Investigating visual preferences: a structured comparison approach

William Fawcett

This paper was published in *Hong Kong Papers in Design and Development*, published by the Department of Architecture, University of Hong Kong.


This is the author’s preprint. Not for circulation. You are advised to refer to the publisher’s version if you wish to cite the publication.

william.fawcett@carltd.com
Investigating visual preferences:  
a structured comparison approach

Published in *Hong Kong Papers in Design & Development*, vol.1, 1998

William Fawcett MA PhD RIBA  
Cambridge Architectural Research Ltd  
The Eden Centre  
47 City Road  
Cambridge CB1 1DP  
UK  
tel +44 1223 460475  fax +44 1223 464142  
e-mail wfawcett@carltd.com

**SUMMARY**
A method of investigating the visual preferences of individuals or groups of individuals is described, which relies on the exhaustive comparison of pairs of images representing different design options. The comparison can be carried out using an interactive computer questionnaire. By analysing the participants’ responses to the images, the underlying structure of their visual preferences can be revealed. The method is applied in a case study which investigated visual preferences for low-rise office buildings in the UK.

**KEYWORDS**
Visual preference, appearance, environmental assessment, environmental aesthetics, aesthetic judgement, user requirements.

1 **INTRODUCTION**

The visual appearance of buildings is one of many design parameters that architects have to deal with, but it is perhaps the single parameter that impacts most widely on other people. Many groups – the building’s users, other architects, the general public – are likely to see a new building and, however fleetingly, judge its appearance. No doubt the architect would like these judgements to be positive, but we know that in matters of visual taste different people react quite differently. So, what criteria can the architect use when making decisions about visual appearance?
A critical issue is what audience the architect wants to design for, but in practice there is minimal information about other peoples' preferences; so the architect usually relies on his or her own personal values, responds to the equally personal reactions of a client, or falls in with conventional prejudices about 'popular taste'. The purpose of this paper is to put forward a more constructive approach, based on the systematic investigation of the visual preferences of a wide range of people. How much use, if any, architects would make of new information of this kind is of course another question entirely.

The investigation of visual preferences is part of the much broader topic of environment-behaviour research (Zeisel 1984), where it can be seen as one aspect of environmental assessment (Craik & Feimer 1987). A review article on visual preferences referred to 73 publications which offer a wide variety of approaches, most of them based on the evaluation of images of buildings or landscapes (Nasar 1994). In these studies, the preferences of the participants are revealed by their evaluations of the images. The choice of images depends on the context of each particular study, for example, houses (Devlin & Nasar 1989), churches (Purcell 1984), urban design (Herzog & Gale 1996), landscape (Kaplan & Kaplan 1989), and so on.

In this paper the main objective is to present a method of research rather than the findings of studies carried out using it, although these are indicated in a case study. The following sections explain the method and use a case study to illustrate it. Some further developments of the method are also explained in the context of the case study.

2 METHOD

2.1 Classifying design options

The first stage in the method is to classify the design options in the field of interest: the method is more suitable for relatively detailed investigations of a well-defined range of options, rather than broad and open-ended studies. The classification should cover the whole spectrum of design options of interest; and within this spectrum the options should be chosen so that every imaginable design corresponds to one option or another.
Setting up the classification of design options must be done by the research team before the investigation begins, and necessarily constrains the whole of the subsequent investigation. If important options are missing from the classification, the research cannot say anything at all about them. A draft classification should normally be drawn up and widely discussed and evaluated before being finalised.

Images must then be selected to represent the design options in the classification. The images are usually photographs of completed buildings – each photo is allocated to one or another of the design options. If any photos cannot be classified it suggests that the classification may need to be modified.

In practice the collection of images may come first, with the classification deriving from an analysis of the images.

An diagrammatic set of four design options for the appearance of tower blocks is shown in the box.

2.2 Comparing design options

Having established the design options, the method is based on comparing every option with every other option, in pairs of images. This exhaustive approach limits the number of options that can realistically be handled. Suppose there are four options, W, V, H and G; then we need to compare six pairs, WV, WH, WG, VH, VG and HG. The formula for n options is \( \frac{n \times (n - 1)}{2} \) pairs. The following table indicates how rapidly the number of pairs rises with more design options:

<table>
<thead>
<tr>
<th>number of design options</th>
<th>number of pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td>8</td>
<td>28</td>
</tr>
<tr>
<td>10</td>
<td>45</td>
</tr>
<tr>
<td>15</td>
<td>105</td>
</tr>
<tr>
<td>25</td>
<td>300</td>
</tr>
<tr>
<td>50</td>
<td>1225</td>
</tr>
<tr>
<td>100</td>
<td>4950</td>
</tr>
</tbody>
</table>

In practice it may be hard to cope with more than 8 design options, and 28 pairs.
The set of image pairs are now presented to the people whose preferences are being investigated. For each pair of images in turn, the participants are asked to indicate which of the two images they prefer. In the studies we carried out, we offered the choice of a ‘weak’ or a ‘strong’ preference for one image or the other – that is, four choices. We did not offer a ‘don’t know’ option, forcing people to make a decision rather than giving them an easy opt-out. The addition of a ‘don’t know’ choice would be quite feasible within the research method.

The results of each paired comparison have to be scored. When one image in a pair is ‘strongly’ preferred we attach the score of +2, and −2 to the other image which is by implication rejected; and +1 to an image which is ‘weakly’ preferred and −1 to the one which is rejected. The actual numbers are not critical, so long as they record the direction of preferences.

A set of image pairs for the diagrammatic tower block design options, and typical selections by an individual, are shown in the box.

2.3 Deriving the structure of preferences

When the set of image pairs have all been evaluated by a given individual, the structure of that individual’s preferences can be analysed. The response to a design option can be quantified by counting all the scores attached to images which show that option. If the aggregate score is a large positive number, then that option was generally preferred, either strongly or weakly, over other design options; conversely, a design option with a large negative number was generally rejected. A design option with a zero or very low score (positive or negative) was preferred or rejected more or less equally often – it is neither very much liked nor disliked.

Then the design options can be ranked according to their scores. If the options are widely spaced, then it shows that the individual has a clear and decisive preference structure with respect to the design options; if, on the other hand, the options are bunched around the zero mark then the individual is more or less indifferent with respect to the design options.

The ranking of the diagrammatic tower block design options, based on the typical image selections, is shown in the box.
The preference structures of a number of individuals can be aggregated and ranked to arrive at the average view of the group. A widely spaced pattern of preferences indicates that the group has a clear and decisive consensus view; if the options are bunched at the zero mark then all individuals in the group are indifferent, or their individual preferences vary so that there is no group consensus.

Note that the preferences of an individual are usually somewhat chaotic. Suppose that someone prefers option A to B, and B to C: we cannot assume that A will be preferred to C. When many design options are being compared the hierarchy of preferences can get very tangled. Partly this is because the judgements are made on the basis of images of particular buildings, and may be affected by extraneous features in the images; but a degree of circularity is often a true representation of preferences.

2.4 Collecting preference data

To study the visual preferences of a group of people the paired comparison tests have to be carried out many times with many individuals. The mechanics of presenting the images and recording preferences become cumbersome. For this reason we decided to develop an interactive computer questionnaire.

The questionnaire comprised a number of ‘cards’ which were designed to be both clear and attractive. After introductory cards, the pairs of colour scanned images were presented on successive cards, and the participants indicated their preferences by using the computer mouse to click on one of four on-screen buttons indicating strong or weak preference for one image or the other. When the selection had been made for one image pair the data was automatically logged by the computer and the next image pair was automatically displayed. No time limit was imposed on the participants; in our research the average time taken for the selection from an image pair was about 10 seconds – reflecting a conscientious response from the participants.

The software for the questionnaire was written using the scripting language and display tools in the SuperCard environment. Results were ported to Excel for data analysis.

This interactive computer-based technique for data collection was found to be extremely successful. There were a number of advantages compared to paper-based questionnaires:
Investigating visual preferences

- participants were involved and interested in the questionnaire
- participants who began the questionnaire always finished it
- interactive prompts were used to minimise inconsistent responses
- data was automatically logged for analysis, saving time and avoiding transcription errors.

The data which is logged automatically in the questionnaire can be analysed 'by hand' in Excel or other environments, but if data analysis requirements were standardised they could also be automated and linked to the questionnaire program.
3 CASE STUDY

3.1 Context

The method set out above was used to investigate the design of new low-rise office buildings in the UK. The project compared the views and preferences of users with members of development teams for these projects. The members of the development team were:

- **architect**: designs the development in accordance with the brief
- **planning consultant**: advises in cases where planning consent is controversial
- **estate agent**: advises on the marketability of offices; markets the development when it is completed
- **developer**: identifies the project; responsible for cashflow until a tenant is found
- **investor**: buys the development when a tenant is found

Other aspects of this research project were reported elsewhere (Fawcett 1992).

3.2 Design options

It was hypothesised that visual preferences are based on the presence or absence of important design attributes, for example, materials, colours, shapes, textures, etc. Many photographs of actual office buildings were sorted in various ways, and eventually a simple classification scheme of design options was adopted. It was based on three attributes, each with two possible values:

<table>
<thead>
<tr>
<th>attribute</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>roof shape</td>
<td>flat or pitched</td>
</tr>
<tr>
<td>walling material</td>
<td>traditional or non-traditional</td>
</tr>
<tr>
<td>architectural character</td>
<td>strong or weak</td>
</tr>
</tbody>
</table>

The first two attributes are self-explanatory, but the third requires explanation. Preferences were felt to be positively affected when the elements of the design are handled skilfully, and negatively affected by clumsy designs, independently of the wall material or roof pitch. Each image was therefore categorised as exhibiting a strong or a weak architectural character. This categorisation was subjective and was carried out by the researchers who came from an architectural background.
By combining the attributes and values in all possible ways, this classification scheme leads to eight different design options for low-rise office buildings:

<table>
<thead>
<tr>
<th>ROOF SHAPE</th>
<th>WALL MATERIAL</th>
<th>ARCHITECTURAL CHARACTER</th>
</tr>
</thead>
<tbody>
<tr>
<td>flat</td>
<td>non-traditional</td>
<td>strong</td>
</tr>
<tr>
<td>flat</td>
<td>non-traditional</td>
<td>weak</td>
</tr>
<tr>
<td>pitched</td>
<td>non-traditional</td>
<td>strong</td>
</tr>
<tr>
<td>pitched</td>
<td>non-traditional</td>
<td>weak</td>
</tr>
<tr>
<td>flat</td>
<td>traditional</td>
<td>strong</td>
</tr>
<tr>
<td>flat</td>
<td>traditional</td>
<td>weak</td>
</tr>
<tr>
<td>pitched</td>
<td>traditional</td>
<td>strong</td>
</tr>
<tr>
<td>pitched</td>
<td>traditional</td>
<td>weak</td>
</tr>
</tbody>
</table>

Photographs were organised into pairs of design types, for all possible pairings – every design option paired just once against every other option. Photos were not re-used, so a total of fifty-six were required (eight design options needing seven photos each), giving 28 pairs. Photos were paired so that images of approximately equal photogenic appeal were together, eg. images in sunshine or with attractive landscaping were not paired with images taken in dull days or lacking in landscape interest. The selection of photographs corresponding to permutations of attribute values is an established technique (for example, Herzog & Gale 1996).

Two refinements may be mentioned, although they were not used in the case study project. First, the sequence in which the image pairs are presented to participants could be varied, in case of a bias in selections between the images presented early on and those presented at the end. The computer questionnaire could, in principle, randomise the sequence between each participant. Secondly, the handing of images on the cards could be swapped, in case of a bias in selecting images on the right hand side or left hand side of the screen. The whole set could be exchanged left-to-
right between each participant. The database would be able to track the results regardless of these transformations.

3.3 Data collection

Data about preferences was obtained by presenting the pairs of images in an interactive computer questionnaire, as described above. The number of participants in each group was as follows:

- Architects: 31
- Planning consultants: 11
- Estate agents: 15
- Developers: 12
- Investors: 7
- Users: 93

TOTAL: 169

3.4 Ranking of design options

From the questionnaire selections, each individual’s ranking of the eight design options was derived, using the scoring scheme outlined above (+2 or +1 for each image that was strongly or weakly preferred, and −2 or −1 for each one that was rejected). The scores were averaged for the participants in each of the groups, and the group rankings are shown in the diagram.

*** INSERT DIAGRAM ONE *** (see end)

There are striking differences between the groups. Note, for example:

- architects had the widest divergence between their most favoured and least favoured selections, that is, they were most decisive; developers and users were least decisive

- users showed a preference for pitched roof buildings over flat roofed ones

- the two design options selected most strongly by architects received low scores from all other professions and users
• the options selected strongly by the other professions and users received low scores from architects.

Further analysis can be carried out to follow up these observations: first, the degree of consensus within groups and divergence between groups; and secondly, the importance of the three design attributes (roof shape, walling material, architectural character).

3.5 Group consensus/divergence

Every participant in the survey made selections for the same set of image pairs. Now, if two participants’ selections were in total agreement, the difference between their scores recorded in the database would be zero. If they disagreed for an image pair their scores would be different, by as much as 4 if an image was selected strongly (+2) by one person but rejected strongly (−2) by the other. If the differences in scores are added for all 28 image pairs we arrive at a measure of the amount of agreement or disagreement; the higher this number the greater the disagreement between the two participants.

If both participants made selections at random the measure would be 49. A number below 49 indicates that the individuals have some agreement (lower numbers mean more agreement); a number higher than 49 indicates definite disagreement.

This analysis was carried out for all pairs of participants (14,196 pairs for 163 participants). For pairs of individuals within the same groups the average measures of agreement/disagreement were as follows:

<table>
<thead>
<tr>
<th>group</th>
<th>average measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Architects</td>
<td>33.6</td>
</tr>
<tr>
<td>Planning consultants</td>
<td>27.1</td>
</tr>
<tr>
<td>Estate agents</td>
<td>33.9</td>
</tr>
<tr>
<td>Developers</td>
<td>31.3</td>
</tr>
<tr>
<td>Investors</td>
<td>32.7</td>
</tr>
<tr>
<td>Users</td>
<td>35.2</td>
</tr>
</tbody>
</table>

As can be seen, users had the weakest consensus and planning consultants the strongest (bear in mind the unequal sizes of the groups).
A similar exercise can be carried out to establish the between-group measure of agreement/disagreement. The results are shown on the histogram:

*** INSERT DIAGRAM TWO *** (see end)

The architects showed the highest average measures of disagreement with other groups, both other professions and users.

3.6 Impact of design criteria

As well as ranking the eight design types, it is possible to identify which of the three attributes carried most weight. Recall that the three attributes each had two values:

- roof shape: flat or pitched
- walling material: traditional or non-traditional
- architectural character: strong or weak

Of the images used in the survey, half had pitched roofs and half had flat roofs; half had traditional walling and half had non-traditional walling; and half had strong and half weak architectural character. Each attribute value can be scored by adding the scores of all the images in which it is present.

The difference between the scores of the alternative values for a given attribute indicates whether that attribute had a major impact on preferences. For example, there would be a wide difference if all pitched roof images are strongly selected (high positive score) and all flat roof images are strongly rejected (high negative score). This would indicate that roof shape was an important factor in visual preference. If the difference between the values for an attribute is low, then images with one of the values were selected about as often as images with the other, and therefore that attribute is not a very important factor.

The relative weight attached to the three attributes can be established for the different participant groups, as shown in the pie charts:

*** INSERT DIAGRAM THREE *** (see end)
The results show that:

- roof shape was the major factor for users and all professions except architects and developers; it is most important of all for users

- roof shape was much less important for architects and developers, who are the only groups responding significantly to architectural character

- the wall material was the aspect which varies least in its relative weight across the different groups of participants.

4 CONCLUSIONS

4.1 Research method

The research method offers an effective way of collecting preference data about the appearance of buildings and other features of the environment. It could be used to evaluate architectural styles, as in the case study, or interiors, landscapes designs, etc. With synthesised images it could be used to establish reactions to design alternatives for a new project.

Because the data is structured according to a classification scheme, it is possible to infer the participants' preference structures, which they are probably not aware of themselves. And because the method is based on the selection of images it avoids any reliance on verbal descriptions of designs or responses to designs. This eliminates one possible source of uncertainty. 'Architects and non-architects use words differently and the differences may be indicative of potential communication problems between the architect and client' (Devlin & Nasar 1989).

However, although the method can say a lot about participants' preferences, the significance of the classification scheme of design options must not be underestimated. Any classification scheme imposes a point of view. This is well illustrated in the case study by the attribute 'architectural character'. The results showed that it meant 'attractive to architects' (and largely irrelevant to non-architects), but shed no light on what gives a building a strong 'architectural character'. Of course, the method could be used in further studies to investigate this interesting point.
4.2 Is there a problem?

The case study investigation of low-rise office buildings in the UK revealed clear differences between the visual preferences of architects, other construction industry professionals, and building users. The preferences of the general public were not sought in the case study project. This isolation of designers confirms a pattern that has often been observed (Nasar 1988 and 1994).

Architects’ distinctive attitudes may derive from shared values acquired in their education. Berlyne (1971:256) pointed out that ‘the [aesthetic] judgements of those who have had specialised training in art reflect intensive exposure to works of art and also intensive exposure to the attitudes that have for historical reasons become implanted in art schools’. A recent study in the UK has shown the dramatic changes in visual preferences of architectural students at different stages of study (Wilson 1996).

Whatever its origin, the mismatch between the visual preferences of those who design buildings and those who ‘consume’ them is well established from previous research. The results of further studies, for example of different building types, in different cultures, or with different participant groups, would add to our overall understanding of the phenomenon.

But more specifically, how might studies of visual preferences be used by designers? I suggest that designers could adopt three possible reactions (there is a fourth, which the method presented here is meant to undermine, that visual preferences are so profound and mysterious that nothing useful can be said or done). The three reasonable reactions are to:

- regard the mismatch as acceptable (arguing, perhaps, that designers are leaders of taste, or more knowledgeable about matters of taste, etc)
- try designing to match the preferences of groups other than architects
- try to educate the taste of non-architects.

The research method is well-adapted for use with efforts in the latter two options, by testing the success of architects who consciously design to achieve positive user or
public reactions, or by investigating the taste of users who have been exposed to architectural enlightenment.

It is hoped that the straightforward research method that has been outlined, especially when coupled with the convenience of interactive computer questionnaires, will encourage more architects and researchers to investigate the fascinating and important topic of visual preferences in architecture.

5 REFERENCES


BOX

VISUAL PREFERENCES STEP-BY-STEP

Suppose we want to investigate how members of the public react to different tower block styles. The first step is to ...

1. Classify design options

We decide that the range of alternative elevations can be represented by four design options. Remember that the investigation will not tell us anything about designs which cannot by classified into these four design options.

- window-in-wall (W)
- vertical emphasis (V)
- horizontal emphasis (H)
- glass skin (G)

We go on to ...

2. Select image pairs

We collect large set of images of tower blocks – usually colour photographs – and classify them into the four options. Then we select pairs of images which show every possible combination of options, six pairs in this case. A typical set of pairs is shown diagrammatically, but actually the images would be real photos of different buildings (images are not repeated).

These pairs are presented to the people who are being surveyed so that they can...
3. Indicate preferences

Each participant is shown the image pairs in turn and, for each pair, decides that the left hand or right hand image is preferable, either strongly or weakly – and indicates this choice. This shown by the highlighted ‘buttons’ under the images.

<table>
<thead>
<tr>
<th>LH image</th>
<th>RH image</th>
</tr>
</thead>
<tbody>
<tr>
<td>![Image 1]</td>
<td>+2</td>
</tr>
<tr>
<td>![Image 2]</td>
<td>-2</td>
</tr>
<tr>
<td>![Image 3]</td>
<td>+1</td>
</tr>
<tr>
<td>![Image 4]</td>
<td>-1</td>
</tr>
<tr>
<td>![Image 5]</td>
<td>-2</td>
</tr>
<tr>
<td>![Image 6]</td>
<td>+2</td>
</tr>
<tr>
<td>![Image 7]</td>
<td>+1</td>
</tr>
<tr>
<td>![Image 8]</td>
<td>-1</td>
</tr>
<tr>
<td>![Image 9]</td>
<td>-1</td>
</tr>
<tr>
<td>![Image 10]</td>
<td>+1</td>
</tr>
</tbody>
</table>

Then the participant leaves and the researchers move on to the ...

4. Scoring of results

The participant’s selections are scored. Every image that is selected strongly is given a score of +2, and the ones that were rejected are given scores of -2. Weakly selected or rejected images are given scores of +1 and -1. In our example these scores are shown beside the image pairs. With these scores it is possible to establish the participant’s ...
5. Preference structure

Each design option appears in the image pairs three times, so it has three scores. These are added, indicating the participant's ranking of the four design options. The ranking are shown diagrammatically on a 'ladder'. It can also be compared with other participants' rankings to investigate the preference structures of groups of individuals, for example young people, old people, men, women, professional people, etc.
Divergences in image selections

- **U** = users
- **A** = architects
- **P** = planning consultants
- **E** = estate agents
- **D** = developers
- **I** = investors