Environment and Planning B: Planning and Design 1994, volume 21, pages 223–246

Operational definitions of style

C-S Chan
Department of Architecture, Iowa State University, 482 College of Design, Ames, IA 50011, USA
Received 3 November 1992; in revised form 27 May 1993

Abstract. It is intended that this research will set up operational definitions of style and will study the fundamental phenomenon of style. Four experiments were conducted. In experiment 1, college students were asked to sort out pictures of buildings having the same style to test operational definition. The results showed that style is recognized by the common features present in pictures. The term 'common features' refers to the same physical forms appearing in many design products created by a designer. In experiment 2, students were asked to sort out pictures into four resemblance scales to test the degree of style. The results showed that the degree of style is in proportion to the number of common features present. In experiment 3, an architectural historian identified a style in pictures that had various feature combinations in order to observe the measurement of style, which refers to the threshold for recognizing a style. Data showed that three features are the lower bound for style recognition. Experiment 4 tested the degree of distortion in order to measure the recognizability of a stylistic feature. An expert was asked to recognize distorted pictures, and the results showed that beyond 40% geometric distortion a feature is no longer the label of the style. Observations made in this study suggest that if an artifact has at least three features attributing to a designer, an individual style exists. If four features repeat in a minimum of three different artifacts, an individual style is visible.

Style has been considered in aesthetics as the mode of expression, which is used for identifying the differences between periods, groups, or individual works. Historians of culture and philosophers of history study the forms and qualities that are shared by all the arts of a culture during a particular span of time. Art historians and critics study the change, formation, and history of style. They use style as a benchmark to date and to locate original works and as a means to trace the relationships among groups of works (Schapiro, 1961). Thus, art historians and critics create classes such as Gothic, Baroque, or Rococo on the assumption that a certain complex of elements common to a group of work is sufficiently stable, distinct, and relevant to justify characterizing it as a style (Ackerman, 1963).

Many studies in different fields have explored the meaning of style, but these studies have been concentrated mainly on the interpretations of the stylistic expressions rather than on the substance of style. For example, the degree of style and the measurement of style have not been covered in the literature. This is because such endeavors have been limited by the available methodologies and because empirical approaches to understanding the related phenomenon of style have been lacking. This study, where scientific methods are applied to explore the fundamental phenomena about style, has a rather different approach: to treat a style as an entity that has attributes and can be measured and recognized. Cognitive psychology, which deals with mental activities, is an appropriate tool for inquiry, particularly in exploring how beholders perceive and differentiate style. In this sense, the studies in this paper serve as a pilot research.

The purpose in this paper is to gain an understanding of the fundamentals of style and to lay foundations for studies on how styles result from creative processes (Chan, 1990; 1992a; 1992b; 1993). Architecture designed by architects is the subject matter used for observation. Three major concepts are addressed: the operational definition of style, the degree of style, and the measurement of style. A hypothesis about each, based on the premises developed by scholars in the field, is presented and then tested through four psychological experiments. In experiment 1, the operational definition is set up to examine if common features in artifacts can represent a style. Experiment 2 tests the phenomenon of the degree of a style. Experiments 1 and 2 were applied to thirty-one subjects. The group of thirty-one is statistically accepted as a modest sample population and the results should represent a larger sample (McGuigan, 1983). Experiment 3 tests how a style can be measured and experiment 4 is focused on the recognizability of features that represent a style. In experiments 3 and 4, the repeated-treatment method was used to test the level of recognizability of a style and of features. A similar method has been successfully used by Chase and Ericsson (1981) to study human memory. Therefore, if the results obtained statistically prove or reject the hypotheses, the conclusions drawn should be justified.

Definition of style: experiment 1

Style has been described in several ways:

> "By style is meant the constant form—and sometimes the constant elements, qualities, and expression—in the art of an individual or a group" (Schapiro, 1961, page 81).
>
> "In the study of the arts, works—not institutions or people—are the primary data; in them we must find certain characteristics that are more or less stable, in the sense that they appear in other products of the same artist(s), era or locale, and flexible, in the sense that they change according to a definable pattern when observed in instances chosen from sufficiently extensive spans of time or of geographical distance. A distinguishable ensemble of such characteristics we call a style" (Ackerman, 1963, page 174).
>
> "Style may be considered as the collective characteristics of building where structure, unity and expressiveness are combined in an identifiable form related to a particular period or region, sometimes to an individual designer or school of design" (Smithies, 1981, page 25).

These definitions share a common idea: that man-made objects can be categorized according to constant and recognizable forms (features). The replicated forms, when consistently occurring, represent an artist's particular and preferred way of expression. Inasmuch as forms are media for the manifestation of expressions, they can be characterized into different groups to symbolize different styles. This holds true for architectural design as well. Often, a designer uses certain features in his or her work to express certain meanings. If there are a number of feature repetitions—repetitions of a set of common features—the repetitive features become trademarks and the different sets of features characterize different styles. Hence, a style is operationally defined by the set of common features appearing in artifacts.

The term 'feature' used in the definition covers a wider range of meanings, such as physical forms, patterns, or certain distinguishable characteristics. Different to the meaning used in structural linguistics (Hampton and Dubois, 1993), the concept of feature developed in this study is limited to cover only forms of design products. Thus, a stylistic feature has the following properties: (1) it is a form or composition distinguished by some particular configuration or proportion; (2) it has some contextual relationship with other features; (3) it is originally generated by a designer through certain creative processes; (4) it is adapted or copied by a designer from other sources; and (5) it is a member of a set of prominent forms repeatedly used by the designer. If a feature is originally generated by a designer, it is a signature of an individual style. If a feature has been copied by many designers, it defines a group style. If a feature has been copied by many designers over a period of time, it defines a style of a period. Likewise, if a feature has been copied by many designers over a period of time in a particular region, it defines a regional style. In this research, the focus is on individual style. Thus, the notion that a style exists in products represented by a set of common features will be tested by the following experiment.

Method

Subjects

The subjects were first-year students from the College of Humanities and Social Sciences at Carnegie Mellon University who were, at the time, taking their first psychology course. Thirty-one students participated in this experiment for course credit. Because they were from fields other than architecture, they were not familiar with all the buildings presented to them. This prevented them from using prior knowledge to distinguish known styles.

Materials

Pictures of buildings designed by famous architects were presented to the subjects. There were three major styles included, and each was represented by eleven buildings designed by two architects. The three styles were the prairie style of Frank L Wright and Vernon S Watson, the modern style of Richard Meier and Michael Graves, and the vernacular style of Charles Moore and Robert Venturi. Wright, Meier, and Moore were the three main architects chosen for study and each had ten buildings included. The buildings, mostly residential, were selected to show common features in the pictures. For example, the number of common features appearing in the stimuli ranged from three to five for Moore,(1) five to six for Meier,(2) and nine to ten for Wright.(3)

The method used to develop the lists of features was to extensively label all namable features and primitive graphic elements on each picture and have them verified by three faculty members. Then the set of common features were identified as the features appearing on at least three pictures. These pictures were selected from publications, and each was mounted on a 4 inch by 6 inch white index card.

Procedure

The subjects were asked to sort out thirty-three cards of buildings into piles, each of which represented a style. They were not told the exact number of cards and were instructed that they could place any number of cards in each pile.

(1) Double-pitch roof, single-pitch roof, vertical redwood siding, protruded small units with single-pitch roof, shingle roofing, left and right declining composition, and white stucco surfacing.

(2) Full-height glass wall with mullions expressing the interior structural grid, round columns, circular staircase, horizontal band of parapet, walls of vertical wooden siding painted white, pipe handrail on patio, overhang circular platform, protruded solid round staircase, protruded staircase with solid handrail, protruded staircase with pipe handrail, and overhang staircase with pipe handrail.

(3) Low hip roof, a band of casement windows, continuous bands of sill, extended terraces with low parapet and coping, watertable, corner blocks, planting urns, massive brick chimney, continuous wall between sill and watertable, overhanging eaves, and symmetric side facade.

Result A: feature-set identification

In this experiment each style was characterized by a distinct set of features, but not all the features appear in each building. Sets of features can vary, permitting the definition of subsets nested in styles. Therefore, features appearing in the buildings also constitute subsets. For example, table 1 lists the set of common features (numbered 1 through 8) and the subsets that appear in the vernacular style represented by buildings designed by Moore and the Trubeck house by Venturi. Presented by sets, features 2, 3, and 4 and features 1, 3, and 4 define subsets B and D, respectively. There are five subsets apparent in the stimuli of the vernacular style as shown in figure 1. A set of features can stand alone, representing one style, or be subordinated to another set, representing a larger group of styles. For example, subset D {1, 3, 4} stands by itself, and it also joins subset C to form a larger group.

Table 1. Summary of features apparent in the stimuli of the vernacular style.

Building name	Feature[a]	Subset	Building name	Feature[a]	Subset
S B Club, 1968	2, 4, 6, 7, 8	A	Johnson, 1966	1, 3, 4, 5	C
Koizim, 1971	2, 4, 6, 7, 8	A	Ranch 11, 1969	1, 3, 4, 5	C
Burns, 1972	2, 4, 6, 7, 8	A	Swan, 1976	1, 3, 4, 5	C
Bonham, 1962	2, 3, 4	B	Bonham, 1962	1, 3, 4	D
Ranch 1, 1965	2, 3, 4	B	Trubeck, 1970	1, 4, 5	E
Ranch, 1966	2, 3, 4	B			

[a] 1 Double pitch roof; 2 single pitch roof; 3 vertical redwood siding; 4 protruded small units with single pitch roof; 5 shingle roofing; 6 left and right declining composition; 7 white stucco surfacing; 8 full opening.

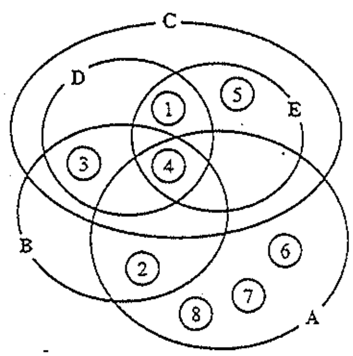

Figure 1. Venn diagram representing subsets in the stimuli of the vernacular style.

Although the subjects did not know the architects and had not seen the pictures beforehand, they had no difficulty in sorting the pictures into piles. They did not place pictures in a random manner; rather, the pictures were clustered into certain fixed groups. In Moore's pictures, for example, there were three buildings having the set of {2, 4, 6, 7, 8}, and twenty-two out of thirty-one subjects (71%) identified this and grouped these cards together. Twelve of these subjects put these three cards into a single pile to represent one style, and ten subjects mixed them as a group with the other style (Meier's cards).

Discussion

The result suggests that the subjects identify styles by distinguishing the same sets of features. In other words, common features were used as clues for sorting and categorizing pictures. According to Lakoff and Johnson (1980, page 163), a categorization is "a natural way of identifying a kind of object or experience by highlighting certain properties, downplaying others, and hiding still others". The subjects may have used different features to categorize the pictures. For clarification, interviews with the subjects were conducted, after the experiment, to determine the features used by them. The interviews showed that the subjects used the technique of comparing features and these features were indeed the ones included on the list. Thus, the lists did cover the possible common features from the stimulus.

Because there are many stylistic features in a given building, some features may have been left out in the analysis of probable factors. This could account for the variations in the data. For instance, 29% of the subjects did not sort subset A into a single pile. Because the analysis method used was based on the concept of sets and is similar to the notion of categorization as studied in psychology, the hypothesis is further tested in the following by the theory of categorization.

Categorization and similarity

The word 'category' means a set of entities or objects that are considered equivalent and are grouped together based on some criterion or rule. According to Bruner et al (1956) to categorize is to render discriminably different things equivalent, to group objects and events into classes, and to respond to them in terms of their class membership rather than their uniqueness. Studies of categorization have included natural-kind categories (Rosch, 1973; Rosch and Mervis, 1975; Rosch et al, 1976; Smith and Medin, 1981) and man-made objects (Reed and Friedman, 1973). These studies have developed the concept of categorization in terms of a prototype model (Posner and Keele, 1968) or the feature-frequency model (Hayes-Roth and Hayes-Roth, 1977). The prototype model suggests that a category is centered around a representative prototype. The feature-frequency model emphasizes a match of individual features and combinations of features. Another major approach to categorization is to compare the similarities between objects (Smith, 1989). The method has been described in terms of geometric analysis of similarity (Shepard, 1974) or by a hierarchical clustering scheme (Johnson, 1967). A more general model of similarity, based on feature matching, has been developed by Tversky (1977). His theory of similarity is that objects are represented as collections of features, and similarity is described as a feature-matching process. In his theory, the similarity is a combination of the measures of common and distinctive features. Thus, the similarity, S, between i and j is monotonically related to

$$S(i, j) = \theta f(I \cap J) - \alpha f(I-J) - \beta f(J-I), \qquad \alpha, \beta, \theta \geq 0 ,$$

where i and j are two objects; I and J are the sets of features in i and j, respectively; $I \cap J$ is the common set of features in i and j; and $I-J$ is the set of features in i but not in j. If $\theta = 1$, and $\alpha = \beta = 0$, the similarity between the objects is determined by their common features. If $\theta = 0$, and $\alpha = \beta = 1$, the dissimilarity is determined by their distinctive features only. The basic concept of Tversky's contrast model is that similarity is an increasing function of the properties common to the objects. Applying this model to the definition of style, we can rewrite the hypothesis so that the number of times buildings i and j appear in one pile, $N(i, j)$, may be a function of the similarity of the two buildings, $S(i, j)$. For example, the representation of the similarity between Moore's buildings Ranch and Johnson is

$$S(\text{Ranch}, \text{Johnson}) = \theta f\{3, 4\} - \alpha f\{2\} - \beta f\{1, 5\}.$$

The parameters in the model can be assigned any number. An increase in the common features increased similarity and decreases difference, whereas an increase in the distinctive features decreases similarity and increases difference. In this experiment, the subjects were not asked to single out any particular building as a referent for comparison during the sorting process. Every building was treated as equally important. Because there is no assigned subject of comparison with the referent, all features are equally salient and the distinctive features count as much as common ones (Tversky, 1977). Thus, the parameters in the model have equal weight (specifically, $\alpha = \beta = \theta = 1$), and the model becomes

$$S(i, j) = f(I \cap J) - f(I - J) - f(J - I).$$

In this modified model, the value of the similarity between two buildings, i and j, is determined by the number of common features and distinctive features in i and j. This value could range from positive (highly similar) to negative (highly dissimilar), depending upon the number of common and distinctive features in i and j. This value of $S(i, j)$ is also assumed to be a function of the number of times that i and j appear in the same pile. The greater the similarity value, the greater the probability that two buildings will be placed on one pile, otherwise the probability of having i and j shown on the same pile is zero. Therefore, this model accounts for both similarity and dissimilarity.

The number of times buildings i and j appear in one pile, $N(i, j)$, as a function of $S(i, j)$ was plotted; the result is shown in figure 2, with $y = 16.529 + 1.2828x$, $R^2 = 0.847$. These data suggest that style can be recognized by the common features between buildings. A larger set of common features appearing in objects will be easily identified as the same style.

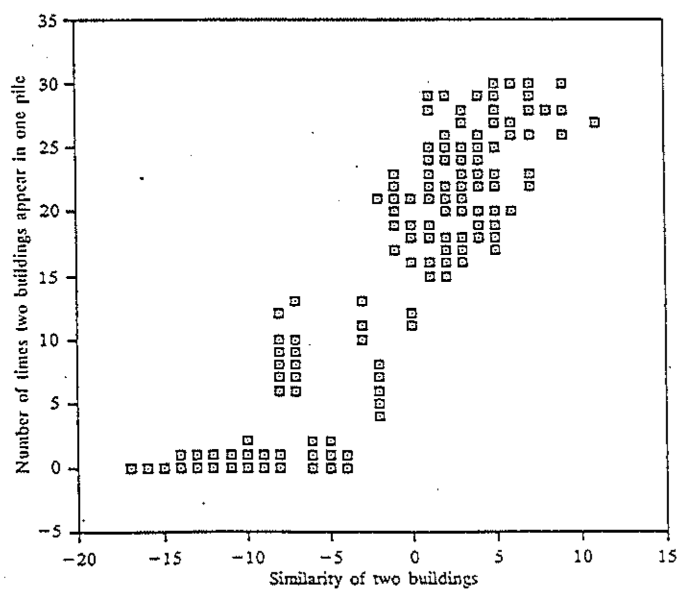

Figure 2. Relationships between the similarity of two buildings and the number of times they appear in one pile.

Result B: number of piles

The 'number of piles' corresponds to the total number of piles per style as sorted by the thirty-one subjects. Its value symbolizes the strength of holding a style. Statistical data of the numbers of piles are presented in table 2. The numbers are very close in Wright's (fifty-one piles) and Meier's (forty-eight piles) styles, but the value increases by 56% in Moore's style, despite the minor difference of feature numbers of between 5–6 and 3–5.

Table 2. Results of sorting the styles.

Name of style	Number of features	Number of piles	Average pile per subject	Number of misplaced piles	Number of single-style piles
Wright	8–11	51	1.65	1 to Moore	16
Meier	5–6	48	1.55	2 to Moore	16
Moore	3–5	75	2.42	15 to Meier	0

Discussion

In comparisons of the feature differences among three different styles, the 56% increase in the number of piles from Meier to Moore is considered significant. This phenomenon suggests that more common features in products would: (1) make a style coherent and less detachable (fewer piles), and (2) provide less possibility of splitting a style into multiple piles. Two other observations support this inference. First, fifteen out of the thirty-one (48%) subjects misplaced Moore's pictures (which have the least number of features) into Meier's, but only one and two subjects, respectively, misplaced Wright's (with the largest number of features) and Meier's (the medium number of features). This indicates that fewer features in artifacts result in poor identification, so that error occurs more frequently. Second, an equal number of subjects (sixteen out of thirty-one, 52%) successfully identified Wright's and Meier's pictures as one style (putting the pictures into a single pile), but none successfully identified all of Moore's pictures. Thus, a larger number of common features tend to more cohesively define a style; or the bigger the number of common features a style has, the easier it will be for subjects to identify it.

Result C: pattern match

As shown in table 1, the features in each picture can be coded by numbers, and a common set is categorized and marked by a letter. For the vernacular style represented by Moore and Venturi, there are five sets of features, which have thirty-one possible ways of being sorted into one pile (set combination). If the subjects were sorting randomly, the thirty-one ways would be expected to appear with equal frequency. But only twenty differently combined patterns appeared within the total of ninety sorted piles (20 is 22% of 90). Thus, it is assumed that there are certain rules governing the subjects' behavior.

For example, there are ten possible ways of combining two sets from A, B, C, D, and E into one pile, (see table 3). But the results showed only six different combinations in twenty-six piles. To compare the probability of occurrence, a computer simulation (a Pascal program with a random-number generator) was conducted to generate random pattern combinations 1000 times (piles). The results showed that the probability range from 8.2% to 12.3% for each pattern combination and that the number of piles ranged from eighty-two to 123. Another simulation generated twenty-six piles, in which each combination appeared at least once, and the probability ranged from 3.8% to 19.2%, which differs from the experimental result (0 to 42.3%). These simulations suggest the existence of rules governing the sorting procedure.

Table 3. Results from simulating a two-set combination.

	AB	AC	AD	AE	BC	BD	BE	CD	CE	DE
Simulate 1000 piles	98	101	99	123	87	103	101	118	88	82
Probability (%)	9.8	10.1	9.9	12.3	8.7	10.3	10.1	11.8	8.8	8.2
Simulate 26 piles	3	1	2	5	5	4	1	1	3	1
Probability (%)	11.5	3.8	7.6	19.2	19.2	15.4	3.8	3.8	11.5	3.8
Experimental result (26 piles)	3	0	0	0	11	1	0	2	4	3
Probability (%)	11.5	0	0	0	42.3	3.8	0	7.7	15.4	11.5

It was also found from the experiment that if the number of common features between two sets of pictures is less than two, these two sets will not be merged into one pile. Because this rule accounts for 93% of the data, it is inferred that the interset of two sets must have at least two common features. Furthermore, subjects assemble cards by using certain pivotal features which become index keys for assembling cards. For example, if one focuses on features {6, 7}, then all cards having features {6, 7} will be put into the same pile and the rest will go to another pile. Table 4 lists some of these inferred rules that explain 65% of the data.

Table 4. Rules for sorting pictures into piles.

Rule 1: If key feature = {6, 7, 8},	then piles = {A} and {B, C, D, E}
Rule 2: If key features = {6, 7, 8} and {2, 3},	then piles = {A} and {B} and {C, D, E}
Rule 3: If key features = {6, 7, 8} and {3, 4}	then piles = {A} and {C, B, D} and {E}
Rule 4: If key features = {2} and {5},	then piles = {A, B} and {C, E} and {D}
Rule 5: If key features = {2, 4} and {1, 4},	then piles = {A, B} and {C, D, E}
Rule 6: If key feature = {2, 4},	then piles = {A, B} and {C, D, E}
Rule 7: If key features = {2, 4} and {4, 5} and {1, 3},	then piles = {A, B} and {C, E} and {C, D}

Discussion

The feature-grouping analyses describe only 65% of the consistency in pattern-matching behavior. It is assumed that the 35% inconsistency is caused by a switch of attention to different pivot features, which causes confusion. Or it may be that the limited span of short-term memory cannot hold the pivot features consistently throughout the sorting process. As the subjects had to work on three different styles for ten to fifteen minutes, each had to use at least three pivot features.

Summary

Style is identified by grouped features. Any set of features may represent a style by itself, or it can combine with other sets to define other styles if, as inferred from the data analysis, there are more than two common members in the interset.

More features tend to make the style coherent and strongly hold the style together. Some features are more salient than others; hence, the salient features provide a good approximation for the recognition of the style. This explains why some styles (Meier's) are easier to recognize than others (Moore's).

The more features that make a style coherent, the greater the degree of style. In other words, some styles are visually more recognizable than others. For example, certain artists' works have a stronger stylistic tendency, whereas others are weak in expression. This leads to the next study of the degree of style, and it is assumed that this relates to the number of common features present.

The degree of style: experiment 2

The stimuli used in experiment 1 consisted of different styles and different architects, thus the sets of features varied accordingly. It might be argued that certain features have a stronger visual impact than others, with distinct results. In the following experiment the stimuli are restricted to buildings designed by the same architect, and features are selected from a common set to balance out the differences between substantially different features from different architects.

Hypothesis The degree of a style is in proportion to the number of common features presented in products.

Method

Subjects

The subjects were the same group as in experiment 1.

Materials

The stimuli were thirty pictures of residential buildings (listed in table 5) designed by Wright, collected from different publications. The pictures were selected to include the eleven features identified in experiment 1. They were divided into eleven groups of different numbers of features, from none to ten. Each group had three different pictures that had the same number of, but different, features drawn from the common feature set. The pictures were mounted on 4 inch by 6 inch white index cards.

Procedure

In the beginning, the subjects were shown three pictures (Fricke House, 1902; Willitts House, 1902; Little House, 1903) that had ten features representing the typical style of the architect. After they had looked at these pictures as long as they wanted, they were asked to use these pictures as referents to sort the other thirty pictures (with from none to nine features) into four different piles that stood for four resemblance scales of strongly like, like, unlike, or strongly unlike the corresponding typical style shown earlier.

Result A: distribution of resemblance

Table 5 provides the names of the stimuli (buildings), the number of responses on each resemblance scale, and the mean scores for each feature group—scores of 1 (strongly unlike), 2 (unlike), 3 (like), or 4 (strongly like) are assigned. The distribution of the four resemblance scales obtained from a plot of the number of features versus the total number of response is shown in figure 3. The ranges of the scales in figure 3 were 0-4 features for strongly unlike, 0-8 for unlike, 1-9 for like, and 3-9 for strongly like. With a 95% interval of the distributions, the four scales were ranged as 0-3 features (strongly unlike), 1-5 features (unlike), 2-9 features (like), and 5-9 features (strongly like). This indicates that the spreads of the scales fall into clusters that move upward along the scale as the number of features increases.

Table 5. Data obtained from experiment 2.

Card index	Name and date of building	Strongly unlike	Unlike	Like	Strongly like	Mean score
Number of features = 0						
1	Madison, 1923	31	0	0	0	1
2	Ennis, 1924	30	1	0	0	1.03
3	Jones, 1929	31	0	0	0	1
Group mean score = 1.01						
Number of features = 1						
4	Jacobs, 1937	10	20	1	0	1.71
5	Willey, 1934	11	18	2	0	1.71
6	Winkler, 1939	19	11	1	0	1.42
Group mean score = 1.61						
Number of features = 2						
7	Winkler, 1939	11	18	2	0	1.71
8	Hickox, 1900	3	24	4	0	2.03
9	Coonley, 1912	7	18	6	0	1.97
Group mean score = 1.90						
Number of features = 3						
10	Bach, 1915	5	17	9	0	2.13
11	Gale, 1909	4	21	6	0	2.06
12	Winslow, 1893	8	10	12	1	2.19
Group mean score = 2.13						
Number of features = 4						
13	Hoyt, 1907	3	11	15	2	2.52
14	Martin, 1902	2	12	16	1	2.52
15	Husser, 1899	5	6	16	4	2.61
Group mean score = 2.55						
Number of features = 5						
16	Allen, 1917	0	6	16	9	3.10
17	Hunt, 1907	0	11	14	6	2.84
18	Dana, 1903	0	7	17	7	3.0
Group mean score = 2.98						
Number of features = 6						
19	Evans, 1908	0	5	16	10	3.16
20	Adams, 1913	0	7	16	8	3.03
21	Thomas, 1901	0	0	18	13	3.42
Group mean score = 3.20						
Number of features = 7						
22	Gridley, 1906	0	1	16	14	3.42
23	May, 1909	0	1	12	18	3.55
24	Boynton, 1908	0	0	11	20	3.65
Group mean score = 3.46						
Number of features = 8						
25	Tomek, 1907	0	1	10	20	3.61
26	Martin, 1902	0	0	7	24	3.77
27	Barton, 1903	0	0	14	17	3.55
Group mean score = 3.64						
Number of features = 9						
28	Robie, 1909	0	0	5	26	3.84
29	Little, 1903	0	0	10	21	3.68
30	Martin, 1902	0	0	6	25	3.81
Group mean score = 3.78						

To test whether the selected features are a significant influence on the resulting resemblance scores, a general linear model was developed as:

$$T = \beta_0 + \beta_1 F_1 + \beta_2 F_2 + ... + \beta_{10} F_{10} + \varepsilon_i ,$$

where T is the score, F represents features, and ε_i represents random errors. The results show that the features scored in the model were significant [$R^2 = 0.693$, $F(10, 919) = 207.55$, $p < 0.00001$], and each individual feature contributed a significant influence on the resulting score (the p values of the features ranged from $p < 0.0185$ to $p < 0.0001$). This proves that the features significantly correlate to the scores.

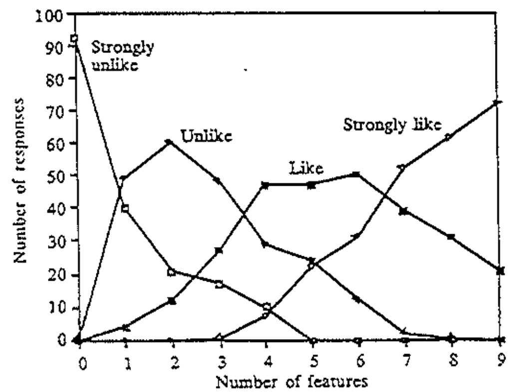

Figure 3. Distribution of the four scales of resemblance.

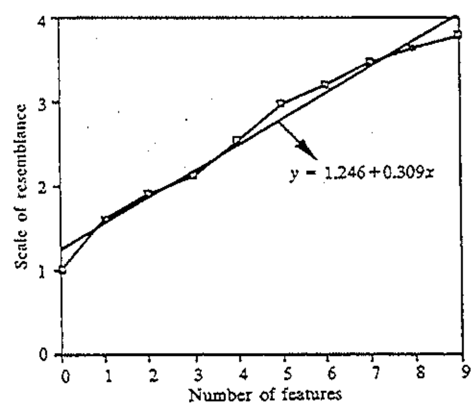

Figure 4. The degree of style versus the number of features.

Figure 4 is a plot of the average score per feature. When the number of features (x) and the resemblance scores (y) were fitted by a simple linear-regression model, the result of the equation was $y = 1.246 + 0.309x$. The slope was significantly different from zero [with $t(928) = 44.87$, and $p < 0.00001$], and the model explained 68.5% of the variance in y.

Discussion

The positive sign in the linear equation indicates the positive relationship between the number of features (x) and the scale of resemblance (y). Because the scale of resemblance denotes the degree of similarity between the subject buildings and the referent buildings that stand for the prototypical style, and a higher score of resemblance represents a stronger similarity to the referent style, the degree of style is in proportion to the number of common features present. Another observation made from figure 3 is that the four scales are all greater than zero at four features, and pictures with fewer than four features tend to be unlike. Hence, four is proposed as the number of common features needed to represent a style.

Result B: subject effect and feature effect

In table 5, the mean score of resemblance varies from picture to picture, regardless of the number of features. It is suspected that this could be caused by two factors: the subjects and the features. In other words, different responses from the individual subjects and different features could cause the variations in the score. A general linear model was designed to test the significance of these two factors. In this model, the resemblance score was modeled as the sum of the student effect and ten feature effects.

$$T = \beta_0 + \beta_1 E + \beta_2 F_1 + \beta_3 F_2 + ... + \beta_{11} F_{10} + \varepsilon_i ,$$

where E denotes the student effect, F are the different features, and ε_i is the random error. Statistical results show that this model explains 73% of the total variability [$F(40, 889) = 60.33$, $p < 0.00001$, $R^2 = 0.731$]. To determine whether the features or the number of features determined the resemblance score, a reduced model was developed. This model fits only the student effect and the total number of features, and had the following equation:

$$T = \beta_0 + \beta_1 E + \beta_2 N_F + \varepsilon_i ,$$

where E denotes the student effect, N_F is the number of features, and ε_i is the random error. This model had a slightly reduced R^2 value [$F(31, 898) = 75.30$, $p < 0.00001$, $R^2 = 0.722$].

Discussion

A comparison of these two models shows only a 0.86% increase in the explanation of the total variability, which is not justified by nine degrees of freedom. Thus, the second model is more suitable for explaining the data. Statistically, this model suggests that the differences between features (as in the full model) do not contribute to the resemblance scores as much as do the number of features (as in the reduced model). Therefore a style is determined by the number of features more than by any particular feature. This result implies that to identify a style, the low hip roof is no more important than the casement windows. It is possible, however, that the feature effect is balanced out by the individual student effect.

Some students may detect more similarities than others and give more responses towards the 'like' or 'strongly like' scale. Thus, different students give different scales of responses for the same pictures. This inference is supported by statistical results in both models. Both are significant at the 0.0001 level [$F(30, 889) = 4.15$, $p < 0.0001$; and $F(30, 898) = 4.06$, $p < 0.0001$], but the number of features is more significant than the student effect in the reduced model [$F(1, 898) = 2212.60$, $p < 0.00001$].

Summary

It can be concluded from the data analysis that the degree of a style is related proportionally to the number of common features that appear in the artifacts. More features in an artifact make a style more expressive. The metaphor is that the number of common features represents the strength of a glue—a larger number of common features will increase the strength of the glue to hold pictures together in a pile. The results do not support the notion that certain features have a stronger impact than others. Rather, the number of features is more significant in identifying a style. Thus, the degree of style is closely related to the number of features and this varies among beholders.

Measurement of style

Experiment 2 was concentrated on the products of an architect through a study of the similarities in style between the products. The next two studies are focused on the style in one product to explore how a style can be measured. Measurement is defined as the use of numbers to describe the attributes of objects or events. The measurement of a style is the threshold for recognizing a style. Two concepts—feature frequency and recognizability—are developed.

Experiment 3: feature frequency

In this experiment style in a product is addressed to discover whether the number of features would determine the recognition of a style. Feature frequency is used to measure an individual style.

Hypothesis Style is measured by the occurrence of features in a product; some features are more effective than others.

Method

Subject

The subject is a faculty member who was teaching architectural history in the Department of Architecture at Carnegie Mellon University when this experiment was conducted. He has a PhD in architectural history and is an expert on Wright's style.

Materials

Wright's side elevation of Little House, designed in 1903, with six recognizable features, was used for the stimuli. These six features in the original elevation were taken away one at a time, two at a time, and so on until all were gone. There were sixty-four permutations by which features could be taken away, such as taking one to three features away as in figure 5 and taking four to six features away as in figure 6. These figures served as the sources for the sixty-four pictures. Each picture was mounted on a 4 inch by 6 inch white index card.

Procedure

The sixty-four pictures from figures 5 and 6 were shown one at a time to the subject, who was asked to make a judgement by answering yes or no about whether the picture could be considered as Frank Lloyd Wright's style. The whole set of stimuli were repeated six times, and each time the cards were shuffled to avoid random error. In the intervals between the sessions, the subject would take a five-minute break to release his visual focus.

Result A: measurement of style

There are seven categories of the number of features, ranging from zero to six. The number of stimuli per category is determined by the permutation drawn from the set

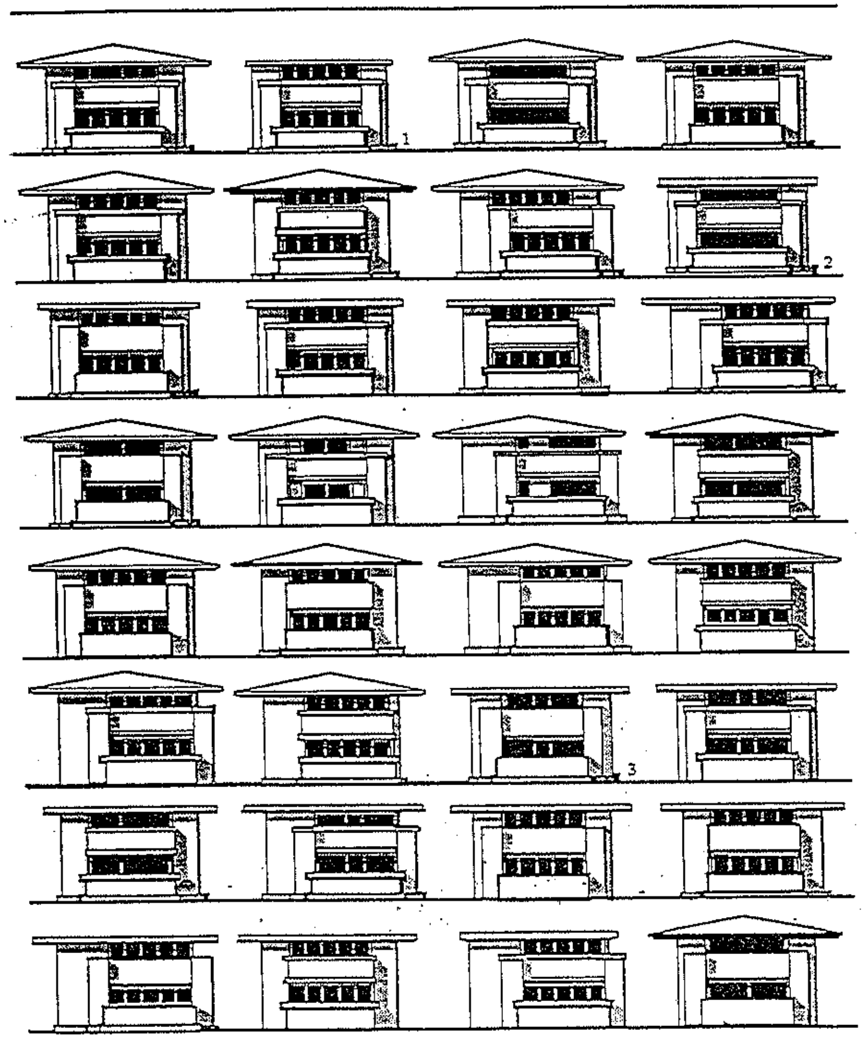

Figure 5. The changes of Wright's elevation in the Little House design used as the stimuli (the number represents the number of features being taken away from the elevation).

of six features. Each category consists of pictures with the same number of, but various, features. Because each stimulus repeats six times, the number of responses per category ranged from six (zero and six features) to 120 (three features). The results of the positive responses per category show that the presence of all six features had a full score. When the number of features is three, the probability of a positive response is only 5%, which stands for six positive responses out of 120 responses and is considered extremely poor. The probability remains constant at two features and drops to 0% with one feature present.

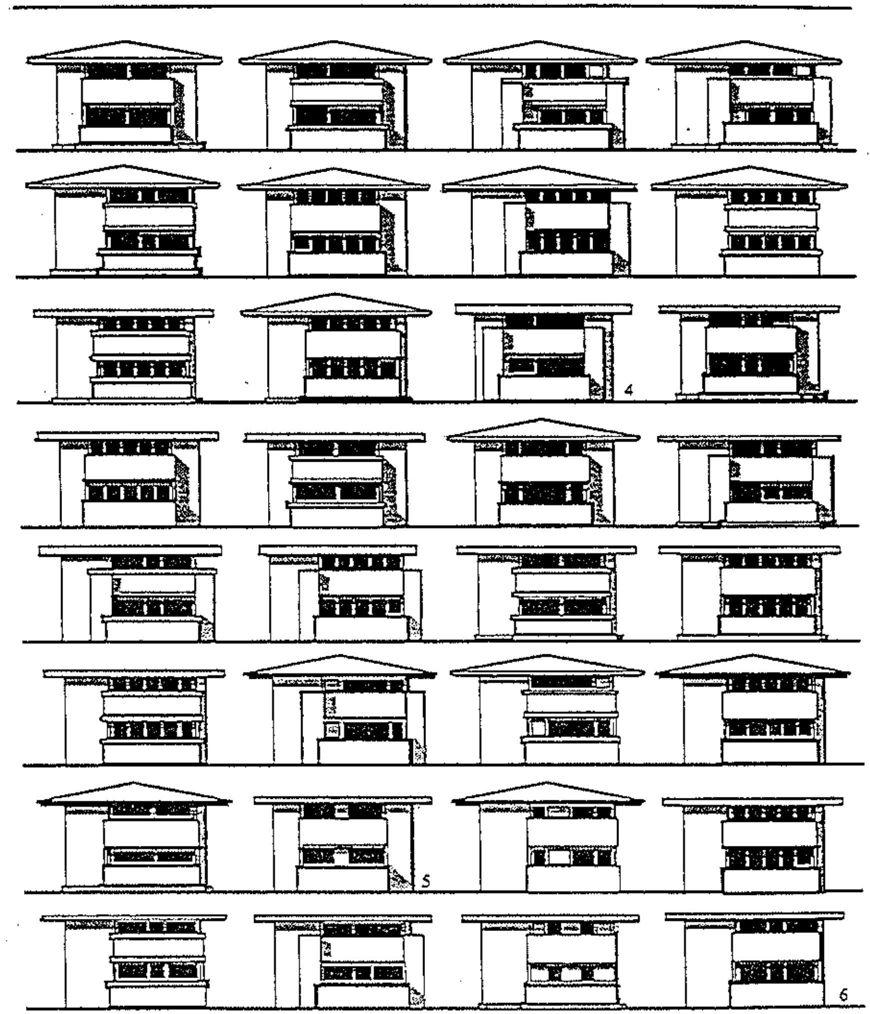

Figure 6. The changes of Wright's elevation in the Little House design used as the stimuli (the number represents the number of features being taken away from the elevation in the preceding pictures).

Discussion

Figure 7 shows that the number of positive responses decreases as the number of features drops. This suggests that the number of positive responses (representing the recognizability of a style) is in proportion to the number of features present in a product. More features will make a style more visible. From the nature of the curve and the value of the probability, it is inferred that three features could be the lower bound for style recognition. This suggests that a style is no longer recognizable when the number of features is three or fewer, despite the contents of the features, and is measurable when there are more than three features.

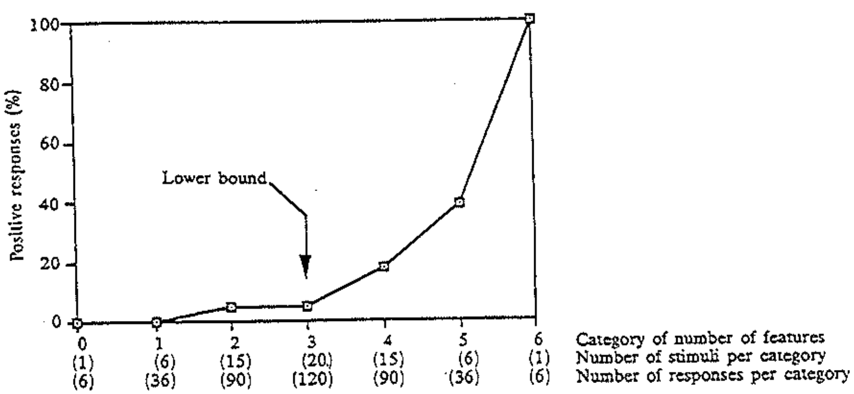

Figure 7. Probability plot of the responses per number-of-features category.

Result B: effectiveness of features

The features in this experiment can be treated as variables with the value of 'absence' or 'presence'. For example, a low hip roof is a variable present or absent in sixty-four stimuli. Table 6 arranges the features in a factorial design where a two-letter code stands for the feature (a negative sign in front indicates absence). Calculated from the table, the grand means of each feature's absence and presence, corresponding to positive responses, are shown on the side of the table. The probability of a positive response is greatest for the presence of a casement window (about 22%, and 2% when it is absent). The weight of six variables is, therefore,

Table 6. Number of responses per case and the grand mean of each feature.

Feature			CB				−CB			
			C		−C		C		−C	
			R	−R	R	−R	R	−R	R	−R
S	WT	W	6	1	2	1	4	2	1	1
		−W	0	0	0	0	0	0	0	0
	−WT	W	5	1	3	0	2	1	2	1
		−W	0	0	0	0	0	0	0	0
−S	WT	W	2	2	0	0	2	0	0	0
		−W	0	0	1	1	0	0	0	0
	−WT	W	2	1	0	0	0	0	0	0
		−W	0	2	0	0	0	0	0	0

1. Mean W = 21.87%; mean −W = 2.08% (casement window).
2. Mean S = 17.18%; mean −S = 6.77% (symmetry).
3. Mean C = 17.18%; mean −C = 6.77% (Coping).
4. Mean R = 16.66%; mean −R = 7.29% (low hip roof).
5. Mean CB = 15.63%; mean −CB = 8.33% (corner blocks).
6. Mean WT = 13.54%; mean −WT = 10.42% (watertable).

ranked by the value of the probability of the responses for feature when present and absent. The order implies that some features are more attractive to the subject than others are.

The grand mean of the absence and presence of each variable in table 6 does not exclude the possibility of interactions among variables. To eliminate the effect of interaction, the responses were studied while a particular feature is missing. For example, when the watertable is absent from the picture, the positive responses are five out of six trials, whereas when the casement window is absent, there is no positive response at all. A reasonable explanation is that the watertable is not an important feature to the subject for judging the style, but the casement window is. Thus, if the responses are sorted by one absent feature and the variables are arranged by the number of responses, the degree of significance of the feature emerges as shown in table 7. This result is very close to the last one in table 6; thus some features are more effective than others.[4]

Table 7. The number of responses per absent feature.

Absent feature	Present features	Number of positive responses
W (Casement window)	S, WT, R, C, CB	0
R (Low hip roof)	S, WT, W, C, CB	1
C (Coping)	S, WT, W, R, CB	2
S (Symmetry)	WT, W, R, C, CB	2
CB (Corner blocks)	S, WT, W, R, C	4
WT (Watertable)	S, W, R, C, CB	5

To verify further the significances of the features to the subject, the analysis is focused on a particular present feature and its relationship with other present features. Thus, when a particular feature appeared, the probabilities of the responses when any number of other features were present was recorded, as in table 8 and figure 8. In the figure, each curve represents the correlation between the responses and the feature-numbers condition for a particular present feature. The shape of the curves suggests a scale of W > R > C = S > CB > WT, which is similar to the result in table 7, but interaction occurs when the number of features is less than two.

Table 8. The probability of responses for a particular present feature.

Number of features	Window	Low hip roof	Coping	Symmetry	Corner blocks	Watertable
0	0	0	0	0	0	0
1	0.033	0	0.067	0.033	0.100	0.033
2	0.083	0.050	0.033	0.067	0.033	0.033
3	0.267	0.167	0.183	0.167	0.150	0.133
4	0.467	0.433	0.400	0.400	0.333	0.300
5	1.000	1.000	1.000	1.000	1.000	1.000

[4] The contradiction between this result and the results from experiment 2 (that a style is determined by the number more than the identity of characteristic features present) may be explained by the fact that the feature effect is balanced by the variable of the individual student effect which is not present in this experiment.

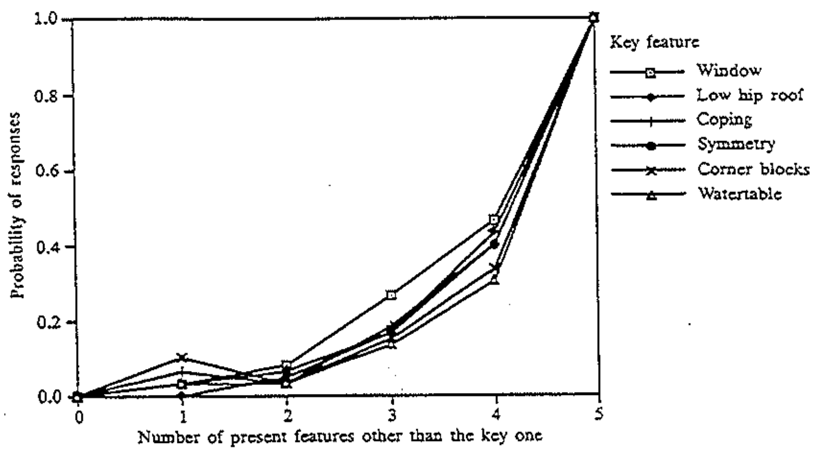

Figure 8. Probability of the subject's responses to a particular feature being present.

Result C: interaction among features

Figure 8 suggests that corner blocks and coping are two features generating the interactions. This result explains why, when the number of present features decreases to three or fewer, the perception of a style is less feasible in figure 7. Explained from another perspective, the feature interactions occur when the total number of features in an artifact is less than three. Because the recognizability of a style is based upon the perception of the style, it is inferred that three is the lower bound for the measurement of a style.

Summary

The data analyses suggest that there is a scale of importance among features in an artifact for a beholder to identify a style. This style judgement does not relate to the sizes or dimensions of the features, but is determined by the number of features present.(5) For example, roofs have the largest volume among the features, and they should be the most noticeable features, but they rank third and second in tables 6 and 7, respectively. When the number of features in an artifact is reduced to three or fewer, interaction occurs and a style is no longer perceptible.

Experiment 4: recognizability of stylistic features

Another measurement of style is the extent to which a borrowed and modified feature remains recognizable as an element of a style. It is argued that whatever changes occur to a feature, as long as its shape is maintained within certain ranges of proportion, it will retain its identity with a perceived style. Thus, topological characteristics of an artifact provide a measurement for an individual style.

Hypothesis Style can be identified within a certain range of geometrical distortion. Topological distortions disable the style identification.

(5) In this experiment, the size is not an issue as long as it is within the normal and expected range for that kind of feature; an unusually massive roof or a set of tiny windows will certainly be noticeable.

Method

Subject

The subject was the same one who participated in experiment 3.

Materials

Wright's stylistic elements apparent in the Little House were used for stimuli (the original picture is shown in the uppermost left of figure 5). The dimension of each feature was either longated or reduced by 10% at a time along either the x or y axis. Thus, $y+2a$ meant that the vertical dimension of an object was increased by 20%, and $x-2a$ meant that its horizontal length was decreased by 20%. The distorted features include horizontality, verticality, roof, coping, corner block, and watertable. The geometrical distortions, ranging from 10% to 50%, are shown in figure 9.

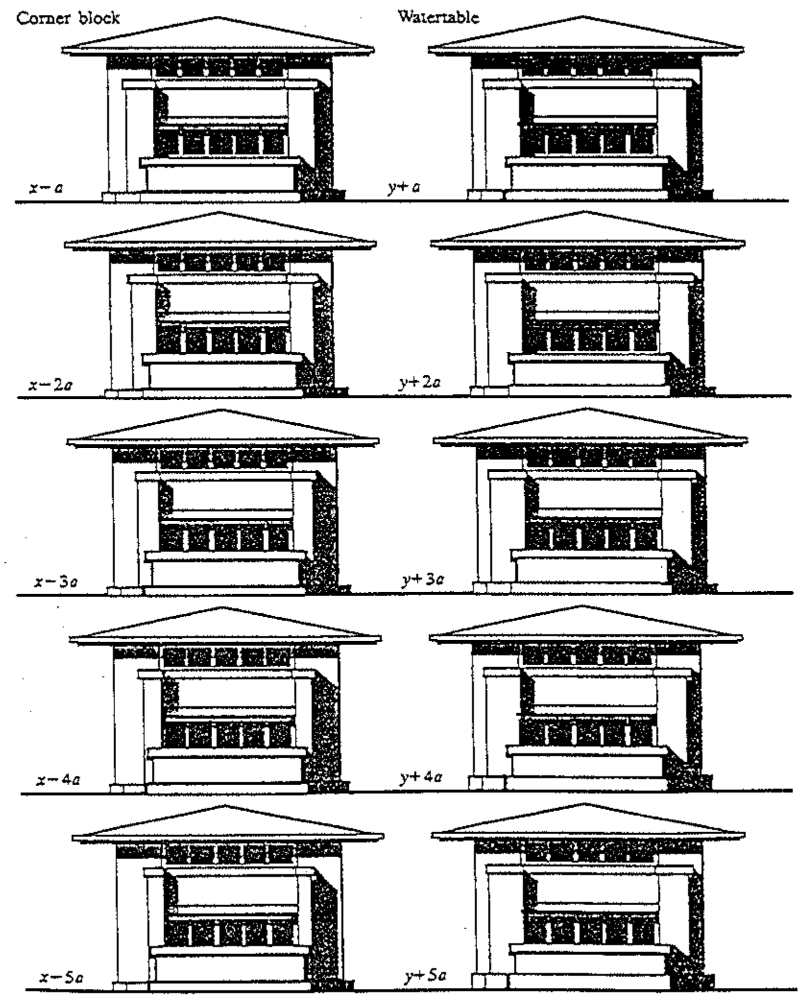

Figure 9. The changed forms of Wright's corner block and watertable.

There were also three pictures having a topological change, as shown in figure 10. These thirty-three figures were the stimuli, and each picture was mounted on a white 4 inch by 6 inch index card.

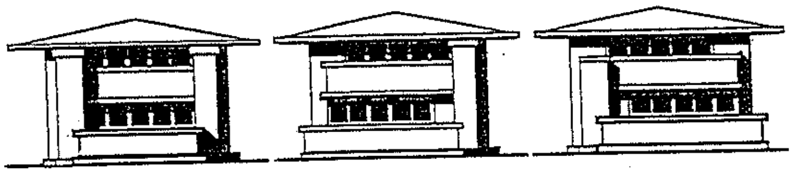

Figure 10. Topologically distorted forms of Wright's elevation.

Procedure
Thirty-three cards were shown to the subject, one card at a time, and the subject was asked to make a judgement about whether the picture could be regarded as Wright's style. This experiment was repeated six times, and the cards were shuffled each time to avoid random sources of error. Between sessions, the subject took a three-minute break to release his visual focus.

Result A: degree of distortion
Plots of the percentage of positive responses versus the degree of distortion are displayed in figure 11. Distortions of the massing, coping, and watertable shared a negative relationship. The curves in the plots of the distortions of the roof and corner blocks were not regular. For example, when the roof increased by 40% of its original height, the response increased. This phenomenon also occurred in the plot for the corner blocks. To verify these results, the same set of these two stimuli was tested again on the same subject a month later. The results of the second experiment are displayed in the last two plots in figure 11. The curve for the corner blocks exhibited a regular negative relationship. But the result for the roof remained the same. An interview with the subject afterward indicated that the picture with the 40% increase in roof height resembles Wright's two early houses—the Heller House (1897) and the Husser House (1899)—thus the subject gave positive responses.

In response to the three topologically distorted pictures (total of eighteen observations), the subject did not attribute any of these forms to Wright's style. Hence it is reasonable to conclude that topological relationships are crucial in maintaining an individual style.

Result B: threshold of recognizability
The stimulus threshold is commonly taken to be the stimulus intensity for which there is a 50:50 chance of detection, or for which the probability of detection is 0.5, as suggested by Kurtz (1966). When this is taken as a criterion, the thresholds for recognizability range from 20% to 60%, and they vary from feature to feature. This range of recognizability is attributable to the fact that certain features stand out and thus are more immediately identifed by the beholder.

Discussion
The proportional distortion of a feature can be tolerated to a certain extent. But topological distortions change the relationships among features, which changes the characteristics of the object and, consequently, alters the style. This suggests that there exists a topological structure (characteristic context) defined by the topological

relationships among stylistic features. After any topological change, a feature will no longer serve the representative role of a style, whereas the geometric distortion of a feature has a more tolerable range of recognizability. Therefore, topological relationships among objects are a crucial factor for sustaining a style.

The fluctuation of the responses to the different objects implies that the recognizability of features used for identifying a style does not correlate to their

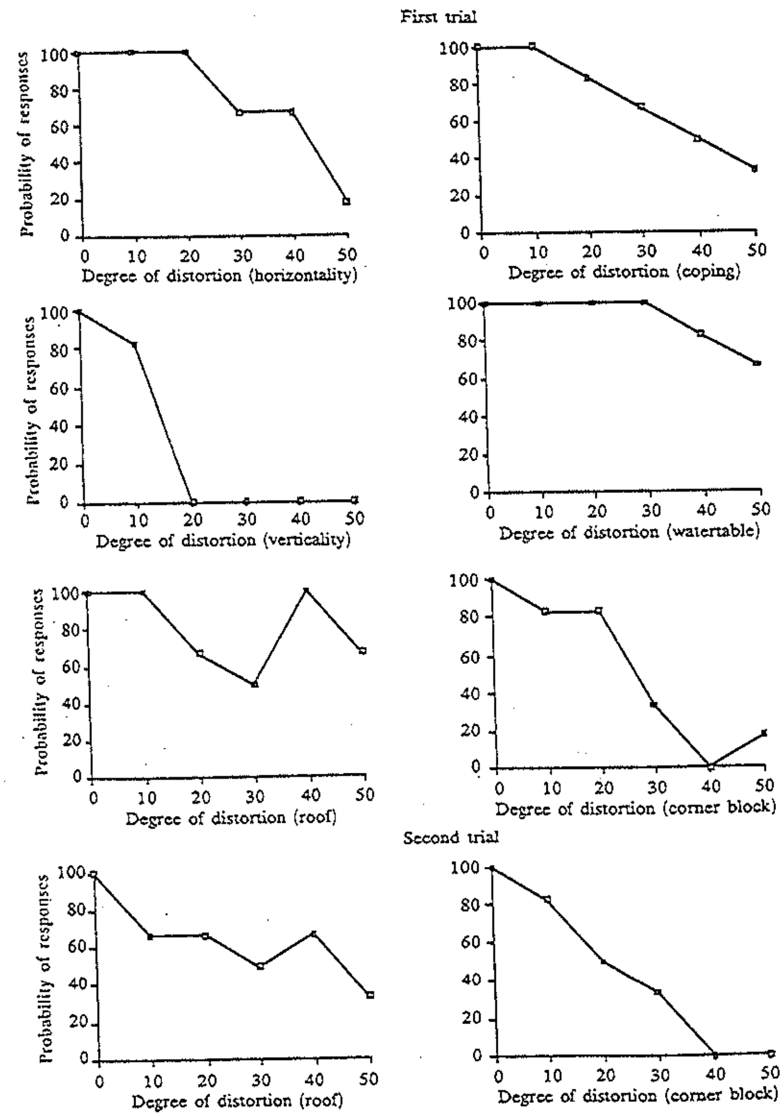

Figure 11. The results of each feature distortion.

dimensions but relates to the beholder's perception. Some objects are more noticeable to the beholder than others. For example, a coping is a small element assumed to be least conspicuous among features, but it had a constant slope of response in figure 11. In this experiment, the subject is more acutely aware of the changing properties of coping and corner blocks than of other features.

The result, that the recognizability of style is a function of the geometric distortion of features, can be explained by the prototype model derived from theories of categorization. In this experiment, the subject had mentally formed a concept (prototype) of the original picture presented in the beginning and used it to categorize distorted pictures. The level of distortion is analogous to the level of generality suggested by Rosch et al (1976) in the comparison of categories. The results obtained here were similar to their finding that it was the most general level at which a prototype could be formed to represent the category as a whole. Thus, a less general (more distorted) level will be less representative of a category (style).

Conclusion

The series of experiments conducted in this research provide some understanding of style, but the conclusions drawn are not intended to represent a formal theory of style. Instead, the study is a start in the research of style. Similar analyses of architectural style as determined by sets of features have been done by Atkin (1974; 1975) using Q-analysis to sort out the clusters of features that characterize Tudor-style architecture in Lavenham (England). One of the criticisms of that work (Couclelis, 1983; Pinkava, 1981) is that mere sets of features do not make a style, because features must also exist within a characteristic context (topological structure). This is shown by experiment 4 such that any topological change will distort a style, and the style will not be recognizable anymore. Therefore, the characteristic context also plays an important role for the identification of a style. This is especially important for the study of historical, regional, and group style because quite often features from many different styles and periods may coexist in an untypical spatial relationship to each other as in these days of eclectic and post-modern architecture.

In summary, experiments 1 and 2 show that a style is represented by a common set of features that appear in objects. This set of common features defines a style. Following the concept developed by Schapiro (1961), the current research suggests that style is the constant form in works of art. It also indicates that the more features used by an artist, the more coherent the style, creating a higher degree of similarity among artifacts. Different individuals perceive features differently, and different features also yield different impacts on the same beholder.

Experiment 3 shows that style is measured by the number of features present in a given set of artifacts. If the number of features drawn from a common set is fewer than three, then a beholder will not be able to recognize the style of the artifact, or the recognizability of style is not measurable. Another inference made here is that if there is a set of features that attributes a style, and a designer picks up at least three elements from this set to form a product, then the final form composed of these elements contains at least one style.

Experiment 4 shows that if a feature drawn from a common set has been geometrically distorted up to approximately 40%, it is still recognizable as the representative stylistic feature. Distorted beyond this level, a feature should not be considered as representative of a style. The value of the degree of distortion varies among features, but basically it is determined by beholders' perceptions.

The experiment also shows that the topological structure among stylistic features is the most important factor for sustaining a style.

Finally, conclusions can be condensed into two sentences to suggest the operational concepts of style: (i) if an artifact consists of at least three recognizable features (result from experiment 3) within a topological structure (result from experiment 4), then a style exists in this artifact; (2) if there are four features (result from experiment 2) with the same topological structure replicated in at least three different artifacts (results from experiments 1 and 4), then these four features and the structure represent an individual style.

Acknowledgments. The author is grateful to John R Hayes for his support in experiments conducted in this research. Thanks to Omer Akin and Herbert A Simon for their criticism concerning the experiments, to David Banks for his help as a statistical consultant, and to the anonymous reviewers for their valuable comments.

References

Ackerman J S, 1963, "Style", in *Art and Archaeology* Eds J S Ackerman, R Carpenter (Prentice-Hall, Englewood Cliffs, NJ) pp 174-186

Atkin R H, 1974 *Mathematical Structure in Human Affairs* (Heinemann Educational Books, London)

Atkin R H, 1975, "An approach to structure in architectural and urban design" *Environment and Planning B* 2 21-57

Bruner J S, Goodnow J J, Austin G A, 1956 *A Study of Thinking* (John Wiley, New York)

Chan C S, 1990, "Cognitive processes in architectural design problem solving" *Design Studies* 11(2) 60-80

Chan C S, 1992a, "Exploring individual style through Wright's designs" *Journal of Architectural and Planning Research* 9(3) 207-238

Chan C S, 1992b, "Exploring individual style in design" *Environment and Planning B: Planning and Design* 19 503-523

Chan C S, 1993, "How an individual style is generated" *Environment and Planning B: Planning and Design* 20 391-423

Chase W G, Ericsson K A, 1981, "Skilled memory", in *Cognitive Skills and Their Acquisition* Ed. J R Anderson (Lawrence Erlbaum, Hillsdale, NJ) pp 141-189

Couclelis H, 1983, "Some second thoughts about theory in the social sciences" *Geographical Analysis* 15 28-33

Hampton J, Dubois D, 1993, "Psychological models of concepts: introduction", in *Categories and Concepts: Theoretical Views and Inductive Data Analysis* Eds I V Mechelen, J Hampton R S Michalski, P Theuns (Academic Press, London) pp 11-33

Hayes-Roth B, Hayes-Roth F, 1977, "Concept learning and the recognition and classification of examples" *Journal of Verbal Learning and Verbal Behavior* 16 321-338

Johnson S C, 1967, "Hierarchical clustering schemes" *Psychometrika* 32 241-254

Kurtz K H, 1966 *Foundations of Psychological Research: Statistics, Methodology, and Measurement* (Allyn and Bacon, Boston, MA)

Lakoff G, Johnson M, 1980 *Metaphors We Live By* (University of Chicago Press, Chicago, IL)

McGuigan F J, 1983 *Experimental Psychology Methods of Research* (Prentice-Hall, Englewood Cliffs, NJ)

Pinkava V, 1981, "Classification in medical diagnostics: on some limitations of Q-analysis" *International Journal of Man-Machine Studies* 15 221-237

Posner M I, Keele S W, 1968, "On the genesis of abstract ideas" *Journal of Experimental Psychology* 77 353-363

Reed S K, Friedman M P, 1973, "Perceptual vs. conceptual categorization" *Memory and Cognition* 1 157-163

Rosch E, 1973, "Natural categories" *Cognitive Psychology* 4 328-350

Rosch E, Mervis C B, 1975, "Family resemblances: studies in the internal structure of categories" *Cognitive Psychology* 7 573-605

Rosch E, Mervis C B, Gray W D, Johnson D M, Boyes-Braem P, 1976, "Basic objects in natural categories" *Cognitive Psychology* 8 382-440

Schapiro M, 1961, "Style", in *Aesthetics Today* Ed. M Philipson (World Publishing, Cleveland, OH) pp 81 - 113
Shepard R N, 1974, "Representation of structure in similarity data: Problems and prospects" *Psychometrika* 39 373 - 421
Smith E E, 1989, "Concepts and induction", in *Foundations of Cognitive Science* Ed. M I Posner (MIT Press, Cambridge, MA) pp 501 - 526
Smith E E, Medin D L, 1981 *Categories and Concept* (Harvard University Press, Cambridge, MA)
Smithies K W, 1981 *Principles of Design in Architecture* (Van Nostrand Reinhold, New York)
Tversky A, 1977, "Features of similarity" *Psychological Review* 84 327 - 352