



ARCHITECTS PREDICT LAY EVALUATIONS OF LARGE CONTEMPORARY BUILDINGS: WHOSE CONCEPTUAL PROPERTIES?

GRAHAM BROWN AND ROBERT GIFFORD

University of Victoria, Canada

Abstract

Evidence suggests that architects as a group cannot predict the public's aesthetic evaluations of architecture. In this study, practicing architects predicted laypersons' responses to large contemporary building, and again these predictions were poorly correlated with ratings by laypersons, although some architects' predictions were better than others, and architects were able to predict accurately that lay ratings in general would be more favourable than their own. To understand why most architects are unable to predict reactions to particular buildings, the architects' predictions were analysed in relation to their own and lay ratings of the buildings' conceptual properties. The results suggest that architects are unable to exchange their own criteria for conceptual properties for those of laypersons when they predict public evaluations, which leads to self-anchored, inaccurate predictions. This was supported by showing that the best-predicting architects related their evaluations to buildings' conceptual properties in a manner similar to that of the laypersons. Implications for design are suggested.

Introduction

The Clock Tower building in San Francisco, designed by David Baker, has been described by some architects as fresh and innovative, and by some members of the public as an abomination. This divergence of opinion between architects and nonarchitects has long been known, but it raises an important question. Do architects, the putative experts in building function and design, understand the preferences of the general public but simply disagree with them, or do they actually not understand the laypersons's view, and are therefore unable to judge buildings as laypersons would? Addressing this question may lead to a better understanding of the reasons for these apparent discrepancies and could help reduce the number of mismatches between designer and lay preferences.

Hershberger (1969) was the first to document the differences between architects and nonarchitects, although discussions of the problem appeared earlier (e.g. Osmond, 1957). Hershberger attributed these differences to the professional education of architects. His study compared the meanings of buildings to architects, prearchitects and nonarchitects, and discovered that architects and nonarchitects differed significantly on 53 of 125 comparisons.

Numerous later studies have confirmed that differences in perceptions, cognitions, and aesthetic preferences exist between architects and other design experts and the public (e.g. Hershberger & Cass, 1974; Groat, 1982; Devlin & Nasar, 1989; Nasar, 1989; Devlin, 1990; Wilson & Canter, 1990; Stamps, 1991a; Hubbard, 1994; Purcell, 1995; Wilson, 1996), although some studies find that certain building qualities, including goodness of example and familiarity, are judged similarly by the two groups (e.g. Devlin & Nasar, 1989; Nasar & Purcell, 1990; Hubbard, 1996). These group differences in preference may be the result of predispositions, design training that alters the architecture student's understanding of symbols and introduces a specialized language (e.g. Wilson, 1996), or on-the-job experience dealing first-hand with design every working day. This apparently leads to a different set of design criteria and values (cf. Groat, 1982, 1995). In particular, architectural education seems to promote a distaste for popular styles (Purcell & Nasar, 1992).

Nevertheless, architects, as the experts in the field, should be able to understand how laypersons

think about design (Izumi, 1965). If architects wish to produce a design that pleases not only themselves and their colleagues, but also the typical person in the street (of course, not all architects do have this goal), it is essential to know the layperson's design values and preferences. But do they? In a study of more experienced architects, Nasar (1989) found that architects misgauged nonarchitects' responses in evaluating six styles of houses in terms of their symbolic meaning (e.g. perceived friendliness and desirability).

This study extends that work in five ways. First, Nasar's (1989) study focused on houses, and he called for studies of other building types; this study examines large office buildings constructed in the 1980s and 1990s. Second, Nasar's study investigated preference for different architectural styles, and it showed that style does have strong effects on observers' preferences. However, style is not the only influence on evaluations; building properties such as coherence, novelty, complexity, and order also influence preference (e.g. Wohlwill, 1974; Devlin & Nasar, 1989; Herzog, 1992). This study extends Nasar's (1989) work by examining the role of such properties on architects' predictions of lay evaluations. Third, Nasar's study used a ranking procedure which did not permit an examination of architects' ability to predict the mean level of lay evaluations across buildings; that is, whether lay and architect observers differ in their mean evaluations of an entire sample of buildings. If laypersons are more (or less) favorable toward an entire set of buildings, can architects predict that difference? Fourth, Nasar's study did not examine the differential ability of individual architects to predict lay evaluations; one presumes that some architects are better at predicting lay evaluations than others. Fifth, although the Nasar study employed style as its dependent variable, it did not consider style as a mediating variable that might help explain why architects were unable to accurately predict lay evaluations. This study examines architects' use of properties such as complexity and novelty as possible bases for their predictive inability.

Method

Participants

Thirty-three practicing architects in a mediumsized Canadian city were sent a letter that informed them of the study's purpose, and 25 agreed to participate. Those who participated were, on average, very experienced, with an excellent range of experience; they had been practicing for an average of 22.04 years, with a range of 1 to 44 years.

Data from a separate set of architects and from lay participants who rated the same set of buildings as architects and laypersons were available from another study (Gifford *et al.* in press). The lay participants (n=27) were selected from the community by randomly telephoning names from the city directory and others were introductory psychology students who received a small amount of credit towards their course grade for participating. None had any architectural training. The architects (n=8) were a separate group of practicing architects from the same city.¹

Buildings

The buildings were 42 large urban structures of diverse styles that were constructed in the 1980s and 1990s in developed countries. Colour slides of each one were made from photos in architectural journals and books. Twenty-one buildings were from the United States, 14 were from England, three were from Switzerland, and one building was located in each of Canada, Austria, and Hong Kong (see Gifford *et al.*, 2000 for some examples).

Measures

The raters were asked to use a global impression rating on a scale of 1 to 10, where 1 represented 'terrible architecture' and 10 represented 'excellent architecture'. The architects were asked to 'predict or try to mimic a typical nonarchitect's global impression of each building'. The laypersons and architects from Gifford *et al.* (in press) were asked to make this same rating themselves.

Results

Reliabilities and means of ratings

The reliabilities of the ratings were computed as intra-class correlations (Shrout & Fleiss, 1979, formula 3, k). Agreement among the architects who predicted lay ratings was 0.79, agreement among the architects judging *as* architects was 0.83, and that among the lay judges was 0.85. The agreement among the laypersons confirms earlier findings by Stamps (1991*b*) and Stamps and Nasar (1997) that quite diverse demographic groups of public raters had very similar responses to high rises and houses. Thus, the architects agree among themselves in their predictions of what laypersons would say, and the architects and laypersons who rated the same buildings as themselves also agreed strongly among themselves. Therefore, the ratings were averaged across all raters within each group to produce an overall preference rating for each building within each group.

Architects' predictions

Mean predictions. Table 1 displays the mean overall (across the 42 buildings) aesthetic ratings by the architects who predicted lay ratings, architects rating for themselves, and laypersons. On the 1 to 10 scale, these were 5.37, 4.11, and 5.55, respectively. A one-way analysis of variance showed an overall significant difference between the groups [F(2,(123) = 22.3, p < 0.001]. A Scheffé multiple range test (0.05 criterion) revealed that the ratings of the architects who rated the buildings as architects were significantly lower than the other two sets of ratings, which were not significantly different from each other. Thus, the architects who predicted layperson ratings apparently were able to adjust accurately their own less favorable ratings-assuming they would have rated the building set similarly to the separate group of architects who rated the same set of buildings *as* architects—to predict accurately the more favorable overall mean aesthetic judgments of the laypersons.

Correlated predictions. Next, the architects' predictions of lay ratings were correlated with the lay ratings on a building-by-building basis to determine the degree to which their predictions matched the lay ratings (see Table 2). The correlation was r=0.16, p>0.30. Thus, although both groups rated

 TABLE 1

 Aesthetic quality ratings by group

Group	М	S.D.
Architects as laypersons $(n=25)$	5.37	0.74
Laypersons $(n = 27)$	5.55	0.93
Architects as architects $(n=8)$	4.11	1.44

Note: The scale ranged from 1 ('terrible architecture') to 10 ('excellent architecture'). Architects as laypersons refers to a group of architects who predicted lay ratings of the 42 buildings. Laypersons refers to lay ratings of aesthetic quality. Architects as architects refers to a separate group of architects who rated aesthetic quality from their own viewpoint. All groups rated the same 42 buildings.

 TABLE 2

 Correlations between group ratings of architectural quality

	Laypersons	Architects as architects
Architects predicting	0.16	0.74
laypersons Laypersons	_	0.14

the building set as a whole at about the same level on the 10-point scale, they agreed very little as to *which* buildings were better examples of good architecture.

These architects' ratings were then compared to those of the separate group of architects who reliably rated the same 42 buildings *as* architects; that is, without trying to simulate lay preferences. This correlation was r=0.74, p<0.001, indicting very strong agreement. The implication of these two correlations is that architects as a group are unable to predict which buildings laypersons will like or dislike because they cannot evaluate the individual buildings differently than they would *as* architects.

Toward understanding the bases of the differences

In an attempt to understand why this happens, the evaluations of the architects and laypersons were related to various properties or qualities of the buildings. The goal was to connect overall aesthetic evaluations of both groups of specific building qualities, and then compare how the two groups' overall evaluations were related to these qualities. This may reveal the basis of each group's overall evaluations.

A long research tradition has sought to identify key aspects of buildings that presumably are based on purely aspects of the buildings, but represent human abstractions from physical form. Examples of such qualities are Kuller's (1980) suggestion that building designs have more or less potency or unity, Berlyne's (1972) 'collative properties' such as congruity and novelty, Nasar's (1994) 'formal variables' such as complexity and order, and Kaplan and Kaplan's (1989) 'informational factors' such as mystery and legibility. We use 'conceptual properties' as an umbrella term for all these concepts which are human cognitive constructs that people easily abstract from the purely physical features of buildings.

Based on an informal review of several such sets of conceptual properties, we selected six conceptual properties that seemed to cover much of the cognitive territory in all the sets: clarity, complexity, friendliness, originality, ruggedness, and meaningfulness. If aesthetic preferences are related to conceptual properties in general, and we know they are (*cf.* Wohlwill, 1974; Devlin & Nasar, 1989; Herzog, 1992), then a comparison of how these conceptual properties relate to aesthetic preference in both groups might reveal something about where the architects go astray in their predictions of lay evaluations.

Fortunately, ratings of the same 42 buildings' conceptual properties by other architects and laypersons (i.e. individuals not part of this investigation) were available from another study (Gifford *et al.*, in press). Each independent judge (16 architects and 9 laypersons) rated all 42 buildings on all six conceptual properties. Judges in each group substantially agreed among themselves about the buildings' conceptual qualities (inter-rater reliability ranged from 0.67 to 0.88, with a median of 0.77).

The results of the comparison may be viewed in Table 3. As can be seen, the pattern of correlations for the six conceptual properties in relation to overall aesthetic quality for the architects who were asked to predict lay judgments (column one) is very dissimilar to the equivalent set of correlations for laypersons (column two). In contrast, the pattern of correlations between the ratings of the architects who predicted the laypersons' responses and (the separate set of) architects' ratings of aesthetic quality (column three) is quite similar to that made by the architects who judged the buildings *as* architects and those of (a different set of) architects who rated the buildings' aesthetic quality (column four).

Thus, even when architects are asked to judge buildings as the public would, they rely—as a group—on conceptual properties as understood by architects. In conclusion, one root of the reason that architects cannot predict lay preferences is that they base their predictions on conceptual properties as conceived by architects, rather than conceptual properties as conceived by laypersons. If architects think of the qualities upon which aesthetic judgement is based differently than laypersons, and cannot adjust to the layperson's criteria, then it is not surprising that differences in aesthetic judgement occur.

Some architects predict better than others

The results so far refer to architects as a group. Obviously, some architects might be able to predict lay preferences better than others. Perhaps these are the more experienced among them, or perhaps it is the less experienced, who are less influenced by years of learning and working among designers. If some architects can predict better than others, then their pattern of correlations described in Table 3 and in the previous section should be closer to that of laypersons.

However, *can* some architects predict better than others? The predictions of the 25 individual architects in this study were correlated with the reliable mean aesthetic ratings of the laypersons. These ranged from r = -0.31 to r = 0.47, with a median of r = 0.08 (recall that the correlation for the pooled architect predictions was r = 0.16). Clearly, some architects predicted lay ratings more accurately than others; some of them predicted lay ratings fairly well whereas nine others' predictions.

Is experience related to accuracy? The correlation between accuracy and architects' years of experience was r = -0.18, a slight trend toward *less* experienced architects predicting lay responses better, although the nonsignificance of the correlation

Property	Architects predicting laypersons with laypersons	Laypersons with laypersons	Architects predicting laypersons with architects	Architects with architects	
Complex	0.16	0.41	0.04	0.21	
Friendly	0.16	0.47	0.41	0.43	
Rugged	-0.06	0.51	0.53	0.61	
Unique	0.19	0.51	0.47	0.62	
Clear	0.01	0.29	0.66	0.52	
Meaningful	0.00	0.71	0.69	0.51	

TABLE 3 The use of conceptual properties in relation to overall aesthetic judgment

Note: Each correlation is between a conceptual property rating and an overall aesthetic rating. In column one, architects' predictions of lay aesthetic evaluations are correlated with lay ratings of the conceptual properties. In column two, laypersons' aesthetic evaluations are correlated with lay ratings of the conceptual properties. Column three's correlations are between architects' predictions of lay aesthetic evaluations and architects' ratings of the conceptual properties. Column four shows architects' aesthetic evaluations with architects' ratings of the conceptual properties.

Architects Predictions

Conceptual property	Laypersons	Most-accurate architects			Least-accurate architects		
Complex	0.41	0.39	0.14	-0.04	-0.40	0.02	-0.16
Friendly	0.47	0.46	0.29	0.28	0.01	0.23	0.03
Rugged	0.51	0.09	0.21	0.04	-0.24	-0.12	-0.29
Original	0.51	0.47	0.09	0.10	-0.43	0.09	-0.30
Clear	0.29	-0.09	0.24	0.15	-0.04	-0.18	0.04
Meaningful	0.71	0.52	0.22	0.14	-0.49	0.04	-0.528
Accuracy (r)		0.47	0.43	0.42	-0.31	-0.18	-0.16

 $T_{ABLE \ 4}$ The use of lay conceptual properties by the three most and three least accurate architects

Note: Each correlation in the top six rows is between a rated conceptual property and judged aesthetic quality. The layperson correlations are between two independent lay groups, one that rated the conceptual properties and one that rated aesthetic quality of the 42 buildings. The architect correlations are between architects' ratings of aesthetic quality as they predicted that laypersons would rate it and conceptual properties as rated by laypersons. Accuracy is the correlation between each architect's prediction of lay ratings of aesthetic quality and the pooled lay ratings of aesthetic quality the architects were trying to predict.

requires the conclusion that experience neither helps nor hinders architects' predictive ability.

Finally, is it true that architects who predicted lay ratings more accurately used a view of conceptual properties that is similar to the layperson's view of conceptual properties? If so, it would suggest that good prediction is founded on adopting lay rather than architect criteria for conceptual properties. To check this, the use of conceptual properties by the three architects with the best predictive records (rs = 0.47, 0.43, and 0.42 with lay judgements of aesthetic quality) were compared with the use of conceptual properties by the three architects with the worst predictive records (rs = -0.31, -0.18, and -0.16). The results of this are displayed in Table 4.

As inspection of the table shows, the more accurate architects' use of conceptual properties resembles the lay use of conceptual properties much more than the least accurate architects, use. This is not to say the resemblance is outstanding; after all, no architect was more accurate than r = 0.47. But at least the more accurate architects' correlations with aesthetic quality are generally positive, like those of the laypersons, whereas those of the least accurate architects are often negative, which indicates that the conceptual property was used in the reverse direction from the laypersons. For example, all three of the least accurate architects seemed to believe that lay evaluations would be higher for less rugged buildings, but the lay evaluations actually were higher for *more* rugged buildings.

Thus, a key clue to the generally poor ability of architects to predict public evaluations of large, contemporary buildings is the failure to understand the manner in which the public mind links conceptual properties of buildings to aesthetic evaluations. For example, the most accurate architect seems to understand that laypersons judge buildings that are more complex, rugged, and original more positively, whereas the least accurate architect believes that the public prefers buildings which are *less* complex, rugged, and original.

Some examples and some hope

Some concrete illustrations may be helpful. A building that fits the general trend of these results is the Disney Headquarters Building in Burbank, CA, a Michael Graves design. The architects themselves gave it a very low rating (2-44 on the 1 to 10 scale), and predicted that laypersons would not be very pleased, either (2-44). However, the public liked the building very much (7-26). In contrast, a building the architects thought people would like (Stockley Park in London, by Foster Associates), with a predicted liking of 6-28, was among the buildings least liked by the public (4-74). If all the predictions showed this pattern, the outlook would be very gloomy.

Fortunately, across all 42 buildings there were exceptions, which might profitably be studied. For example, the architects predicted that the public would not be excited by the Chicago Bar Association Building in Chicago, by Tigerman and McCurry (predicted rating 4.16), and they were correct; the public's rating was 4.56.² Incidentally, the architect's own rating of this building was also low (2.44), so the two groups agreed that this is not a pleasing building. Furthermore, the architects predicted that the public would like the Bank of China Tower in Hong Kong, by I. M. Pei and Partners (predicted rating 6.24), and they were correct (actual rating 7.67). The architects' own rating of the tower

was 7.33, so all were agreed that the tower is a very pleasing building.

Discussion

As expected, architects were unable as a group to predict which large 1980s and 1990s buildings the public would evaluate positively or negatively. They were able to predict that the public would give the set of buildings as a whole higher ratings than would the average architect. The median architect accuracy correlation was about 0.10, but some architects predicted the public's evaluations better (up to r=0.47) and others were worse (down to r=-0.31). However, the main purpose of this study was to examine possible reasons for the architects' inability to predict, with the longer-term goal of improving their understanding of the public taste.

The evidence suggests that an important reason for this pattern is that architects employ conceptual properties as architects do when they try to predict lay preferences, instead of thinking of conceptual properties as laypersons do. This does not mean that architects and laypersons always think differently. For example, Nasar and Purcell (1990) showed in a study of house styles that architects and laypersons thought about buildings similarly in terms of their goodness of example and familiarity. But this study suggests that the same building has different conceptual properties for different groups, even though members within a given group may agree to a high degree. This reiterates the point that conceptual properties are human constructs, even though they are derived from objective, physical building features. The study shows that those who would hope to predict the aesthetic responses of others should use the conceptual properties in the same manner as the group they hope to predict.

Applied to architecture, this certainly does not imply that architects should develop their predictive skills in order to design buildings solely to the public taste. The purpose of the profession, in part, is to creatively advance design. As the data from this study shows, even the public positively values originality in building design, and untrained laypersons are not likely to create designs that work. Nevertheless, it is possible to create designs that both architects and laypersons like (the Bank of China Tower, Figure 1, is a good example). Post-modernism in general had this as a goal, although postmodern buildings do not always fulfill the goal. If pleasing both groups is possible, there seems little point in designing a building to please either group alone. In order to understand what laypersons like and how they evaluate buildings, architects may wish to learn more about conceptual properties in the eyes of laypersons. A good direction for future research is to examine in greater depth just which physical, or formal, properties in a building lead to which conceptual properties in the eyes of architects and laypersons. This would lead the way toward specifying conceptual properties in physical terms, and would facilitate an objective grammar of design which would in turn produce more buildings that please both architects and the public.

Notes

Correspondence and reprint requests should be addressed to: Robert Gifford, Department of Psychology, Box 3050, University of Victoria, Victoria, BC Canada V8W 3P5. Tel: 250-721-7532; Fax: 250-721-8929; E-mail: rgifford@ uvic.ca

(1) The number of architects judging *as* architects (8) may seem small but, if subjects are chosen more or less at random from the population in question, which they were, then high inter-rater agreement (0.83 in this instance) strongly suggests that adding more architects would be redundant because the results would be the same (*cf.* Guilford, 1954). The same is true for the other rating groups.

(2) Lest the reader fret that perhaps building ratings were influenced by such knowledge, for example that this building is a professional home of lawyers, we wish to clarify that no raters were informed of any building's owners, tenants, or architects during the rating exercise.

References

- Berlyne, D. E. (1972). Aesthetics and Psychobiology. New York: Appleton-Century-Crofts.
- Devlin, K. (1990). An examination of architectural interpretation: architects versus non-architects. Journal of Architectural and Planning Research, 7, 235–244.
- Devlin, K. & Nasar, J. L. (1989). The beauty and the beast: some preliminary comparisons of 'high' versus 'popular' residential architecture and public versus architect judgments of same. *Journal of Environmental Psychology*, 9, 333–334.
- Gifford, R., Hine, D. W., Muller-Clem, W., Reynolds, Jr., D. J. & Shaw, K. (2000). Decoding modern architecture: A lens model approach for understanding the aesthetic differences of architects and laypersons. *Environment and Behavior*, **32**, 163–187.
- Gifford, R., Hine, D. W., Muller-Clemm, W. & Shaw, K. T. (in press). Why architects and laypersons judge buildings differently: cognitive and physical bases. *Journal* of Architectural and Planning Research.
- Groat, L. (1982). Meaning in post-modern architecture: an examination using the multiple sorting task. *Journal* of Environmental Psychology, **2**, 3–22.

- Groat, L. (1995). Introduction: place, aesthetic evaluation, and home. In L. Groat, (Ed.), *Giving Places Meaning*. New York: Academic Press, pp. 1–26.
- Guilford, J. P. (1954). Psychometric Methods. New York: McGraw-Hill.
- Hershberger, R. G. (1969). A study of meaning and architecture. In J. L. Nasar, (Ed), *Environmental Aesthetics: Theory, Research, and Application.* New York: Cambridge University Press, pp. 175–194.
- Hershberger, R. G. & Cass, R. (1974). Predicting user responses to buildings. In J. L. Nasar, (Ed.), *Environmen*tal Aesthetics: Theory, Research, and Applications. New York: Cambridge University Press, pp. 195–211.
- Herzog, T. R. (1992). A cognitive analysis of preference for urban spaces. *Journal of Environmental Psychology*, 12, 237–248.
- Hubbard, P. (1994). Professional vs lay tastes in design control—an empirical investigation. *Planning Practice* and Research, 9, 271–287.
- Hubbard, P. (1996). Conflicting interpretation of architecture: an empirical investigation. Journal of Environmental Psychology, 16, 75–92.
- Izumi, K. (1965). Psycho social phenomena and building design. Building Research, 2, 9–11.
- Kaplan, R. & Kaplan, S. (1989). The Experience of Nature: A Psychological Perspective. New York: Cambridge University Press.
- Kuller, R. (1980). Architecture and emotions. In B. Mikellides, (Ed.), Architecture and People. London: Studio Vista.
- Nasar, J. L. (1989). Symbolic meaning of house styles. *Environment and Behavior*, 21, 235–257.
- Nasar, J. L. (1994). Urban design aesthetics: the evaluative qualities of building exteriors. *Environment and Behavior*, 26, 377–401.
- Nasar, J. L. & Purcell, T. (1990). Beauty and the beast extended: knowledge structure and evaluations of house by Australian architects and nonarchitects. In H.

Pamir, V. Imamoglu & N. Teymur, (Eds), *Culture, Space, History*. Ankara, Turkey: Sevki Vanh Foundation.

- Osmond, H. (1957). Function as the basis of psychiatric ward design, *Mental Hospitals*, 8, 23–30.
- Purcell, T. (1995). Experiencing American and Australian high- and popular-style houses. *Environment and Be*havior, 27, 771–800.
- Purcell, T. & Nasar, J. L. (1992). Experiencing other people's houses: a model of similarities and differences in environmental experience. *Journal of Environmental Psychology*, **12**, 199–211.
- Shrout, P. E. & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*, 86, 420–428.
- Stamps, A. E. (1991a). Comparing preferences of neighbors and a neighborhood design review board. *Environ*ment and Behavior, 23, 618–629.
- Stamps, A. E. (1991b). Public preferences for high rise buildings. stylistic and demographic effects. *Percep*tual and Motor Skills, 72, 839–844.
- Stamps, A. E. & Nasar, J. L. (1997). Design review and public preferences: effects of geographical location, public consensus, sensation seeking, and architectural styles. *Journal of Environmental Psychology*, 17, 11–32.
- Wilson, M. A. (1996). The socialization of architectural preference. Journal of Environmental Psychology, 16, 33–44.
- Wilson, M. A. & Canter, D. V. (1990). The development of central concepts during professional training. An example of a multivariate model of the concept of architectural style. *Applied Psychology: An International Review*, **39**, 431–455.
- Wohlwill, J. (1974). The Place of Aesthetics in the Study of the Environment. Paper presented at the Symposium on Experimental Aesthetics and Psychology of the Environment at the International Congress of Applied Psychology, Montreal.