectivity…" (Brown 1993). It is a way of quantifying what would otherwise be considered qualitative data. By quantifying qualitative data, Q methodology allows us to cross reference the opinions of multiple subjects in order to achieve a group opinion. Because of the way Q sorting is structured and analyzed, it allows us to compare and analyze multiple variables simultaneously. And it all begins with Concourse Theory.

Without Concourse Theory, there could be no Q methodology. The definition of a concourse as it pertains to Q methodology according to Brown (1993) is: "the flow of communicability surrounding any topic"; the running together of ideas. Essentially, this pertains to any opinions about any subject in any form of commentary or media, including conversation, in daily life. Subjects that make for a good concourse, leading to use as a topic involving Q methodology, are those which do not have a definitive answer. A topic in which one cannot simply open a book and say, “This is the answer, this is fact.” A concourse is built of opinion, of subjectivity.

As Brown (1993) states, “…Q methodology’s task [is] to reveal inherent structure of a concourse.” By taking a concourse from sources such as interviews, newspaper commentaries, essays, normative theories, article reviews, etc., one can use Q methodology to analyze the opinions around a subject and give it a structure, i.e. find a group opinion and to what degree that opinion applies to the whole. It is concourse, in its rawest form, which gives us the tools to create Q samples.

Q samples are a set of statements, images, sound clips, video clips, etc. that are derived from concourse. They are presented in a medium that can be lumped, initially, into one, all encompassing pile, such as a deck of cards, and eventually sorted. Traditionally, Q samples have been presented as words or statements printed on cards, each card containing one word, or one statement, about the same subject. Since its inception in 1935, Q samples have evolved, in some research, to be printed images on cards, one image per card. Images may or may not have obvious commonalities depending on the topic being researched. For example, research involving forest management applications to a specific plot of land may involve multiple images of the same area whereas research involving visual quality may include images from all over the world ranging from urban plazas to residential front lawns to natural landscapes showing no human impact whatsoever.

Regardless of what topic is being researched, one thing that is consistent across Q methodology is that Q samples need to encompass a generalization of the topic being studied. Due to the nature of concourse, the simple fact that it involves an open-ended topic, it is impossible to include all information in any one study about any individual topic. To attempt to do so would create a Q sample set so massive, it would be extremely difficult to find willing participants. Therefore, it is important to create a Q sample set that represents the research topic at hand from all basic angles in a general, concise fashion.

Figure 2 displays an example of what a Q sample deck might look like using images. For this example, let us assume we are researching visual quality. Particularly, we are looking for a correlation between respondents' reactions to the presence of concrete and the presence of vegetation in the Q samples. This example is fictional and is based only loosely on existing studies.

Figure 2: An example of what a Q sample deck may look like in a visual quality study. For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this thesis.

Once a Q sample set is established, the next step is to have it Q sorted by respondents. Each respondent is given the Q sample (for the sake of this description we will call it a deck of cards) and asked to sort them into, first, three piles from *least agree* to *neutral* to *most agree*. The number of cards that are placed in each pile at this point is irrelevant.

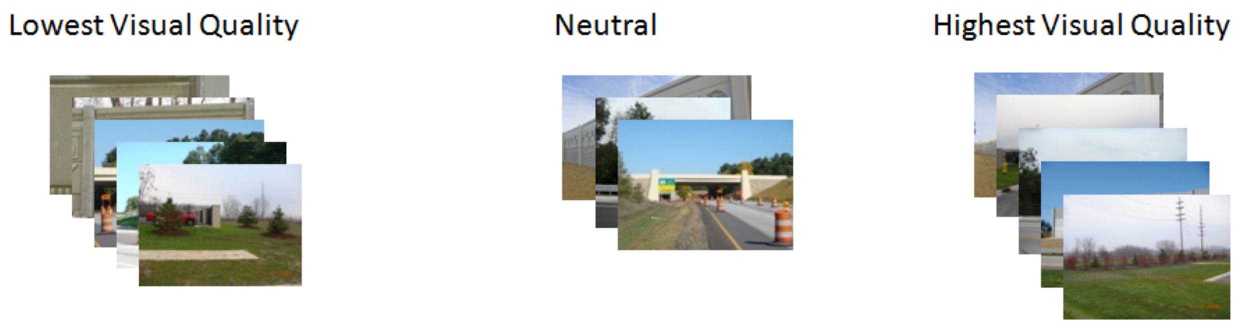

Figure 3: An example of what the first step in Q sorting may look like in a visual quality study.

Once the first step in the Q sorting process is complete, the subjects are asked to further distinguish which cards in the deck they most/least agree with by placing them into specific piles. For example, there may be seven piles, each with a specific number cards it is allowed to hold. There is no magic to having seven piles, studies may have any number of piles the researcher sees fit. As stated by Brown (1993), "The distribution is symmetrical about the middle… Both the range and distribution shape are arbitrary and have no effect on the subsequent statistical analysis, and can therefore be altered for the convenience of the Q sorter…"

In Figure 3, our fictional respondent has organized the thirteen cards in the Q sample deck into three piles. Because we are studying visual quality in this example, we ask respondents to evaluate the cards not by *least agree* to *most agree* but instead by *lowest visual quality* to *highest visual quality*. In other words, which images do they like least and which images do they like most. Again, at this stage in the sorting process, the number of cards in each pile is irrelevant.

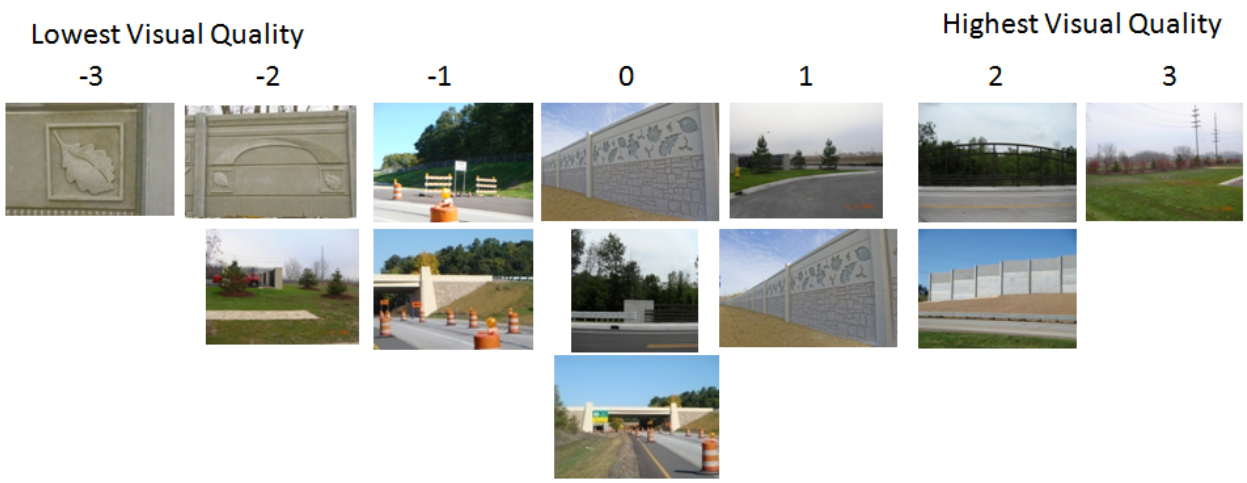

Figure 4: An example of what the second step in Q sorting may look like in a visual quality study.

Atop each pile is a ranking, or weight. For example, with seven piles the ranking would be as follows from *least agree* to *most agree*: -3,-2,-1, 0, +1, +2, +3. These are used for two purposes: (1) helping the respondent sorting the cards decide to what degree they agree/disagree with each card, and (2) to weight each pile, and therefore, each image during statistical analysis.

Figure 4 displays seven piles, each with a predetermined number of images allowed to be placed in it. This number was determined by the researcher. Each pile is weighted as described above. Our fictional respondent was asked to place the images in the piles based on which one was most favorable (3), which one was least favorable (-3), and a range of images in between.

Upon completion of data collection from all respondents, the data is analyzed. To do this, the data is first correlated. This correlated data is then factor analyzed. This process allows researchers to identify relationships between data sets of each individual. It also groups them into factors. Brown (1993) suggests thinking of these factors as "schools of thought." The number of factors is dependent upon how the respondents performed on their Q sorts. For example, if the study has twenty respondents, then there may be between one and twenty factors depending on whether some respondents seemed to 'agree' with one another when analyzing their Q sorts.

These factors are graphed, and then rotated in order to identify further relationships. The result is *rotated factor loadings*. Rotated factor loadings allow researchers to group respondents together overall based on their Q sorts. The rotated factor loadings (f) are weighted within their 'group' based on Equation 1. This step allows researchers to interpret the data for each Q sample in the deck. For a more in-depth explanation of the analysis process, including some of the reasoning behind each step, see Brown (1993).

$$w = f/(1-f^2)$$

Equation 1

This entire process used to be done by hand which was very tedious and took a great deal of time. This, of course, was dependent upon the amount of respondents and number of Q samples. Now, however, the whole statistical analysis can be done using computer software. Stricklin's PDQ for Windows is a paid-license software that appears to be the most up-to-date (last updated in 2004). Also out there, is PQMethod for DOS. Other software, which can be used online and is written in HTML, is Flash Q, QSorter, Web-Q, WebQSort, and Epimetrics Q-Assessor. All web-based software programs listed here require some knowledge in HTML programming language and many can be incorporated as a piece into an independent research website. A literature review revealed one study by Roth (2006) which linked results of on-site testing with results of web-based testing.

Once the statistical analysis has been completed, each Q sample is given a value based on how the respondents as a whole reacted to it. Using these individual values for each Q sample, the Q samples are placed back into the seven piles in order of how they performed from *least* to *most*. It is this final step in the analysis that allows the researcher to begin interpretation of the results.

In Figure 5, it is possible see how each image performed based on the averaged results of the fictional respondents. By placing the images back into the weighted Q sort piles, trends in the data can be identified. For example, it appears the respondents did not favor images that contained mostly concrete. On the other end of the scale, it can be observed that respondents tended to favor images which contained large quantities of grass. The center piles show us that respondents tended to feel neutral about images that contained relatively equal amounts of concrete and grass together.

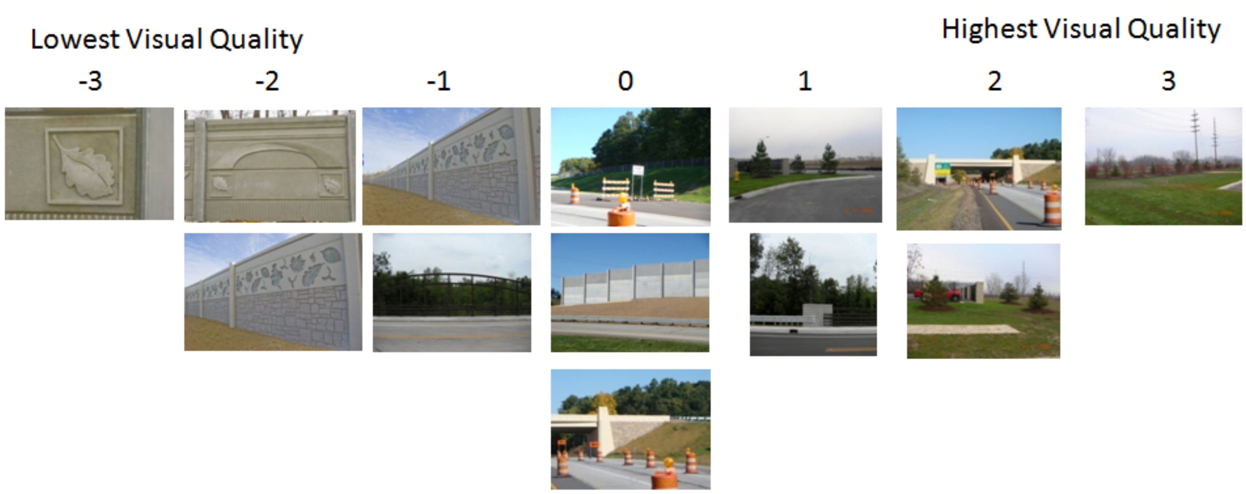

Figure 5: An example of what final results may look like in a visual quality study. Patterns may be identified using these piles and interpreted.

Although Q methodology began as a scientific theory, it has become an invaluable tool to many researchers since its introduction in 1935. What Q methodology allows researchers to do is identify and analyze distinct ways of thought related to a subject being researched without having to know specific information about the respondents who participate in a study. In addition, information that was previously only seen as qualitative is now able to be quantified and backed by concrete, statistical methods. Q methodology is a research tool that has been further strengthened by its continued used and accuracy since its inception.

PURPOSE OF STUDY

The purpose of this study is to determine the relationship, as perceived by human participants, between photographs of a site and 2D images of 3D Google SketchUp 7 (hereafter referred to as SU) models. Previous research has studied the relationships between photographs vs. drawings (Smardon, Palmer, & Felleman 1986, pp. 190-194), photographs vs. physical models (Smardon, Palmer, & Felleman 1986, pp. 190-194), photographs and/or slides vs. site visits (Shafer & Richards 1974, Hull & Stewart 1992, Stewart et al. 1984), photographs vs. computer-generated images (Oh 1994, Bergen et al. 1995), and a 3D fly-through of a model vs. a site visit (Rohrmann & Bishop 2002, Bishop & Rohrmann 2003) as discussed in the *Literature Review* section of this paper. Although there is no perfect substitute to visiting a site, it has been found that photographs, both black & white and color, and slides are the most effective alternative when a site visit is not practical (Shafer & Richards 1974).

This study hopes to determine if 2D images of an SU model are an effective substitute for color photographs. In most cases, it would make little sense to build an SU model of an existing site simply to present to somebody for the sake viewing the site when, in most cases, it would be both faster and cheaper to visit the site and take photographs. However, in the case of a designed but undeveloped site, where photographs are impossible, SU is an acceptable method for presenting the future development.

With this in mind, if SU images are found to relate closely with photographs, then it would be possible to suggest that SU is a more effective method for representing a design than other previously studied methods that have been found to have a poor relationship with photographs such as drawings or paintings. However, if SU images are found to have a poor relationship with photographs, then further research will be required to determine if the relationship between SU images and photographs correlate stronger or weaker than the relationship between other methods and photographs. This would help in determining the most effective method of those presentation tools currently available to designers.

This study attempts to determine the concordance between SU images and corresponding photographs using the Q-sort method. Participants are asked to sort groups of images that contain both photographs and corresponding SU images according to aesthetic preference. It is hypothesized that the SU images will receive similar scores as the corresponding photographs. The null hypothesis predicts that SU images and corresponding photographs will receive scores that are extremely different.

In an attempt to further enhance the results of this experiment, 50 seeded images from Burley (1997) are included in the Q-sort as well. It is hypothesized that the seeded images will receive scores that have strong concordance with scores received in the original study. The null hypothesis predicts that seeded images will receive scores that have no concordance or poor concordance with the scores received in Burley's (1997) original study.

This research will add to the body of knowledge that exists about which forms of media are effective in presenting landscape designs. The focus of this study is to discover, in general, where SU falls in level of effectiveness compared to other methods. Although other forms of 3D modeling have been tested against photographs such as Bergen et. Al (1995), SU, in particular, had yet to be tested prior to this study using these methods in determining SU's relationship to photographs.

Hypothesis 1: SU images will receive similar scores as the corresponding photographs.

Null Hypothesis 1: SU images and corresponding photographs will receive scores that are extremely different.

Hypothesis 2: Seeded images will receive scores that have strong concordance with scores received in Burley's (1997) original study.

Null Hypothesis 2: Seeded images will receive scores that have no concordance or poor concordance with the scores received in Burley's (1997) original study.

METHODS

Image Creation and Data Collection

Twenty-five digital photographs were chosen that included a variety of landscape elements such as grass, trees, sky, road, construction barrels, bridge, fence, concrete sound wall, buildings, vehicles, sidewalk, and road signs. From these images, several 3D models were created to imitate a corresponding scene for each image, resulting in 50 image-pairs. All models were created using Google SketchUp 7 (SU). After creating each model, a scene was set up in SU and exported as a 2D, digital image. If necessary, images were cropped in order to relatively match the dimensions of the corresponding photo. No other manipulations were made to the 2D images. An additional 50 images were randomly selected from Burley's (1997) image set and added to this image set. The final image set in this study contains 100 images. All images were numbered on the reverse side, placed in clear sheet protectors (to prevent smudging or weather damage), and presented to participants as 8.5" x 11" full page printouts. The complete set of images can be viewed in Appendix A, 1.1-1.3.

Fourteen participants were randomly selected to participate in this study on October 29, 2011 at the Meridian Township Farmers Market in Okemos, Michigan. Before beginning, participants were asked to sign a waiver stating their consent to participate (an example of the waiver can be seen in Appendix B). Participants were also made aware that they would receive a free cupcake upon completion of the sorting process.

Each participant was asked to sort 50 images, randomly selected from the 100-image set. The 100-image set was divided into two, randomly selected, groups of 50 images without replacement. Once both groups had been sorted, two new groups were created again using random selection. Sorting continued until the full 100-image set has been sorted for a total of seven times.

From the group of 50 images, five 10-image subgroups were created at random. Participants were given one subgroup at a time and asked to sort the scenic quality of the images from *most aesthetically pleasing* (or high scenic quality) to *least aesthetically pleasing* (or low scenic quality). A score of 1 represented high scenic quality while a score of 10 represented low scenic quality as compared to the other nine images in the subgroup. Each image received a score between 1 and 10; no two images were able to receive the same score within a subgroup. The score was determined by the order in which the participant placed the printed images on a table top. Fifteen minutes were allotted to sort each subgroup, if a participant completed the sorting process before the fifteen minutes was up, they were allowed to move to the next subgroup. The participant was finished when all five subgroups had been sorted. The entire 50 image group took roughly fifteen minutes to sort on average.

Once a participant finished sorting a 10-image subgroup, the order was recorded by the researcher and the participant was given the next subgroup. When all subgroups had been sorted and recorded, the participant was allowed to take a cupcake. At this time, participants were allowed to ask questions about the study if they desired. If questions were asked, they were answered away from current sorting participants so as not to influence decisions made by current sorters.

Calculations

Data was organized into a table displaying the score for each image by judge. Scores for each image were then totaled. From there, the data was analyzed in two ways. First, SU Images were compared with SU Photographs using Kendall's Coefficient of Concordance (Daniel 1978) to determine the relationship between the two 25-image sets. Second, the ranks of the 50 images from Burley (1997) were compared with the ranks from the original study using the same method.

SU Images vs. SU Photographs

A raw score was given to each image based on the total of the ranks from all judges. Based on this raw score, a new rank was given to each SU image out of 25 (because there are 25 SU images in total). If the raw score of an image was equal to that of another image in the set, the average of the ranks was assigned to all images. This was referred to as a tie. This process was repeated for SU photographs as well. A complete list of raw scores and ranks for SU Images and SU Photographs can be seen in Table 1 of the *Results* section.

Following the process in Daniel (1978) for Kendall's Coefficient of Concordance, the sum of the ranks for each corresponding image was found. Corresponding images were determined by matching views as in Figure 6 and had been originally numbered as such, so corresponding pairs can be represented as images 1:26, 2:27, 3:28, … , 23:48, 24:49, 25:50. The sum of the ranks were then squared (Rj^2).

Figure 6: Corresponding images, pair 11:36. (Left) SU Image. (Right) SU Photograph.

The general value for W was found using Equation 2.

$$W = (12*S)/[m^2 n(n^2-1)]$$

Where:
W = Kendall's coefficient of concordance
S = (sum of the ranks – mean)^2
m = number of sets (SU Images and SU Photographs)
n = number of images in each set

Equation 2

The value for W was then adjusted for ties using Equation 3.

$$W = (12*S)/[m^2 n(n^2-1)-m*\Sigma(t^3-t)]$$

Where:
W = Kendall's coefficient of concordance adjusted for ties
S= (sum of the ranks – mean)^2
m = number of sets (SU Images and SU Photographs)
n = number of images in each set
$\Sigma(t^3-t)$ = the sum of the number of observations in all sets of rankings tied for a given rank

Equation 3

Finally, the chi-square value was determined using Equation 4.

$$X^2 = m(n-1)W$$

Where:
X^2 = chi-square value
m = number of sets (SU Images and SU Photographs)
n = number of images in each set
W = Kendall's coefficient of concordance adjusted for ties

Equation 4

The X^2 value of was then compared to "A Table of Percentage Points of the X^2 Distribution" in Daniel (1978, p. 452).

Seeded Images vs. Original Score

The scores received for the 50 seeded images borrowed from Burley (1997) were compared to the scores in the original study. For the sake of this explanation, the set of scores received in this study will be referred to as *Seeded Images A* (SIA) and the set of original scores will be referred to as *Seeded Images B* (SIB).

For SIA, a raw score was given to each image based on the total of the ranks from all judges. Based on this raw score, a new rank was given to each SIA out of 50 (because there are 50 SIA images in total). If the raw score of an image was equal to that of another image in the set, the average of the ranks was assigned to all images. This was referred to as a tie. This process was repeated for SIB as well. A complete list of raw scores and ranks for SIA and SIB can be seen in Table 3 of the *Results* section.

Following the process in Daniel (1978) for Kendall's Coefficient of Concordance, the sum of the ranks for each corresponding image was found. Corresponding images in this case are the exact same image with unique scores from SIA and SIB. The sum of the ranks were then squared (Rj^2).

The general value for W was found using Equation 5.

$$W = (12*S)/[m^2 n(n^2-1)]$$

Where:
W = Kendall's coefficient of concordance
S = (sum of the ranks – mean)^2
m = number of sets (SIA and SIB)
n = number of images in each set

Equation 5

The value for W was then adjusted for ties using Equation 6.

$$W = (12*S)/[m^2 n(n^2-1)-m*\Sigma(t^3-t)]$$

Where:
W = Kendall's coefficient of concordance adjusted for ties
S= (sum of the ranks – mean)^2
m = number of sets (SIA and SIB)
n = number of images in each set
$\Sigma(t^3-t)$ = the sum of the number of observations in all sets of rankings tied for a given rank

Equation 6

Finally, the chi-square value was determined using Equation 7.

$$X^2 = m(n-1)W$$

Where:

X^2 = chi-square value

m = number of sets (SIA and SIB)

n = number of images in each set

W = Kendall’s coefficient of concordance adjusted for ties

Equation 7

The X^2 value of was then compared to “A Table of Percentage Points of the X^2 Distribution” in Daniel (1978, p. 452).

RESULTS

Following the explanation in the *Methods* section, calculations were performed on both data sets; *SU Images vs. SU Photos* and *SIA vs. SIB*.

SU Images vs. SU Photographs

As discussed in the *Methods* section, a rank was given to each SU image out of 25 (because there are 25 SU images in total) based on the raw scores from all judges. If ties in raw scores occurred, they were accounted for (i.e. SU Images 9 and 11 both had a raw score of 51, therefore the average of ranks (17+18)/2 = 17.5 was given). This process was repeated for SU photographs as well. Table 1 shows the relationship between all image pairs with the raw scores and individual ranks for each pair. Image pairs are aligned horizontally and can be read as image pair 1:26, 2:27, 3:28, … 23:48, 24:49, 25:50.

Table 1: Raw scores and ranks for SU Image and SU Photo image pairs.

SU Images (m1)			SU Photographs (m2)		
SU Image	Raw Score	Rank	SU Photo	Raw Score	Rank
1	40	7.5	26	32	2
2	41	9	27	29	1
3	64	25	28	54	22
4	61	23.5	29	62	25
5	53	19.5	30	55	23.5
6	39	6	31	38	6.5
7	44	11	32	50	17
8	50	16	33	46	12
9	51	17.5	34	50	17
10	56	22	35	50	17
11	51	17.5	36	48	15
12	49	15	37	51	19.5
13	53	19.5	38	51	19.5
14	34	3.5	39	38	6.5
15	40	7.5	40	33	3

Table 1 (cont'd)

16	32	1	41	41	8.5
17	48	13.5	42	45	10
18	34	3.5	43	37	5
19	33	2	44	46	12
20	54	21	45	41	8.5
21	38	5	46	46	12
22	61	23.5	47	55	23.5
23	45	12	48	47	14
24	43	10	49	52	21
25	48	13.5	50	35	4

Following the process in Daniel (1978) for Kendall's Coefficient of Concordance, the sum of the ranks for each corresponding image was found. The sum of the ranks were then squared (Rj^2). The *mean value of the ranks*, *m*, and *n* values are given in Table 2 where *m* is the number of sets (SU Images and SU Photographs) and *n* is number of images in each set.

Table 2: Mean value of the ranks, m, and n values for SU Images vs. SU Photographs.

Mean value of the ranks	26
m	2
n	25

The general value for W was found using Equation 2. By substituting the S, m, and n variables for their corresponding values, it was computed:

$$W = (12*4336)/[2^2*25(25^2-1)] = 0.834$$

Where:
W = Kendall's coefficient of concordance

The value for W was then adjusted for ties using Equation 3. By substituting the S, m, n, and $\Sigma(t^3-t)$ variables for their corresponding values, it was computed:

$$W = (12*21236)/[2^2*25(25^2-1)-2*108] = 0.837$$

Where:
W = Kendall's coefficient of concordance adjusted for ties

Finally, the chi-square value was determined using Equation 4. By substituting the X^2, m, n, and W (adjusted for ties) variables for their corresponding values, it was computed:

$$X^2 = 2(25-1)0.837 = 40.164$$

Where:
X^2 = chi-square value

The X^2 value of 40.164 with 24 degrees of freedom was then compared to "A Table of Percentage Points of the X^2 Distribution" in Daniel (1978, p. 452). It was found that SU images and SU photos have significant concordance to a p-value of <0.025. The primary null hypothesis is therefore rejected and the alternate hyothesis is accepted.

These findings were then graphed. Figure 7 shows the rank values for each image pair. For each image pair, the red bar indicates the rank value of the SU Image in the image pair while the blue bar indicates the rank value of the SU Photo. From this, Figure 8 is derived, displaying the differences between image ranks compared to the average difference in ranks among all image pairs. For each image pair in this figure, the blue bar represents the difference of ranks between the images in the pair, the green bar represents the average difference among ranks for the entire set, and the red bar represents the deviation of the difference of ranks in an image pair from the average difference among ranks for all pairs.

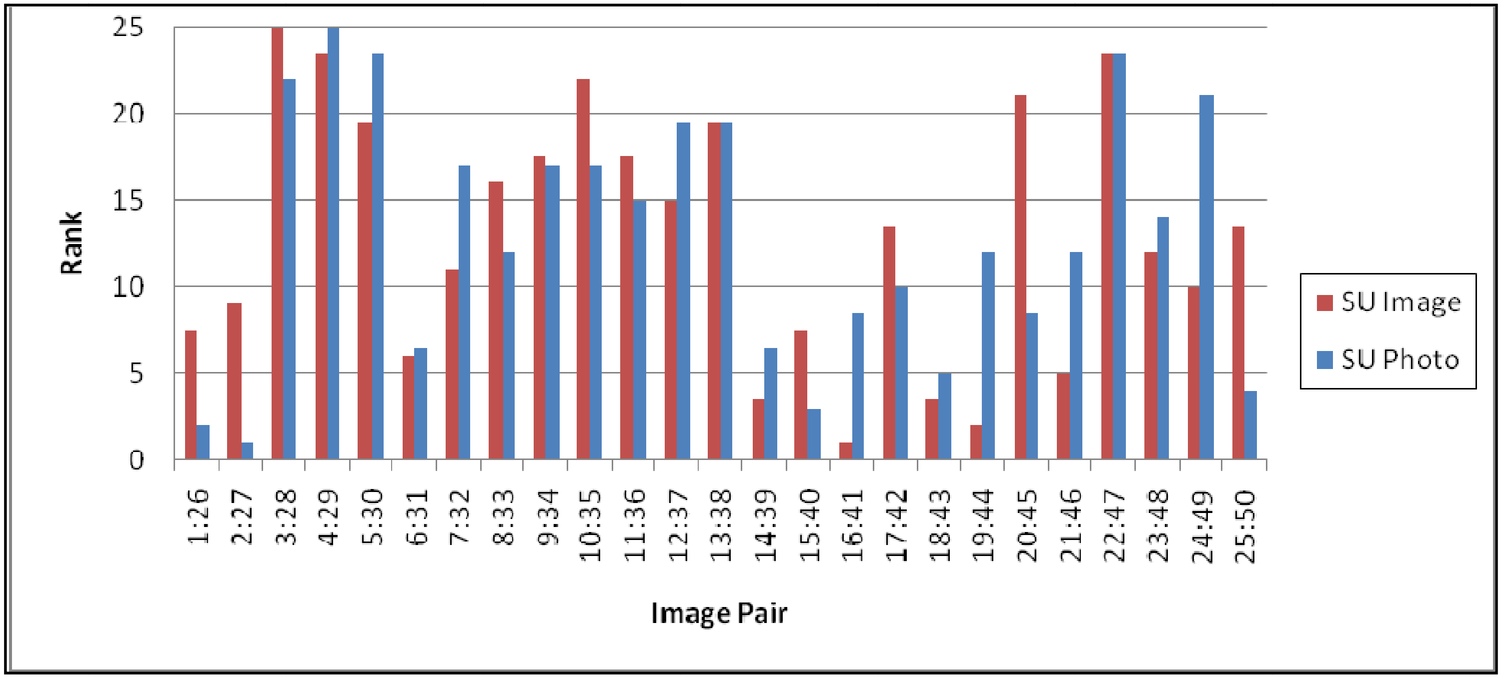

Figure 7: Image pair rankings for SU Image and SU Photo image pairs.

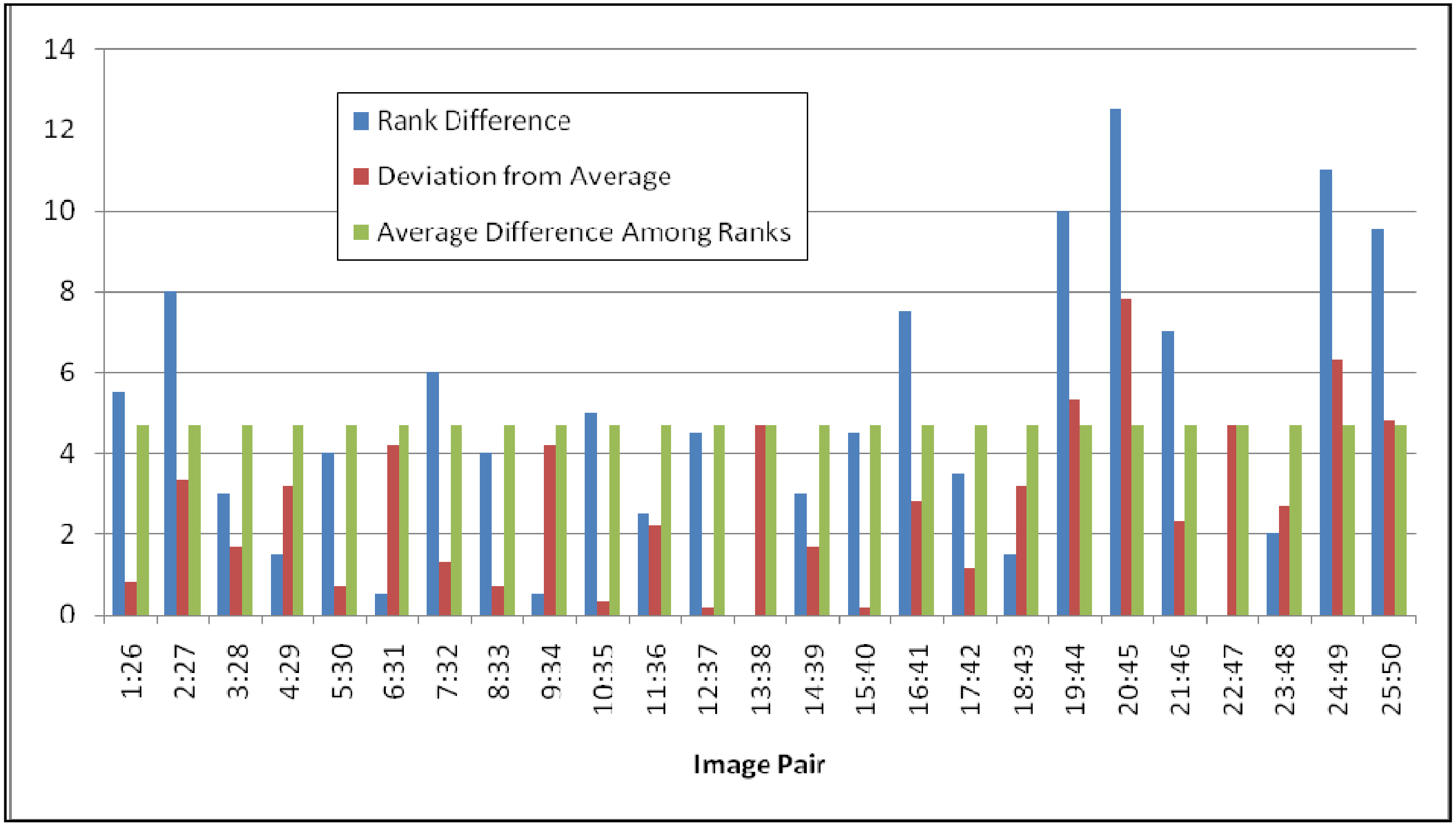

Figure 8: Rank differences by image pair for SU Images vs. SU Photos.

SIA vs. SIB

As discussed in the *Methods* section, a rank was given to each SIA image out of 50 (because there are 50 SIA images in total) based on the raw scores from all judges. If ties in raw scores occurred, they were accounted for (i.e. SU Images 59 and 67 both had a raw score of 38, therefore the average of ranks (36+37)/2 = 36.5 was given). This process was repeated for SIB as well. Table 3 shows the SIA vs. SIB relationship for all images with the raw scores and individual ranks for image.

Table 3: Raw scores and ranks for SIA and SIB images.

Seeded Images A (m1)			Seeded Images B (m2)	
Image	Raw Score	Rank	Score from Burley (1997)	Rank
51	21	15	64	21
52	23	20	48	13
53	20	12.5	36	4
54	13	2.5	40	7
55	14	4.5	42	9
56	9	1	51	14
57	52	44	122	45.5
58	27	26	97	34
59	38	36.5	103	37
60	18	10	62	18
61	44	39	96	33
62	27	26	45	11.5
63	33	33	108	41
64	39	38	106	39.5
65	26	23.5	73	25
66	63	48	128	49
67	38	36.5	102	36
68	17	8	62	18
69	21	15	75	28
70	65	49.5	124	47.5
71	22	17.5	41	8
72	18	10	62	18
73	28	28.5	65	22
74	65	49.5	106	39.5
75	26	23.5	37	5

Table 3 (cont'd)

76	23	20	26	2.5
77	24	22	26	2.5
78	22	17.5	74	26.5
79	16	7	74	26.5
80	18	10	54	15
81	32	31	63	20
82	21	15	43	10
83	14	4.5	45	11.5
84	15	6	60	16
85	48	42	120	44
86	62	47	137	50
87	20	12.5	66	23
88	33	33	90	29.5
89	58	46	113	42
90	27	26	98	35
91	33	33	104	38
92	53	45	124	47.5
93	34	35	92	31
94	23	20	39	6
95	46	40.5	93	32
96	30	30	90	29.5
97	13	2.5	72	24
98	28	28.5	25	1
99	50	43	122	45.5
100	46	40.5	116	43

Following the process in Daniel (1978) for Kendall's Coefficient of Concordance, the sum of the ranks for each corresponding image was found. The sum of the ranks were then squared (Rj^2). The *mean value of the ranks*, *m*, and *n* values are given in Table 4 where *m* is the number of sets (SIA and SIB) and *n* is number of images in each set.

Table 4: Mean value of the ranks, m, and n values for SIA vs. SIB.

Mean value of the ranks	51
m	2
n	50

The general value for W was found using Equation 5. By substituting the S, m, and n variables for their corresponding values, it was computed:

$$W = (12*36638)/[2^2*50(50^2-1)] = 0.880$$

Where:
W = Kendall's coefficient of concordance

The value for W was then adjusted for ties using Equation 6. By substituting the S, m, n, and $\Sigma(t^3\text{-}t)$ variables for their corresponding values, it was computed:

$$W = (12*166688)/[2^2*50(50^2-1)-2*240] = 0.880$$

Where:
W = Kendall's coefficient of concordance adjusted for ties

Finally, the chi-square value was determined using Equation 7. By substituting the X^2, m, n, and W (adjusted for ties) variables for their corresponding values, it was computed:

$$X^2 = 2(50-1)0.880 = 86.290$$

Where:
X^2 = chi-square value

The X^2 value of 86.290 with 49 degrees of freedom was then compared to "A Table of Percentage Points of the X^2 Distribution" in Daniel (1978, p. 452). It was found that there is a significant concordance between the results from this study and those found in Burley (1997) to a p-value of >0.005. The secondary null hypothesis is therefore rejected and the alternate hyothesis is accepted.

These findingds were then graphed. Figure 9 shows the ranking relationships from this study (SIA) and from Burley (1997) (SIB) for each image. For each image in this figure, the red bar indicates the rank earned in this study while the green bar indicates the rank earned in the original study by Burley (1997). From this, Figure 10 is derived, displaying the rank differences for each image compared to the average difference among ranks for all images in the set. For each image in this figure, the blue bar indicates the rank difference between SIA and SIB, the green bar indicates the average difference among ranks for the entire set, and the red bar indicates the deviation of the difference of ranks for an image from the average difference among ranks for all images.

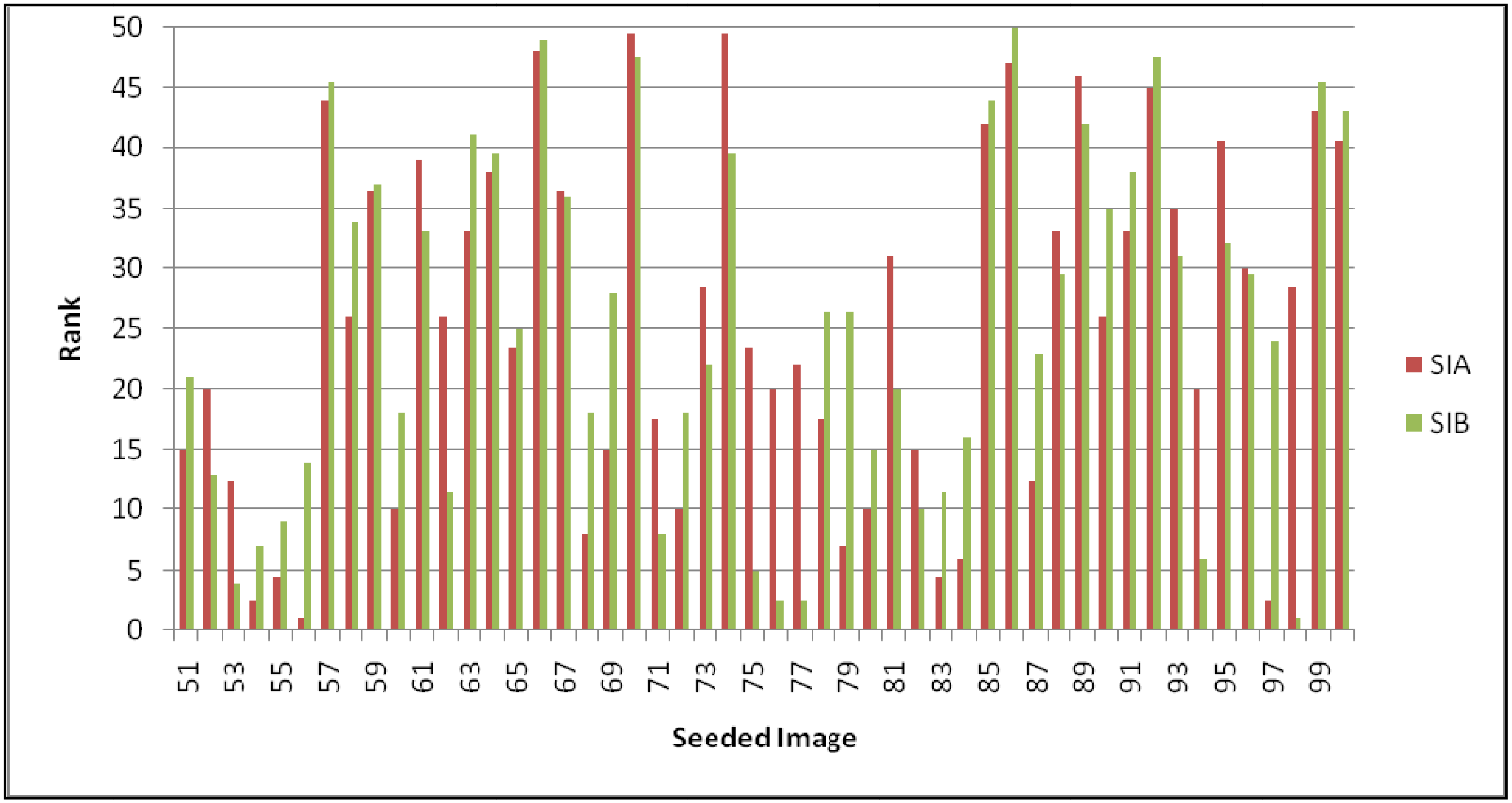

Figure 9: Seeded image rankings for SIA vs. SIB

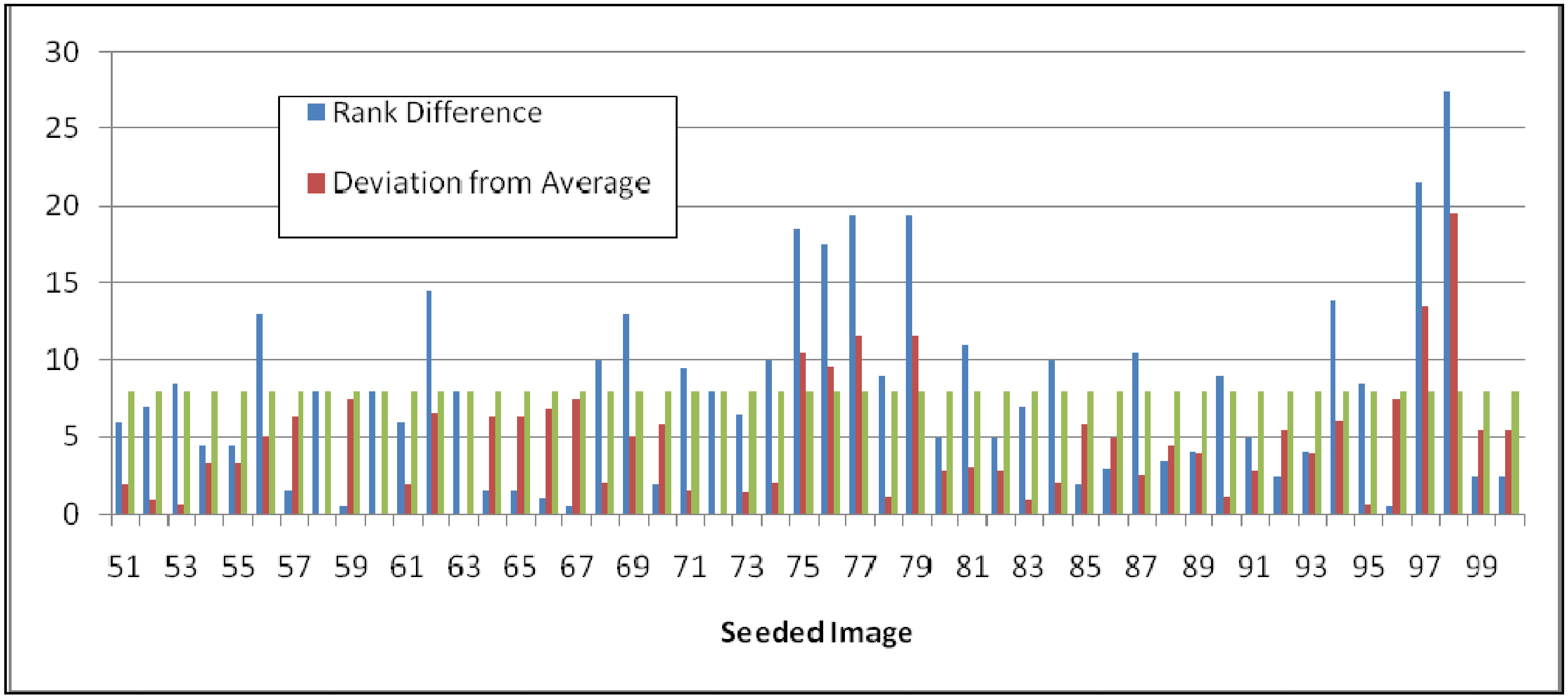

Figure 10: Rank differences by image for SIA vs. SIB.

DISCUSSION

The results show there is significant concordance between SU images and SU photographs to a p-value of <0.025, suggesting a strong link between 2D images of a 3D model and photographs. As noted by Bergen et al. (1995), "Photographs have been found to be adequate surrogates for actual scenes." Therefore, it can be inferred that 2D images of a 3D model may also be acceptable surrogates for actual scenes.

Additionally, the results imply that the average person, when presented with an SU image of a design, whether proposed or existing, will more than likely see an image of a design rather than an image of a model. To clarify, when presented with an image of a model, people will base their opinion of the design on the content of the image rather than on the quality of the SU model.

To better understand the relationship between SU images and SU photographs, six pairs of corresponding images are analyzed below. The six pairs, chosen to represent the full spectrum of rank differences, are 13:38 (with a rank difference of 0), 22:47 (with a rank difference of 0), 12:37 (with a rank difference of 4.5), 5:30 (with a rank difference of 4), 24:49 (with a rank difference of 11), and 20:45 (with a rank difference of 12.5). Pairs 12:37 and 5:30 are nearly in line with average rank difference of 4.68 while pairs 13:38, 22:47, 24:49, and 20:45 are outliers. Table 5 displays rank differences for all image pairs.

Table 5: SU Image and SU Photo image pairs with rank difference, average rank difference, and deviation from the average rank difference.

Image Pair		Rank Difference	Deviation from Average
1	26	5.5	0.82
2	27	8	3.32
3	28	3	1.68
4	29	1.5	3.18
5	30	4	0.68
6	31	0.5	4.18
7	32	6	1.32
8	33	4	0.68
9	34	0.5	4.18
10	35	5	0.32
11	36	2.5	2.18
12	37	4.5	0.18
13	38	0	4.68
14	39	3	1.68
15	40	4.5	0.18
16	41	7.5	2.82
17	42	3.5	1.18
18	43	1.5	3.18
19	44	10	5.32

Table 5 (cont'd)

20	45	12.5	7.82
21	46	7	2.32
22	47	0	4.68
23	48	2	2.68
24	49	11	6.32
25	50	9.5	4.82
Average Rank Difference		4.68	

Image pair 13:38 in Figure 11 depicts a highway sound wall and contains wall, fence, guardrail, car, dirt, straw, trees, and sky. The rank difference of these images is 0, meaning that both the SU image and SU photograph received the exact same rank out of 25 possible ranks. Obvious similarities are the angle from which the image or photograph was taken and the apparent materials of the wall and fence. Differences include shadow patterns, ground texture, and the presence or absence of clouds. Despite these rather noticeable differences, these images were ranked equally by participants.

Image pair 22:47 in Figure 12 depicts a highway rest area and includes grass, mulch, dirt, concrete sidewalk, trees, building, picnic table, trash receptacle, truck, and parking lot. The rank difference of these images is 0, meaning that both the SU image and SU photograph received the exact same rank out of 25 possible ranks. Obvious similarities are the angle from which the image or photograph was taken, lack of clouds, apparent material of building, and general shadow pattern. Differences include ground texture and color, amount of detail in the truck, exact placement of the picnic table and trash receptacle, and form of foreground trees. Again, despite these rather noticeable differences, these images were ranked equally by participants.

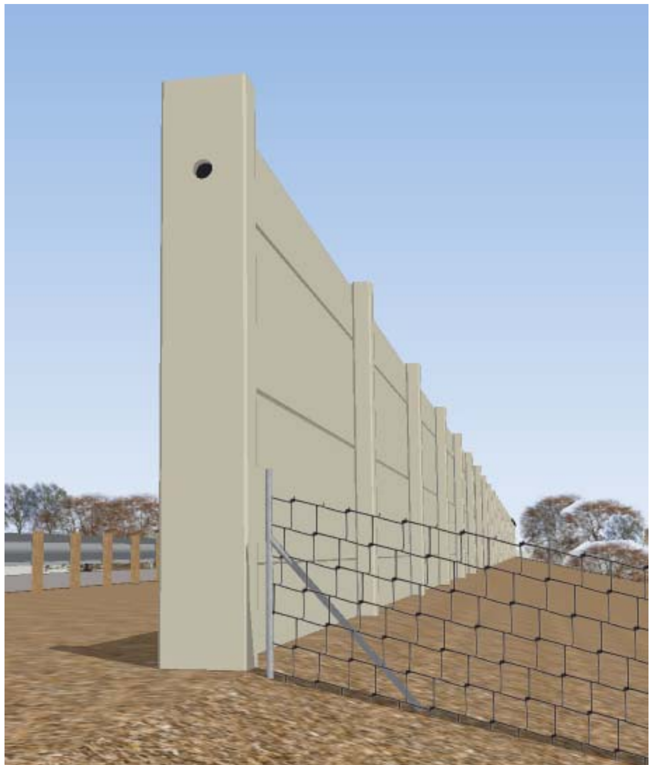

Figure 11: Image pair 13:38. (Top) SU Image 13. (Bottom) SU Photograph 38.

Figure 12: Image pair 22:47. (Top) SU Image 22. (Bottom) SU Photograph 47.

Image pair 12:37 in Figure 13 depicts a bridge with a road sign and contains barrier wall, fence, guardrail, bridge, sign, concrete, grass, trees, and sky. The rank difference of these images is 4.5, meaning that both the SU image and SU photograph received similar scores out of 25 possible ranks but had a difference of 18% overall. Obvious similarities are the basic architectural features of the bridge, a limited view of the sky, and the text on the road sign. Differences include shadow patterns, ground texture, the amount of visible road, grass type, jointing patterns on the bridge, the cleanliness of the barrier wall, and the angle from which the image was taken. These differences may have contributed to the difference in rank between the two images.

Image pair 5:30 in Figure 14 depicts a highway going under a bridge and contains guardrail, gravel, road, construction barrels and signs, road signs, trees, grass, and sky. The rank difference of these images is 4, meaning that both the SU image and SU photograph received similar scores out of 25 possible ranks but had a difference of 16.67% overall. Obvious similarities are the general angle from which the image or photograph was taken, the general placement of the construction barrels and signs, architecture of the bridge, lack of clouds, and the shadow patterns. Differences include the distance from which the image was taken, the color of the concrete, and the amount and soft-appearance of vegetation. These differences may have contributed to the difference in rank between the two images.

Figure 13: Image pair 12:37. (Top) SU Image 12. (Bottom) SU Photograph 37.

Figure 14: Image pair 5:30. (Top) SU Image 5. (Bottom) SU Photograph 30.

Image pair 24:49 in Figure 15 depicts a highway rest area building and contains grass, trash receptacles, ash urn, person, mulch, concrete sidewalk, trees, and sky. The rank difference of these images is 11, meaning that both the SU image and SU photograph received very different scores out of 25 possible ranks. Obvious similarities are the angle from which the image or photograph was taken, basic shadow pattern, absence of clouds, and the apparent materials of the building. Differences include the absence or present of a person, the absence or presence of display cases, the location of trash receptacles, color of mulch, color of windows, and amount and color of background trees. These differences may have contributed to the difference in rank between the two images.

Image pair 20:45 in Figure 16 depicts a grassy patch near a driveway and contains grass, fence, stone pillars, asphalt, catch basin, fire hydrant, buildings, sign, street light, trees, and sky. The rank difference of these images is 12.5, meaning that both the SU image and SU photograph received very different scores out of 25 possible ranks. Obvious similarities are the angle from which the image or photograph was taken, absence of defined clouds, and the texture of fencing materials. Differences include the presence or absence of buildings, presence or absence of a sign, texture of the driveway, cleanliness of the curb and driveway, presence or absence of the date image was taken, form of the street light, shadow pattern, background trees, and fullness, transparency, and species of foreground trees. These differences may have contributed to the difference in rank between the two images.

Figure 15: Image pair 24:49. (Top) SU Image 24. (Bottom) SU Photograph 49.

Figure 16: Image pair 20:45. (Top) SU Image 20. (Bottom) SU Photograph 45.

The most obvious difference between all image pairs is the lack of realism in the SU images. The photographs look unsurprisingly more realistic. This too may have had an effect on the degree to which the ranks between image pairs differ.

It can be observed that, as the difference between ranks grew, as did the amount of observable differences between the images. It therefore seems feasible that two photographs with the same differences might also receive similarly different rankings, as suggested by Shafer & Richards (1974) in their study. Although the SU images could have been changed to better match the photographs they each represent, they were instead created to a level of detail as would commonly be achieved for a basic design presentation.

In looking at the calculations for SIA vs. SIB, the results show there is a significant relationship between the two image sets to a p-value of <0.005. This concordance serves to further validate the above findings by reproducing the results of Burley's (1997) study nearly identically. This suggests not only that there is a concordance between SU images and SU photographs, but it also solidifies the validity of the method used to find these results.

Furthermore, it may be possible to assume that when presenting a design to the general public or an average client, many people do not understand how to read technical drawings or plan-view graphics. Here, the use of alternate display graphics may be necessary to ensure improved understanding. Again, having SU images of the proposed design may benefit the project by providing this added understanding necessary to progress decision-making.

Designers who choose to spend time mastering the technique of creating SU models may benefit further, learning skills such as how to construct a model that can be easily modified to accommodate design changes, how to customize textures and face-me components, or the speed at which the model is completed to name a few. However, the basics needed for presenting a conceptual design can be learned in a relatively short period of time.

Errors and future research

It is possible that errors may have occurred during this study. Potential errors may include unintentional bias when creating SU images that may incline a participant to favor either the SU image or the SU photo, inconsistent direction given to participants if said participant asked questions during the instructional or sorting period, regional bias based on the location of the testing area, and bias to participants who like cupcakes.

Unintentional bias may have occurred during the construction of SU images based on the researcher's technological capabilities or preferences. Every effort was taken to create SU images in the likeness of the corresponding SU photo to the fullest reasonable extent possible. As mentioned earlier, SU models were created to what the researcher believed to be a reasonable limit where a model of a proposed design would typically be created within the constraints of the SU photo. To clarify, no additional "artistic flare" was intentionally added and no additional detail was added where it did not already exist in the corresponding SU photo.

Initial directions were given consistently to each participant at the start of each session. However, several participants asked individualized questions or asked for the directions to be repeated. If it was deemed by the researcher that answering a particular question would affect the data, the question was not answered. However, it may be possible that by repeating directions or answering questions that seemed unrelated to the results, this may have skewed that particular data set. Future studies could avoid this by only providing written or prerecorded instructions.

There is unavoidable regional bias in this study based on the location of the data recording site. Data was recorded in Okemos, MI at the Meridian Farmer's Market. The intension of choosing this location was to get a variety of participants from a variety of backgrounds. While it seems this was achieved, it is possible that citizens who do not attend farmer's markets could be considered biased against. For a comprehensive study, future research should be conducted at a variety of venues at multiple locations around the globe.

Finally, cupcakes were given as an incentive for participants to supply their input for the study. It is possible that a potential participant who does not like cupcakes may not have participated due to the lack of proper incentive. Future studies could include a wider variety of incentives for participants.

Other future studies should include testing using other 3D modeling software such as 3D Studio Max, Revit, or other industry equivalents. Future studies could also include an additional level of rendering detail by applying advanced textures and lighting such as those found in SketchUp Podium, Kerkythea, or Maya. Still other studies may include the use of video animations or interactive 3D models vs. walking through an actual site.

CONCLUSION

The relationship between SU images and SU photos was found to be significant to a p-value of <0.025, suggesting that the average person, when looking at an image from an SU model will see the image for its content rather than for the quality of the SU model itself. Hypothesis 1 was accepted. These findings suggest that SU models are a valid and useful tool for designers when presenting designs to the general public or a typical private client. With this in mind, advantages of using an SU model could arguably outweigh the disadvantages.

Furthermore, the inclusion of images from Burley (1997) in this study and relation of the ranks back to the original study only serve to strengthen the validity of the methods used here. There was found to be a strong concordance ($p<0.005$) between the ranks achieved in this study (SIA) and those achieved in the original study (SIB) by Burley (1997). Hypothesis 2 was accepted.

APPENDICES

APPENDIX A

SECTION 1.1: SU IMAGES

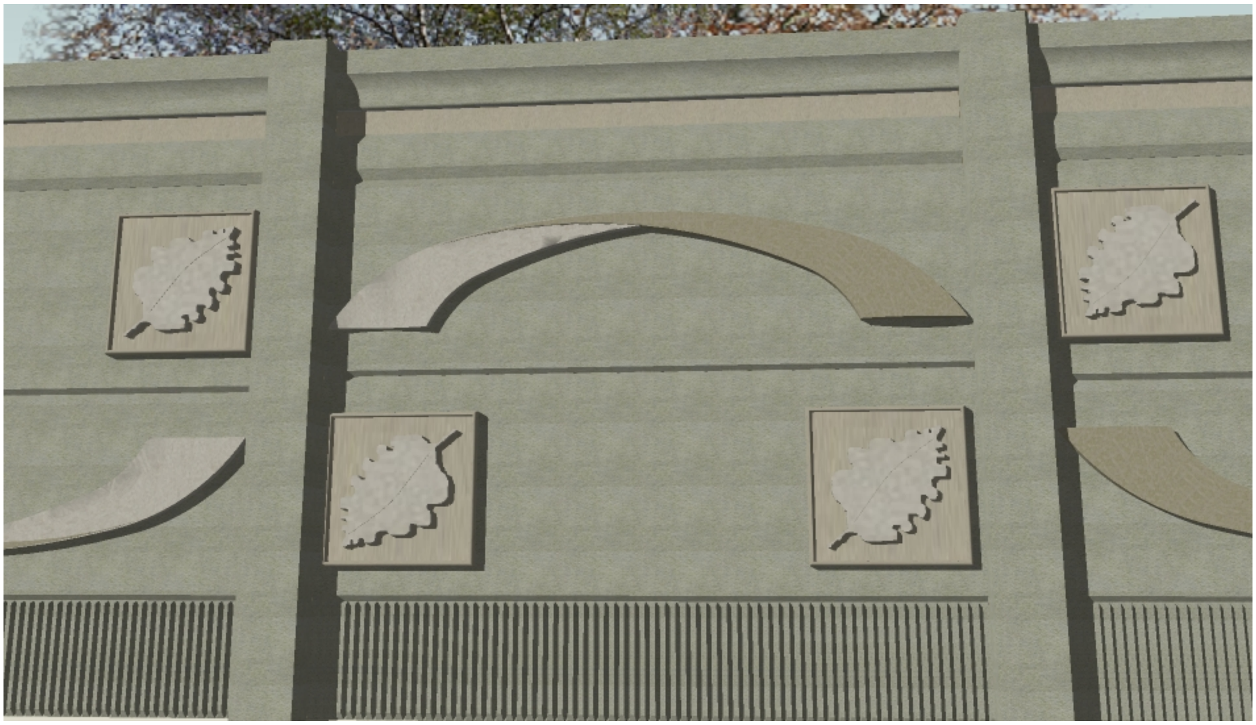

Figure 17: Image 1.

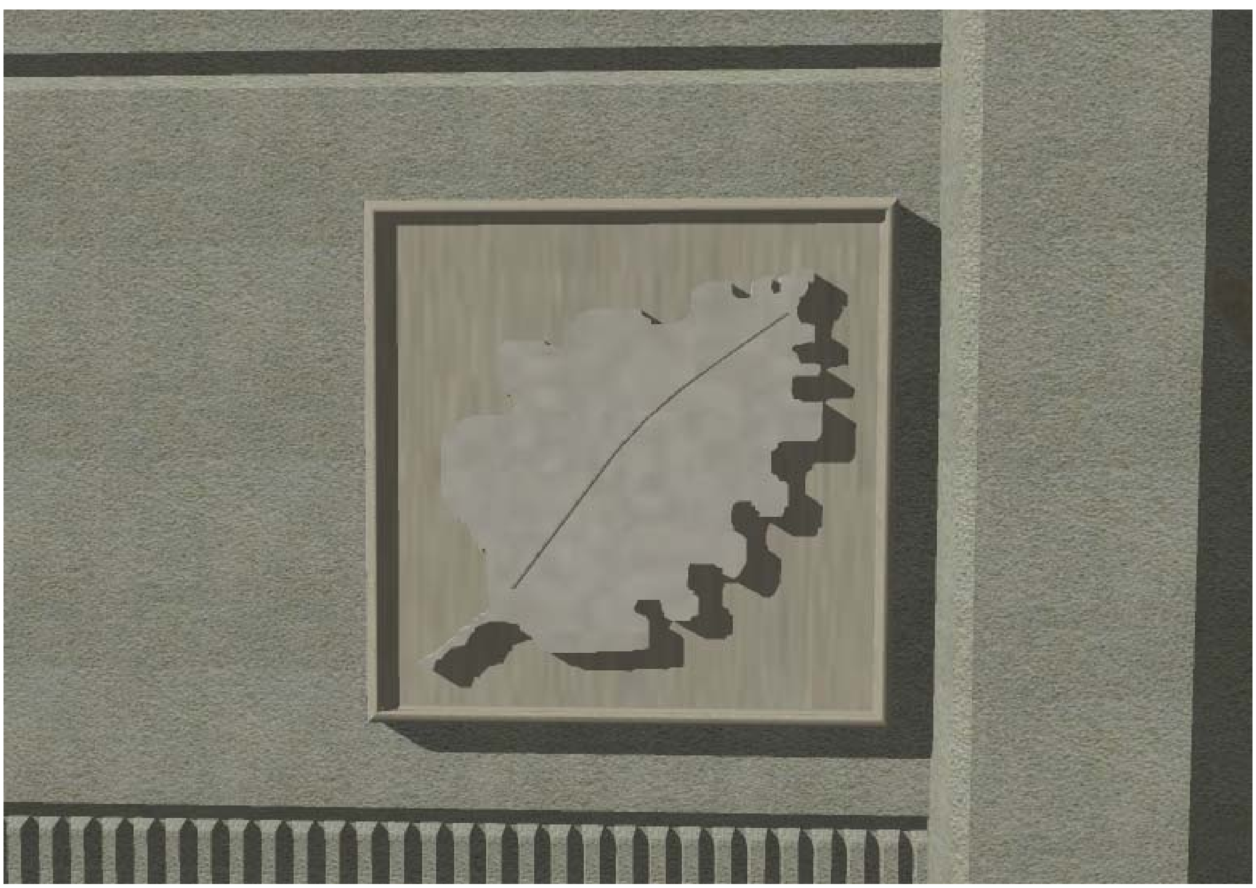

Figure 18: Image 2.

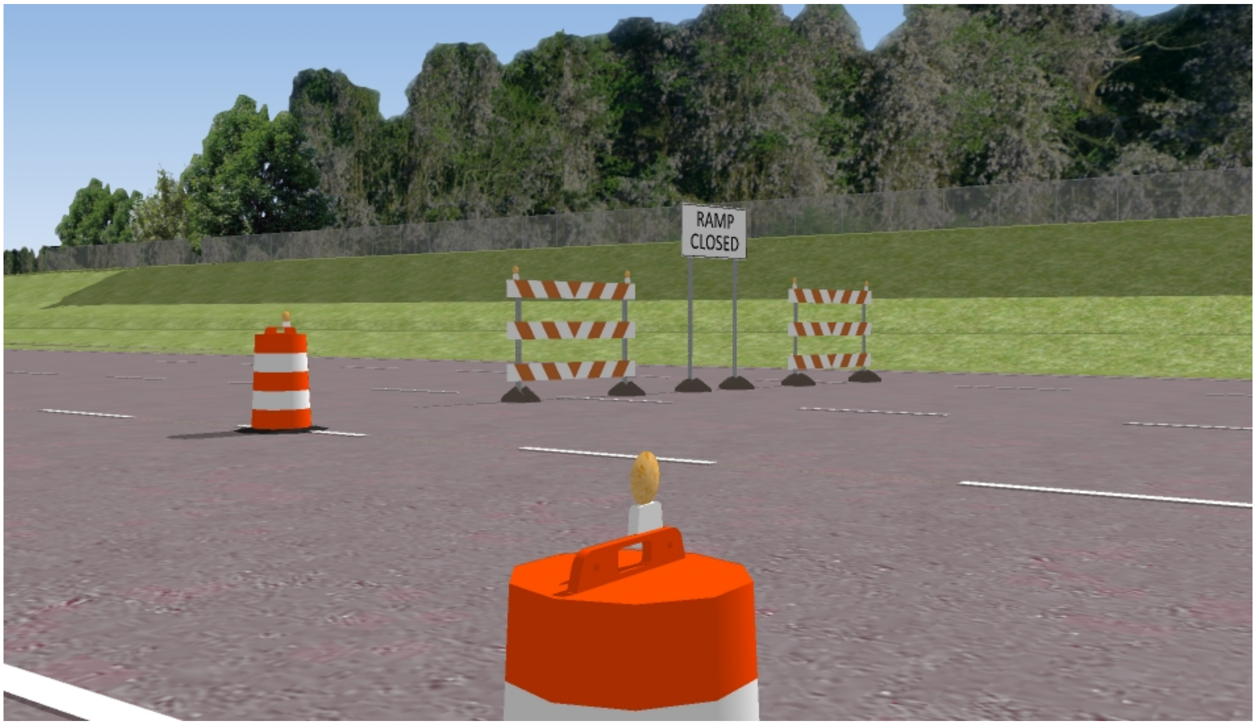

Figure 19: Image 3.


Figure 20: Image 4.

Figure 21: Image 5.

Figure 22: Image 6.

Figure 23: Image 7.

Figure 24: Image 8.

Figure 25: Image 9.

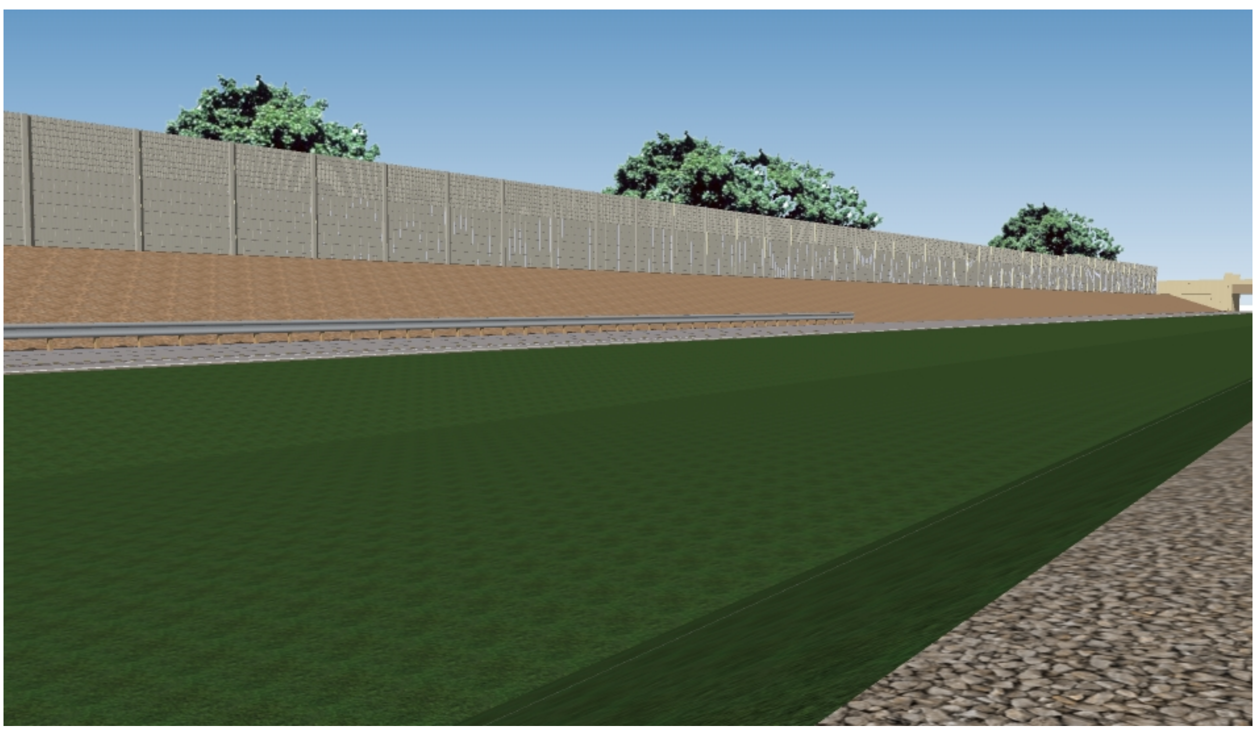

Figure 26: Image 10.

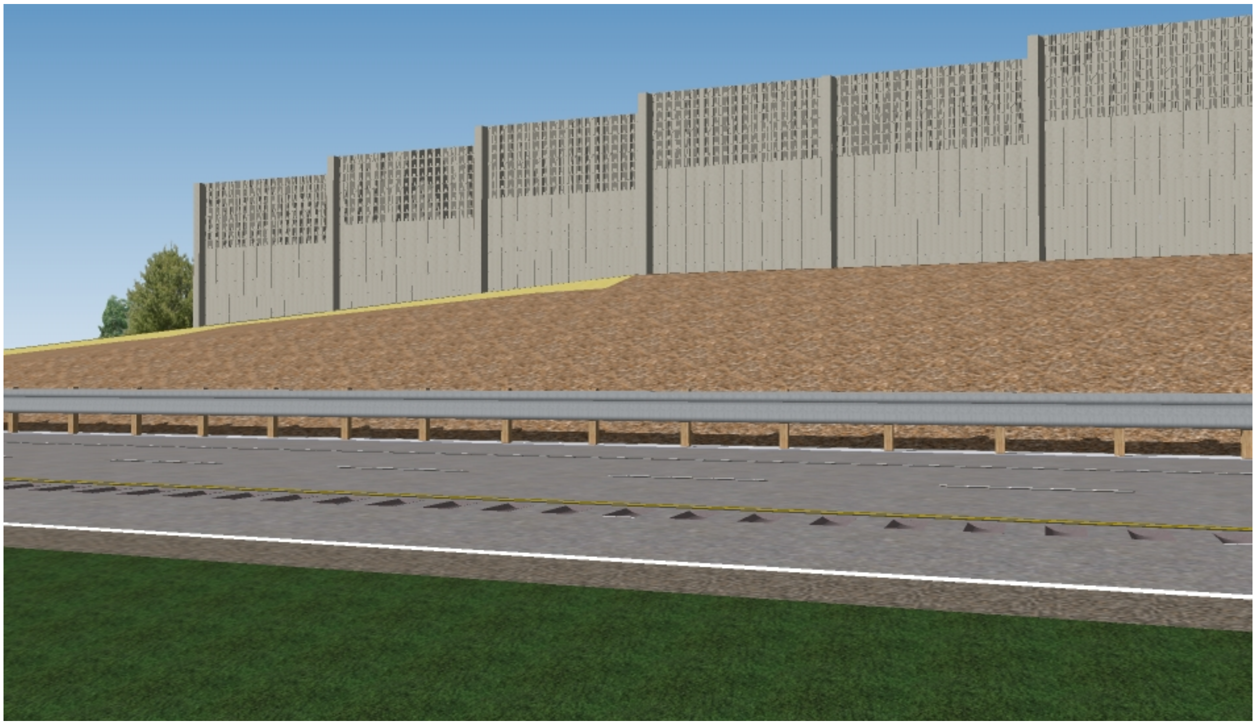

Figure 27: Image 11.

Figure 28: Image 12.

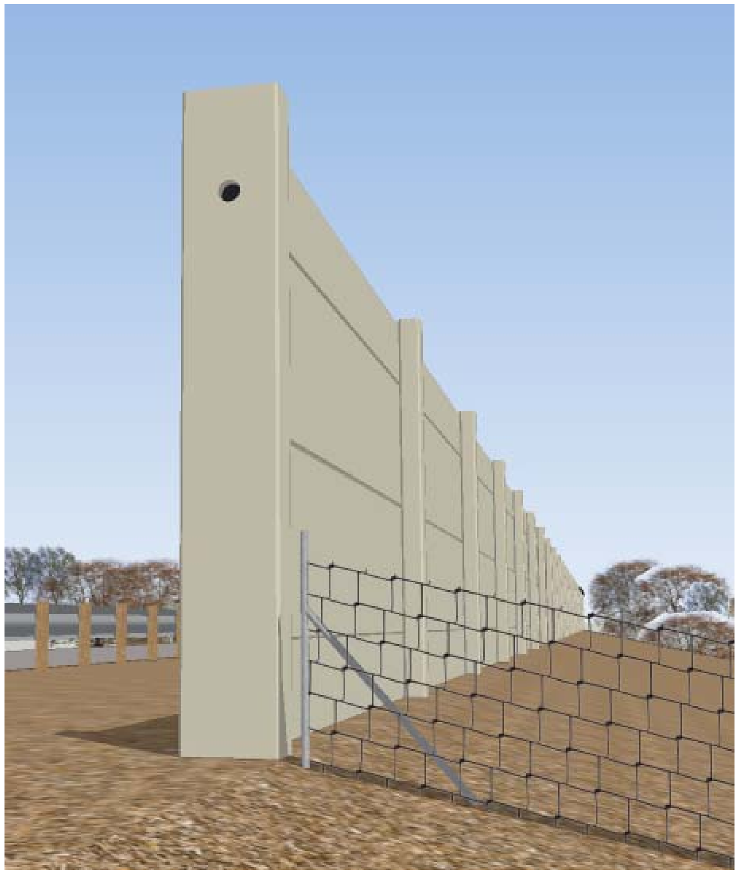
Figure 29: Image 13.


Figure 30: Image 14.

Figure 31: Image 15.

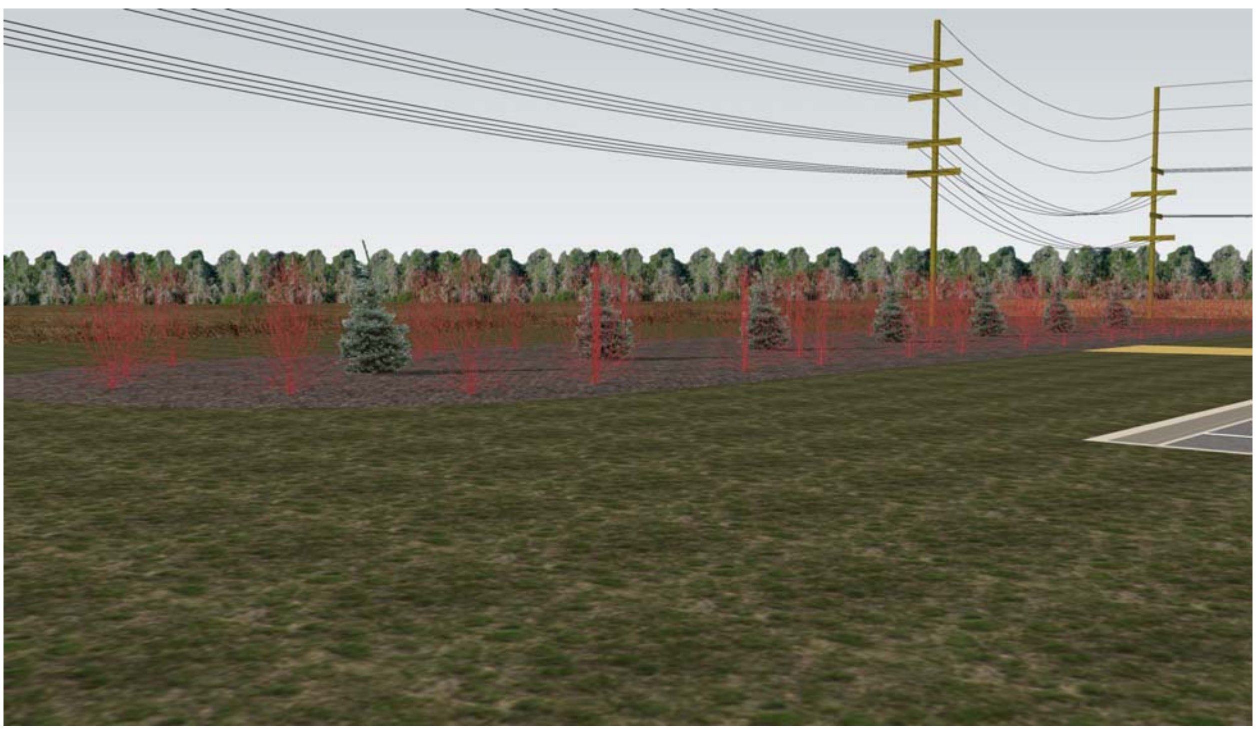

Figure 32: Image 16.

Figure 33: Image 17.

Figure 34: Image 18.

Figure 35: Image 19.

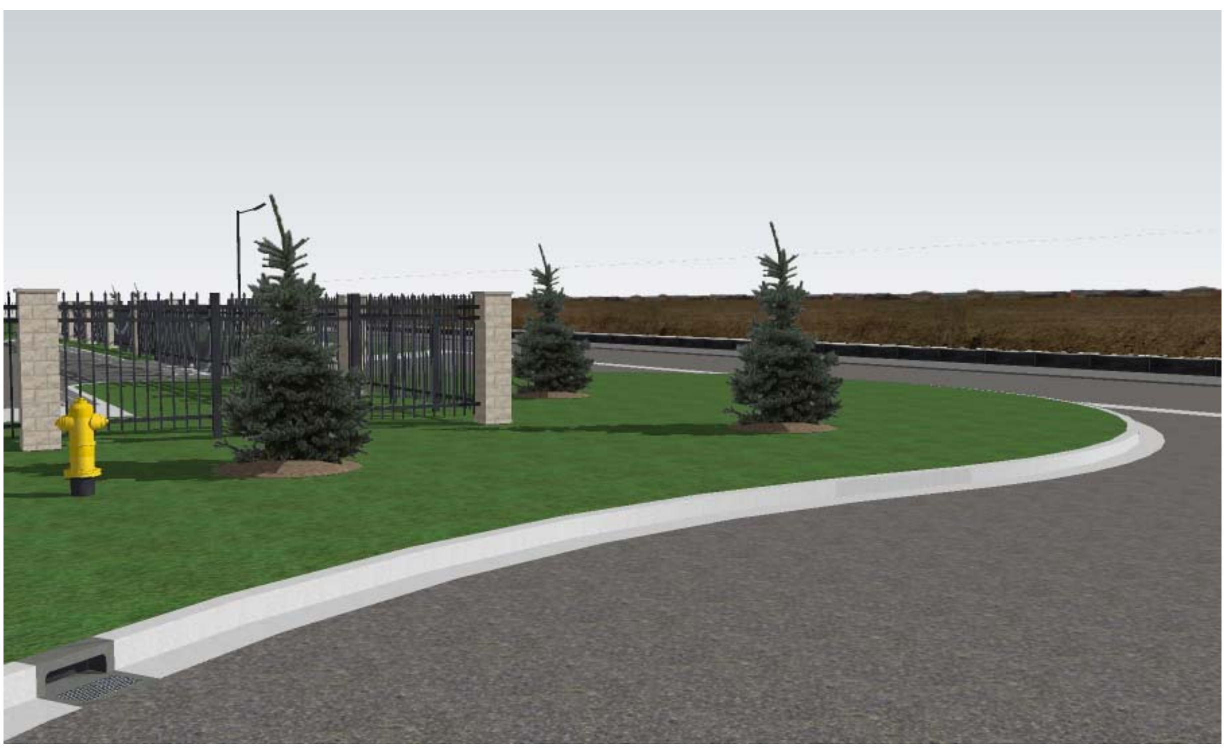

Figure 36: Image 20.

Figure 37: Image 21.

Figure 38: Image 22.

Figure 39: Image 23.

Figure 40: Image 24.

Figure 41: Image 25.

APPENDIX A

SECTION 1.2 SU PHOTOGRAPHS

Figure 42: Image 26.

Figure 43: Image 27.

Figure 44: Image 28.

Figure 45: Image 29.

Figure 46: Image 30.

Figure 47: Image 31.

Figure 48: Image 32.

Figure 49: Image 33.

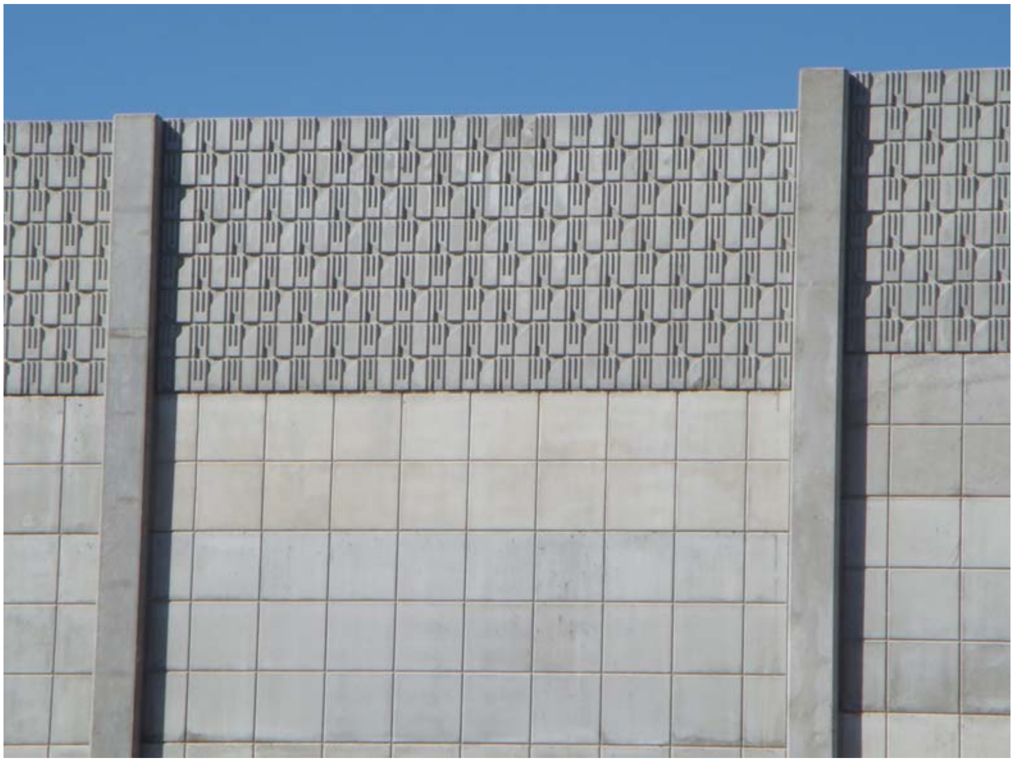

Figure 50: Image 34.

Figure 51: Image 35.

Figure 52: Image 36.

Figure 53: Image 37.

Figure 54: Image 38.

Figure 55: Image 39.

Figure 56: Image 40.

Figure 57: Image 41.

Figure 58: Image 42.

Figure 59: Image 43.

Figure 60: Image 44.

Figure 61: Image 45.

Figure 62: Image 46.

Figure 63: Image 47.

Figure 64: Image 48.

Figure 65: Image 49.

Figure 66: Image 50.

APPENDIX A

SECTION 1.3: SEEDED IMAGES

Figure 67: Image 51.

Figure 68: Image 52.


Figure 69: Image 53.


Figure 70: Image 54.


Figure 71: Image 55.


Figure 72: Image 56.

Figure 73: Image 57.

Figure 74: Image 58.

Figure 75: Image 59.

Figure 76: Image 60.

Figure 77: Image 61.

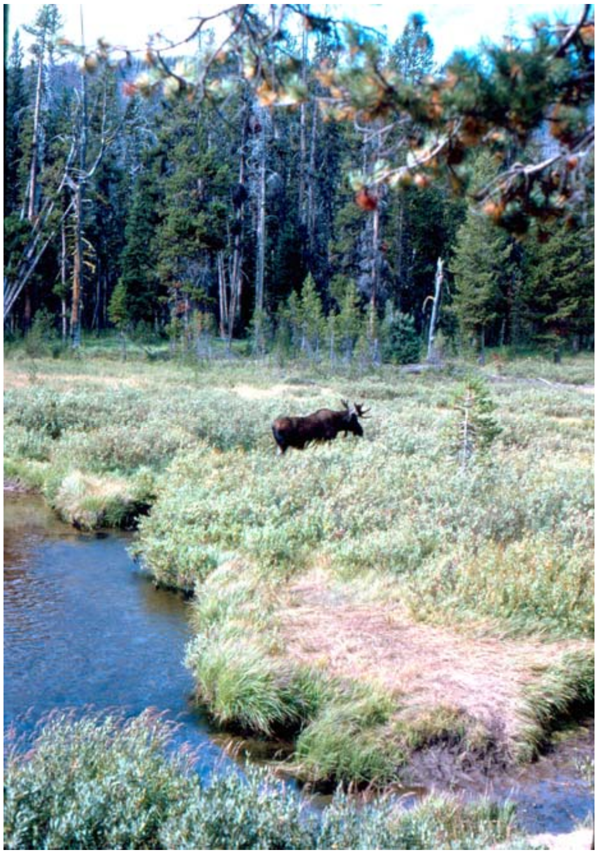

Figure 78: Image 62.

Figure 79: Image 63.

Figure 80: Image 64.

Figure 81: Image 65.

Figure 82: Image 66.

Figure 83: Image 67.

Figure 84: Image 68.


Figure 85: Image 69.

Figure 86: Image 70.

Figure 87: Image 71.

Figure 88: Image 72.

Figure 89: Image 73.

Figure 90: Image 74.


Figure 91: Image 75.

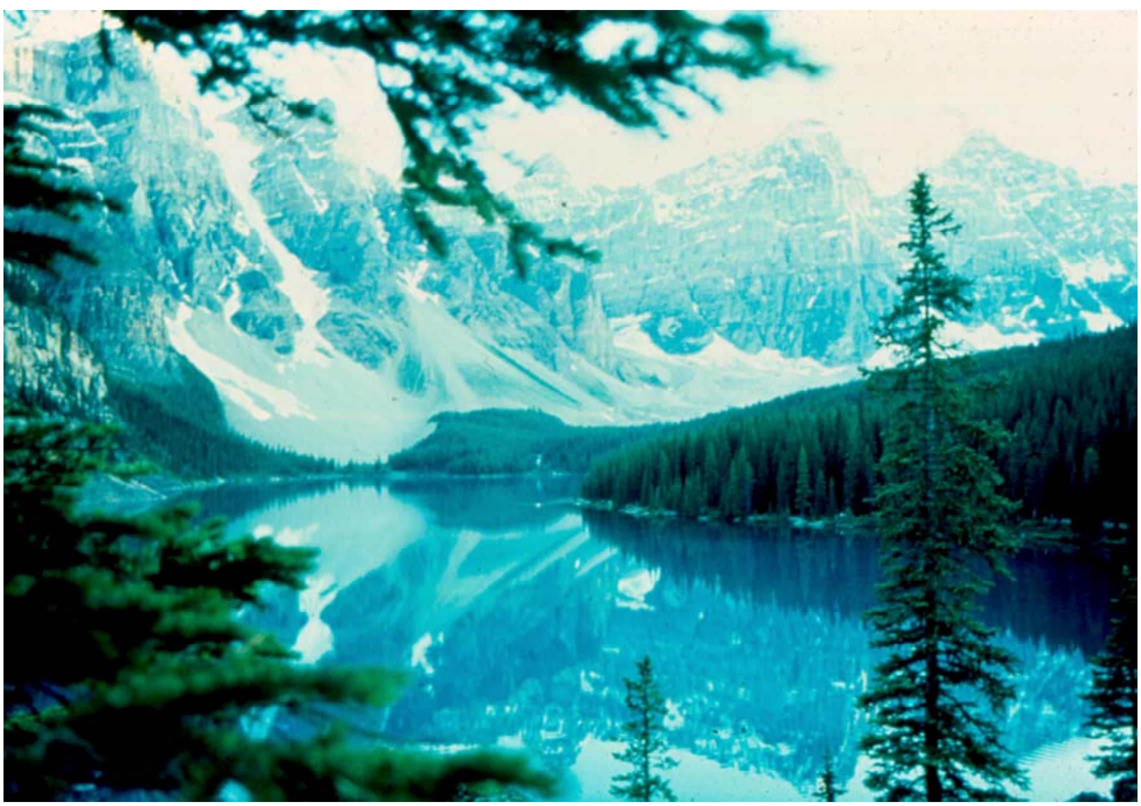
Figure 92: Image 76.

Figure 93: Image 77.

Figure 94: Image 78.

Figure 95: Image 79.

Figure 96: Image 80.

Figure 97: Image 81.

Figure 98: Image 82.

Figure 99: Image 83.

Figure 100: Image 84.

Figure 101: Image 85.

Figure 102: Image 86.

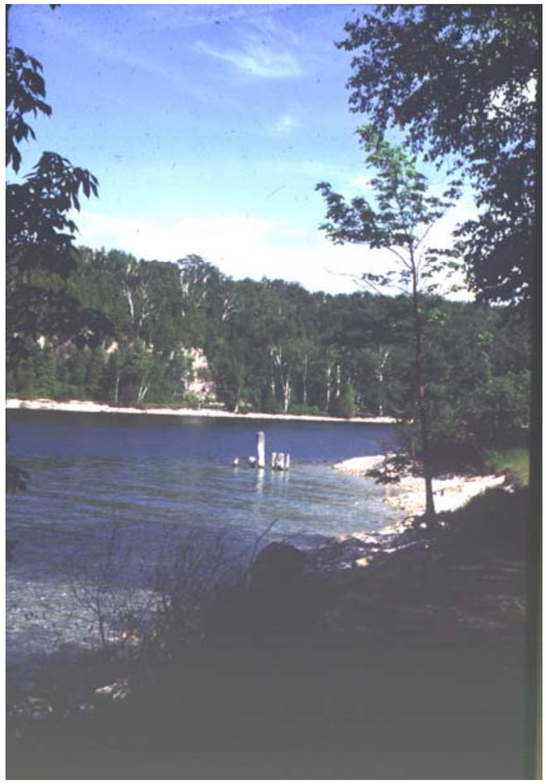

Figure 103: Image 87.

Figure 104: Image 88.

Figure 105: Image 89.

Figure 106: Image 90.

Figure 107: Image 91.

Figure 108: Image 92.

Figure 109: Image 93.

Figure 110: Image 94.

Figure 111: Image 95.

Figure 112: Image 96.

Figure 113: Image 97.

Figure 114: Image 98.

Figure 115: Image 99.

Figure 116: Image 100.

APPENDIX B

WAIVER FORM

Consent to Participate in Research

The activity in which you may choose to participate involves research concerning image preference. You will be asked to arrange five groups of ten images each in order from *least aesthetically pleasing,* on the left, to *most aesthetically pleasing,* on the right. Images will be presented in the form of full-page, color printouts. Fifteen minutes is the estimated time to sort each group of images. It may be possible to complete each group in less time. You have completed this activity when all ten images in each of the five groups have been arranged.

There are no known adverse side-effects or physical harm by participating in this study. Your participation in this research is voluntary. If you feel, for any reason, that you are unable to complete the activity, you may stop at any time without penalty.

It is requested, but not mandatory, that you answer the following question:

Are you a professional designer?   Yes: ________   No: ________

In order to participate in this study, you are required to print and sign your name below. *Your name will not be used and you will not be identified in this study.* By signing your name, you are stating that you have read and understand the above statements and that you are willing to participate in the described research activity.

Name (Please Print): ____

Signature: ____

If you have any comments, questions, or concerns, please feel free to contact any or all of the following:

Name/E-mail:	Mailing Address:
Shawn Partin Researcher Graduate Student E-mail: partinsh@msu.edu	Michigan State University 101 Human Ecology East Lansing, MI 48824
Dr. Jon Burley Professor E-mail: burleyj@msu.edu	Michigan State University 101 Human Ecology East Lansing, MI 48824
Harry McGee, MPH SIRB Chairperson Phone: 571.355.2180 E-mail: irb@msu.edu	Michigan State University 205A Olds Hall East Lansing, MI 48824

REFERENCES

REFERENCES

1. Akhtar-Danesh, N., et al. "Q-Methodology in Nursing Research: A Promising Method for the Study of Subjectivity." *Western Journal of Nursing Research* Vol. 30, No. 6. (2008): 759-773.

2. Bergen, S., C. Ulbricht, J. Fridley, and M. Ganter. "The Validity of Computer-Generated Graphic Images of Forest Landscape." *Journal of Environmental Psychology* 15. (1995): 135-146.

3. Bishop, I., B. Rohrmann. "Subjective responses to simulated and real environments: a comparison." *Landscape and Urban Planning* 65. (2003): 261-277.

4. Boscolo, P. and L. Cisotto. "Instructional strategies for teaching to write: a Q-sort analysis." *Learning and Instruction* 9. (1999): 209-221.

5. Brown, Stephen."A Primer on Q Methodology." *Operant Subjectivity* (April/July) 16(3/4). (1993): 91-138.

6. Buckley, M., et al. "Development and Validation of a Q-Sort Procedure to Assess Temperament and Behavior in Preschool-Age Children." *Journal of Clinical Child and Adolescent Psychology* Vol. 31, No. 4 (2002): 525-539.

7. Burley, Jon. "Visual and ecological environmental quality model for transportation planning and design." Transportation Research Record 1549. (1997): 54-60.

8. Daniel, Wayne. <u>Applied Nonparametric Statistics</u>. Houghton Mifflin, 1978.

9. Dearden, Philip. "Factors Influencing Landscape Preferences: an Empirical Investigation." *Landscape Planning* 11. (1984): 293-306.

10. Doyle, Michael. <u>Color Drawing</u>. New York: John Wiley & Sons, 1999.

11. Hull, R., W. Stewart. "Validity of Photo-Based Scenic Beauty Judgements." *Journal of Environmental Psychology* 12. (1992): 101-114.

12. Lin, Mike. <u>Drawing and Designing with Confidence</u>. New York: John Wiley & Sons, 1993.

13. Liu, C. "Mobile phone user types by Q methodology: an exploratory research." *Int. J. Mobile Communications* Vol. 6, No. 1. (2008).

14. Magritte, Rene. <u>Les trahison des images</u>. 1928-1929. Wikipedia: The Treachery of Images. October 12, 2011. November 13, 2011. <http://en.wikipedia.org/wiki/File:MagrittePipe.jpg>.

15. McKnight, M. "Applying the Q Sort Method: A Qualitative Classifiaction of Factors Associated with the Organization Training Support Inventory (OTSI)." *Online Journal of Workforce Education and Development* Vol. III, Issue 2. (2008).

16. Oh, Kyushik. "A perceptual evaluation of computer-based landscape simulations." *Landscape and Urban Planning* 28. (1994): 201-216.

17. Pitt, D., and E. Sube. "The Q-Sort Method: Use in Landcape Assessment Research and Landscape Planning." National Conference on Applied Techniques for Analysis and Management of the Visual Resource. Scientific Article No. A2592. (1979).

18. Previte, J., et al. "Q Methodology and Rural Research." *European Society for Rural Sociology* Vol. 47, No. 2. (2007).

19. Reid, Grant. Landscape Graphics. New York: Watson-Guptill Publications, 2002.

20. Rohrmann, B., and I. Bishop. "Subjective Responses to Computer Simulations of Urban Environments." *Journal of Environmental Psychology* 22. (2002): 319-331.

21. Roth, Michael. "Validating the use of Internet survey techniques in visual landscape assessment—An empirical study from German." *Landscape and Urban Planning* 78. (2006): 179-192.

22. Shafer, Jr., Elwood, and Thomas Richards. "A Comparison of View Reactions to Outdoor Scenes and Photographs of Those Scenes." USDA Forest Service Research Paper NE-302. (1974).

23. Smardon, Richard, James Palmer, and John Felleman. Foundations for Visual Project Analysis. New York: John Wiley & Sons, 1986.

24. Stewart, Thomas, Paulette Middleton, Mary Downton, and Daniel Ely. "Judgments of Photographs vs. Field Observations in Studies of Perception and Judgment of the Visual Environment." *Journal of Environmental Psychology* 4. (1984): 283-302.

25. Swaffield, S., and J. Fairweather. " Investigation of attitudes towards the effects of land use change using image editing and Q sort method." *Landscape and Urban Planning* 35. (1996): 213-230.

26. Westin, D. et al. "Affect Regulation and Affect Experience: Individual Differences, Group Differences, and Measurement Using a Q-Sort Procedure." *Journal of Consulting and Clinical Psychology* Vol. 65, No. 3. (1997): 429-439.

27. Zube, Ervin, David Simcox, and Charles Law. "Perceptual Landscape Simulation: History and Prospect." *Landscape Journal* Vol. 6. (1987): 62-80.