CONCEPTUALIZATION, MEASUREMENT, AND NOMOLOGICAL VALIDITY OF VISUAL ORGANIZATION GRAMMAR FOR DESIGN: AN ANALYSIS OF WEBSITE DESIGN

Research-in-Progress

Julian Lin
Communications and New Media Department
National University of Singapore
11 Computing Drive, Singapore
lin@nus.edu.sg

Abstract

The aim of this paper was to provide an operational definition, to develop and to validate a measurement tool for assessing visual aesthetics—specifically contrast, rhythm, balance, and proportion (CRBP)—proposed by Stebbing (2004). A 12-item scale was developed to measure CRBP. In a pilot study with over 300 responses, exploratory and confirmatory factor analyses were performed on the items. This research-in-progress paper showed that the final 12 items had adequate convergent and discriminant validity. Implications for practice and research derived from the paper are also discussed. Future study is proposed to look at the relationship between the CRBP and other related constructs.

Keywords: Visual Organization, contrast, balance, rhythm, proportion
Introduction

Art is the imposing of a pattern on experience and our aesthetic enjoyment is recognition of the pattern

(Whitehead and Price 1954, p. 229)

The aforementioned quotation epitomizes the purpose of this study. Even the most unstructured, unrestricted and undefined subject—art and our aesthetic enjoyment—can be analyzed through etic dimensions. These dimensions are described by Alfred North Whitehead (1954), as patterns, which we can and should recognize. Art, in this study, is defined in its broad scope; it is defined as qualities, productions, expressions, or realms—according to aesthetic principles—of what is beautiful, appealing, and of more than the ordinary significance. We believe that no object, which is beautiful and appealing, is mysterious. The mystery is often in the interpretation of one's eye. Although many artists, designers, and researchers' laborious attempt to come up with aesthetics dimensions that are realistic, universal, and able to encompass all nuances in design (e.g., Cyr et al. 2006; Cyr et al. 2010; Hsiao and Chou 2006; Lavie and Tractinsky 2004; Lindgaard et al. 2011; Moshagen and Thielsch 2010; Park et al. 2005; Schenkm an and Jonsson 2000; Thuring and Mahike 2007; Tractinsky et al. 2006; van Schaik and Ling 2009), none has yet offered a systematic view of aesthetic principles based on patterns.

The late Paul Cezanne, a French artist and Post-Impressionist painter, once said: "Light is a thing that cannot be reproduced, but must be represented by something else, by color" (Sérullaz 1960, p. 119). Likewise, all areas of design are but representations and manifestations of a designer's idea. The represented is then by logic directly inaccessible to us. Nonetheless, with the use of etic dimensions, we could ascribe the representatives into various categories, break them up, and segregate them into comprehensible patterns. By understanding these patterns, IS researchers could inform practitioners to design visually appealing websites that are important for establishing web presence for their business. Especially, previous studies have shown that users do not trust badly designed websites (Lindgaard et al. 2011; Sillence et al. 2004a; Sillence et al. 2004b). For instance, 94 percent of the factors that lead users to quickly reject a website are design while only about 6 percent are related to content (Sillence et al. 2004b). Furthermore, it only takes users 30 seconds to leave an ugly webpage (Lindgaard et al. 2011; Sillence et al. 2004a).

Many researchers have investigated constructs and dimensions of aesthetic principles and their relationships with attitude, intention, and/or use (e.g., Cyr et al. 2010; Hartmann et al. 2008; Lee and Kou bek 2010; van der Heijden 2003; van Schaik and Ling 2011; Verhagen et al. 2012). Although these studies advanced our understanding about aesthetics and use, they only examined the overall constructs as well as the overall dimensions of aesthetic principles. For instance, most of these studies looked at a single construct such as perceived aesthetics (Cyr et al. 2010; van der Heijden 2003), expressive and classical aesthetics (Hartmann et al. 2008; Lavie and Tractinsky 2004), or visual attractiveness (Verhagen et al. 2012). Multiple dimensional constructs of aesthetic principles with established theoretical concepts are often not considered.

The aim of this paper is to consolidate and develop a scale for multiple dimensional constructs of aesthetics—a cohesive practical set of patterns—that Stebbing (2004) called as the visual organization grammar for design. The scope of this paper would cover only patterns of visual composition; being the arrangement of visual elements on a medium. Other patterns of designs—such as, but not restricted to, color and style—will not be discussed in this paper.

Literature Review

Stebbing’s Contrast, Rhythm, Balance and Proportion (CRBP)

Stebbing (2004) developed the following principles as an attempt to identify the core patterns of a visual organization grammar for the design curriculum (Stebbing 2004). He isolated four specific patterns through content analysis of words from the content pages of 50 books on visual composition from various art and design subjects.
The four patterns are stipulated as follows:

1. **Contrast** – refers to the arrangement of two or more distinct design patterns in identifiable opposition to other patterns – by means of exhibiting unlikeness through comparison or juxtaposition of different form or lines. Contrast is not to be confused with graduation (lacks distinction) or with variation (lacks unlikeness).

2. **Rhythm** – refers to the arrangement of two or more design patterns in a recurring manner in one or a combination of the four basic symmetry operations – shift, reflection, mirror and rotation – of their own patterns to form a consistent or characteristic arrangement of qualities.

3. **Balance** – refers to the arrangement of two or more similar or dissimilar design patterns so that it equalizes or neutralizes the tensions, and results in a sense of equilibrium. Related concepts include, but not restricted to equilibrium or symmetry.

4. **Proportion** – refers to a ratio composed of two or more contrasting quantities used repeatedly in either the same and/or different measures in a design. Related concepts or synonyms include, but are not restricted to, ration, section or variety.

Using an analytical investigation procedure, the results supported his hypotheses that contrast, rhythm, balance and proportion (CRBP) were the most commonly used design patterns for visual composition (Stebbing 2004). Specifically, Stebbing (2004) highlighted that CRBP was not only about consistency among multiple cultures’ design vocabulary, but they also affected our aesthetic behavior in creating and appreciating visual compositions that has evolved from our innate ability to recognize the diversity of organic forms through these basic organizing principles (i.e., CRBP). In addition, our aesthetic behavior may be an evolutionary “spandrel”, an aimless evolutionary by-product of another ability, which evolved to recognize organic form (Stebbing 2004).

**Related Studies on CRBP Design Patterns**

Previous studies related to contrast, rhythm, balance and proportion design patterns usually put much focus on the development of computation models (i.e., algorithm) to measure different page layout. Specifically, Ngo, Teo and Byrne (2000b) proposed algorithms for measuring 14 screen layout which include, rhythm, balance and proportion design patterns. They determined aesthetic values, which were measured in pixels, by calculating the measures of five websites with their proposed algorithms. They also compared these values with the subjective ratings of almost 200 undergraduate students. Their results showed a relatively high relationship between the two studies. In the same vein, Ngo (2001) replicated their previous study by constructing a model (i.e., the regression equation) to predict the overall measure of design patterns with 57 screens that were rated on a 0 to 1 scale by 7 GUI designers. He also used the regression equation to predict the subjective ratings with new screens and subjects. The results showed that the two measures were related to the viewers’ perception of design patterns. Along the same line, Ngo and Byrne (2001) examined aesthetic values for data entry screens instead of websites. The results were highly consistent with their previous results.

Park et al.(2005) examined several design patterns using algorithms proposed by Ngo and Byrne (2001) with more than 50 webpages. The result of regression analysis revealed that the bold personality was associated with simplicity and cohesion; analytical personality with contrast, density and simplicity; friendly personality with contrast, cohesion, density and regularity; and sophisticated personality with regularity and balance. The following subsection elaborates on these four design patterns.

**Contrast**

Contrast can be created using different design elements including colors, sizes, shapes, locations or relationships. For text, contrast can be achieved by using different type fonts such as serif for text body (e.g., Georgia) and san-serif for heading (e.g., Arial). Previous studies have measured contrast by examining the numeric value that is determined by the degree of difference among objects in size, color and location (Lindgaard et al. 2011; Park et al. 2005; Wong 1987). Specifically, very low contrast means objects and text are hardly distinguishable; medium contrast refers to although objects and text can be distinguished but there is no start relations exist; and high contrast means that objects and text can be distinguished clearly.
Rhythm

Successful designs with rhythm have an effective ebb and flow of composition through a pattern of identical or similar objects. For instance, text and graphics should seem to be patterned and have repetition elements (Garrett 1975; Lauer and Pentak 2011; Wallschlaeger et al. 1992; Wong 1972; Wong 1987). To measure rhythm, previous studies have used the numeric value to determine the repetition of, or regularity in, the sizes, colors, or locations of objects (Ngo 2001; Ngo and Byrne 2001; Ngo et al. 2000a; Ngo et al. 2000b; Park et al. 2005).

Balance

Balance has been equated with harmony and good color combinations (i.e., harmony is well-balance, and balance between manipulated colors) (Cyr et al. 2010). It has been defined as the measurement of the center layout with an equal weighting of components on each side of the horizontal and vertical axis (Ngo et al. 2000a). It also refers to the distribution of visual weight across the whole picture or the distribution of optical weight in a picture (Ngo and Byrne 2001; Park et al. 2005); and it is achieved by providing an equal weight of screen elements, left and right, top and bottom (Lindgaard et al. 2011; Ngo and Byrne 2001; Park et al. 2005). To measure rhythm, previous studies have used the numeric value to determine when balance is maximized by evaluating the equality of visual weights on the right and left sides or in the upper and lower regions on a screen (Ngo 2001; Ngo and Byrne 2001; Ngo et al. 2000a; Ngo et al. 2000b; Park et al. 2005).

Proportion

Proportion is the relative visual size and weight of particular graphical elements, the ratio of an object's height to its width (Park et al. 2005), and it deals with the size relationship of one part to another (Ngo and Byrne 2001). To measure proportion, previous studies have used the numeric value to determine the variation of design patterns by examining the differences among colors, fonts, shape and sizes (Ngo 2001; Ngo and Byrne 2001; Ngo et al. 2000a; Ngo et al. 2000b; Park et al. 2005).

Method

Instrument Development Process

While most IS studies conceptualize constructs as reflective, there has been an increase in conceptualizing them as formative (Diamantopoulos 2011; Marakas et al. 2007). In a formative construct, the direction of causality is from items to constructs; in contrast, in a reflective construct, the direction of causality is from construct to items. The same construct could also be measured formatively or reflectively. For instance, Diamantopoulos (2011) explained that intoxication could be reflectively measured by assessing the symptoms of feeling drunk or be formatively measured by examining a person's intake of different alcoholic drinks and other substances. As a reflective construct, intoxication is measured by self-perceived intoxication while the actual (objective) intoxication is an indicator of the formative construct. Similarly, CRBP could be conceptualized as either reflective or formative constructs. When we ask users to evaluate their perception of web design, CRBP are conceptualized as reflective. On the other hand, when we ask developers what constitute CRBP, they are conceptualized as formative. Because the purpose of this study is to evaluate users' perception of CRBP, in this paper, they are being conceptualized as reflective.

Following the three-step instrument development process for reflective constructs proposed by Moore and Benbasat (1991), the development of instrument that includes item creation, scale development and instrument testing was carried out. In the item-creation stage, we reviewed about two dozen of books that described the design patterns of contrast, rhythm, balance and proportion (CRBP). From these books—written by artists, architects, photographers, graphic designers and web designers—we created pools of items that identified the dimension of CRBP. The review of these books provided us with about 50 items (available upon request). With the help of two professional designers, these items were evaluated, and only 36 items were retained.

In the scale-development stage, card sorting was performed by ten judges who categorized items into their own created dimensions (i.e., each judge was asked to provide the names of the dimensions) and into the
provided dimensions (i.e., labeled sorting by providing judges with the names of the dimensions). Their level of agreement (multi-rater reliability using Cohen’s Kappa) was over .78 for both unlabelled and labeled sorting, and these numbers were higher than accepted threshold of .70 (Miles and Huberman 1994). After conducting the two rounds of sorting, the remaining items were 21 items that include five items for contrast; six items for rhythm; five items for balance and five items for proportion (see Table 1).

In the instrument-testing stage, we analyzed exploratory factor analysis (EFA) using SPSS version 14.0 and confirmatory factor analysis (CFA) using LISREL version 8.3 to ensure the unidimensionality of items for the visual organization grammar for design. To obtain the data for these analyses, we created an online survey, and asked 69 undergraduate students to rate six webpages. A total of 312 data points from 52 undergraduate students (31 percent was male and 69 percent was female) were collected within two weeks. These data were divided into half randomly using a split-half validation technique. We used the first set of the data to perform EFA as a procedure of measure purification (Bagozzi 1994, p. 331) while the second set of the data was used for CFA to complement the EFA approach by assessing the unidimensionality, reliability, convergent and discriminant validity of the constructs.

| Table 1. Items within the Constructs of Contrast, Rhythm, Balance and Proportion |
|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| **Contrast**                    |                                 | **Rhythm**                      |                                 | **Balance**                    |                                 |
| 1. Some parts of the layout are denser than others. | 6. The same pattern appears at many places. | 7. All elements on the page are obviously related. | 8. Similar shapes, forms and lines are used repetitively. | 12. The design achieves a sense of equilibrium (i.e. balance). | 13. The proportion of the information distributes equally. |
| 2. The dark part stands out against the bright part, or the bright part stands out against the dark part. | 9. Visual details give a consistent structure to the layout.* | 10. Similar visual cues are used throughout the layout.* | 11. Elements of the layout are tied together by visual cues.* | 14. The information at the top and at the bottom is arranged evenly. | 15. There is some kind of equal visual weight of information. |
| 3. Opposite colors appear close to each other. |                                 |                                 |                                 | 16. The amount of information in the left and right are about the same. |                                 |
| 4. Adjacent elements are clearly distinguishable.* |                                 |                                 |                                 |                                 |                                 |
| 5. Elements on the page are clearly differentiated. * |                                 |                                 |                                 |                                 |                                 |
| **Proportion**                  |                                 |                                 |                                 |                                 |                                 |
| 17. There are texts of different colors, fonts and sizes.* |                                 |                                 |                                 |                                 |                                 |
| 18. The design includes a variety of dissimilar objects. |                                 |                                 |                                 |                                 |                                 |
| 19. Various shades, hues and moods of the same color are applied in the design.* |                                 |                                 |                                 |                                 |                                 |
| 20. There are objects of different colors, values and shapes. |                                 |                                 |                                 |                                 |                                 |
| 21. The elements in the layout are not monotonous. |                                 |                                 |                                 |                                 |                                 |

* Deleted items during the instrument testing state

Before the EFA was conducted, we looked at the reliability of constructs. The results showed that all constructs have adequate reliability with alphas greater than the required number of .70 (Cronbach 1951) . The first random split data set created by SPSS yielded 150 responses. An exploratory principal factor analysis was conducted using the first set of the data. After deleting nine items shown in Table 1, the results showed the rotated factor analysis generated four distinct and clean components for the visual organization grammar for design.

The second split set of responses (n=162) was used to conduct CFA. The convergent validity was evaluated by examining whether the composite reliability of the construct is larger than the required value of .70. The discriminant validity was assessed by comparing the $\chi^2$ of the CFA with four latent variables against other CFAs with three different latent variables. In addition, seven fit indices, which consists of five model-fit indices and two error term magnitude estimates, were used to examine the adequacy of the
The results showed that all factor loadings of each item to each construct were found to be statistically acceptable. Specifically, the factor loading values ranged from 0.69 to 0.95; and t-values was significant ranging from 5.27 to 17.29). The ratio of chi-square value to degrees of freedom was 1.60 (i.e., df=48 and chi-square value was 77); the Goodness of Fit Index (GFI) =.93; the comparative fit index (CFI) = .97; incremental fit index (IFI) = .97; the normed fit index (NFI) = .92; the Tucker-Lewis Index (TLI) = .95; the root mean square of approximation (RMSEA) = .06; and the standardized root mean square residual (SRMS) = .05.

Field Study (In Progress)

In the field study, subjects will be asked to view and rate a few websites. CRBP scales that were developed earlier and existing aesthetic measures will administered to the subjects. It is possible that CRBP may be accounted for a common higher order aesthetics factors. The collected data will be used to assess the assumption. If possible, other theoretical models will also be explored.

Data Analysis and Results

Analytical Procedures (Our Plan)

After we collect the data for the field study, the data analysis that we will conduct aims to achieve two main objectives: the first objective was to examine the reliability, convergent validity, and discriminant validity of the development of scales for visual organization grammar for design that has four dimensions: contrast, rhythm, balance and proportion. We will evaluate the convergent validity of the items used to measure the dimensions of the visual organization grammar for design dimensions by examining four aspects. First, we will look at whether the hypothesized 4-factor structure adequately accounted for the data. Second, we will examine whether each item has a statistically significant loading of substantial size on the hypothesized factor. Third, we will analyze whether item cross-loadings on non-hypothesized factors are substantial, and lastly, whether each latent construct is able to account for a large proportion of the variance in its measured indicators.

We will also examine the reliability of each dimension of the visual organization grammar for design by calculating the internal consistency reliability, composite reliability, for each subscale. The discriminant validity of the scales will be assessed by comparing the \( \chi^2 \) of the CFA with four latent variables against other CFAs with three different latent variables. In addition, the average variance extracted (AVE) values should be larger than the required value of .50 (Fornell and Larcker 1987). Essentially, Fornell and Larcker (1987) argued that if two constructs are distinct, the average variance in a construct's indicators accounted for by the hypothesized construct should be greater than the amount of variance that the construct shares with any other construct. The second objective is to look at the nomological validity of the final scales by empirically examining whether the scale scores correlated with other variables in their nomological net in the manner expected.

Discussion

The results of exploratory and confirmatory factor analyses in the pilot study (i.e., instrument development process) revealed that the scales provide adequate evidence of the convergent and discriminant validity of the 12 items. All of the items loaded on their constructs as well as the latent factors accounted for a substantial proportion of the variance in their indicators. The results from the
instrument development process show that the scale is promising because these initial data reveals that it possess an adequate level of convergent and discriminant validity. Future research should continue refining the items and evaluate the nomological and predictive validity of factors in the visual organization grammar for design.

This study contributes to theory and practice in several ways. First, the visual organization grammar for design developed in our study evaluates visual aesthetics in a systematic manner. To our knowledge, none study has offered a systematic view of visual aesthetics. Second, the psychometric properties and our initial data indicate that the visual organization grammar for design scale can be useful for both academics and practitioners. Specifically, the scale could be a valuable diagnostic tool for assessing visual aesthetics in websites. Third, the measures could be applied to examine whether a systematic view of visual aesthetics can influence users’ attitude, intention and use of websites. Fourth, the scales can be also used to serve as a method to evaluate designs in different platforms such as flyers, brochures, or motion pictures. Last but not least, the scale development described in this study is derived from solid theoretical concepts after reviewing more 70 books written by practitioners. Specifically, Stebbing (2004) reviews 50 books and comes out with CRBP concepts. In addition, this study reviews more than 20 books and creates items based on these books. In summary, we believe the study of exogenous and endogenous factors of the visual organization grammar for design is a promising direction for research on aesthetics.

Acknowledgements

I gratefully acknowledge the support of NUS, Faculty of Arts and Social Sciences Start-up Grant R-124-000-043-133. I thank Hanson and Hoang Lian for research assistance, and to the reviewers and editors for helpful comments.

References