Outside the Frame: A Critical Analysis of Urban Image Surveys

Phoebe Crisman

Aided by digital technology, designers and planners are making ever greater use of photographic images to depict and envision proposals for urban change. Yet, even as such presentations have become increasingly central to community planning efforts, the methods for employing photographs and their precise construction have not been adequately analyzed.

Discussing the new reliance on imaging techniques in architecture, Cambridge Professor Andrew Saint recently noted, "the long-term challenge for the architectural profession…is to ride this exciting, undisciplined, licentious, and dangerous beast, to control this irresponsible lust for image that pervades our culture."[1]

Several imaging methods have gained widespread currency in the design of public places. This essay specifically examines what have become known as urban image surveys. The inclusive and participatory intentions of these methods are laudable, as is their emphasis on well-designed physical environments. But the predominance of image over bodily experience; the exclusion of intertwined socioeconomic, historical and political specificity; and the commodification of place they engender all raise serious concerns.

Above: +6.1 This mixed-use boulevard from Denver was the highest-rated image in the street category of the survey.

Technique for Community Engagement

Perhaps the most widespread image-survey methodology today is the Visual Preference Survey (VPS), trademarked by Anton Nelessen and first tested in 1979.[2] Residents of a town or neighborhood are shown a fast-paced sequence of photographic images and asked to rate each numerically as "acceptable" or "unacceptable." Tabulated and analyzed survey results lead to the attachment of a calculated value for each photograph, while an optional questionnaire elicits marketing and demographic information and written comments. According to Nelessen, the process is intended to "articulate the residents' impression of the present community image and build consensus for its future character. The

conclusion of the process is called a Vision Plan."[3]

A second popular methodology, Community Image Survey (CIS) has been developed by the Center for Livable Communities (CLS). Their website describes it as follows:

> *The Community Image Survey consists of forty slides from a community or region. Approximately eighty percent of the slides come from the specific locale in which the survey is administered. Taken as a whole, the forty slides present contrasting images of our living environment—its streets, houses, stores, office buildings, parks, open space and key civic features.*[4]

Numbers on how many such image surveys have been conducted are unavailable, but A. Nelessen Associates has estimated that approximately 50,000 people had filled out VPS sheets by 1994. It is likewise difficult to determine precisely when the desire to quantify image "appropriateness" or "likeability" began, since several practitioners developed processes for evaluating visual appearance simultaneously.[5]

A key intent of image surveys, however, has been to facilitate community involvement in placemaking. According to the Center for Livable Communities:

> *The Community Image Survey is a powerful planning and public participation tool that can help decision-makers and their constituents. Rather than using words to describe places, the Survey uses visual images to help people better understand crucial planning elements and make more informed, proactive decisions about creating places where they want to live, work, shop and play.*[6]

As a tool for community involvement, however, the trouble with such surveys lies in the definition of what constitutes a place. Sense of place incorporates a range of engaged bodily experiences, not merely passive appreciation of visual imagery. Image-based approaches to urban planning thus run the risk of fostering an inattentiveness to—and subsequent undervaluing of—socioeconomic, historical and political realities.

Aristotle and Heidegger once bestowed place with important symbolic and political values representing the structure of social relations, or the *res publica*.[7] How can one truly understand and rate an urban place without knowing more about what lies outside the frame—realities that cannot be captured by the camera?

The predominance of vision creates other problems. Understanding place involves multisensory experience in time. Describing the impact of contemporary ocularcentrism, Juhani Pallasmaa has argued that the senses are not independent, but interactive and synergetic.[8] He has also suggested that peripheral vision may be as important to spatial experience as focused vision.[9] By contrast, texture, sound, weight, and bodily measure all disappear from image surveys, resulting in a severely limited perception of place.

While photo surveys thus aspire to engage "outsiders"—especially those unfamiliar with more technically demanding methods of environmental evaluation—their worthy intent may compromise their basis as "science." As decontextualized images, they may simply leave too much outside the frame.

A Case Study

While a quick Web search yields hundreds of examples of urban image surveys, this essay examines one case with which the author has first-hand experience. The Milwaukee Downtown Plan, prepared for the City of Milwaukee by A. Nelessen Associates in 1999, employed a "Public Visioning Process" based on a Visual Preference Survey.

Over a period of three and one-half months, more than 1,600 Milwaukee residents viewed a rapidly timed sequence of 230 simulated images and carefully framed photographs of existing local, national and international urban conditions, including scenes from Venice, Denver, Boston and Sacramento. Using a computer scan sheet, participants numerically rated each image from +10 to -10 as "appropriate and acceptable or inappropriate and unacceptable" for their city. In a subsequent "Visual Translation Workshop" conducted over three days, participants were then asked to locate appropriate sites for applying the positive VPS images.

On reflection, it became apparent that this process had several critical weaknesses. Most obviously, apart from a brief "Demographic, Market and Policy Questionnaire," participants were given little context for understanding broader socioeconomic, historical, racial, cultural or political conditions beneath the images' surface—especially those showing scenes from places other than Milwaukee.

But the image-survey methodology also did not explicitly acknowledge that *all* photographs—whether family snapshots or "fine art"—are actively configured constructions of the referent (the thing to which they refer), and so can never be evaluated as stand-ins for absent realities. Instead, the carefully framed image-survey scenes deliberately sought to establish an apparently unmediated tableau—a city of isolated, picturesque moments without the often harsh contrasts and complex continuities that occur in every city.

The potential effect of such photographic artifice became evident to me when I examined a CIS survey prepared for Georgetown, Texas. The survey happened to contain a highly rated photograph from my own city, Charlottesville, Virginia.[10] By carefully excluding a large expanse of pavement and an immense hospital across the street, the image depicted a place very different than that I experience every day at lunchtime—a place free of busy traffic and car exhaust, the rumble of an elevated railway line, and the irregular urban fabric across the street.

This single example suggests the need to grasp more fully the power and responsibilities that accompany the use of photographic images. By analyzing images from the Milwaukee Visual Preference Survey through concepts of formal structuring, translation, meaning, and the relationship between caption and photograph, it is possible to see how a photograph projects a particular range of readings and conceptual content. In other words, although image-survey photographs seem to depict an unmediated "this is," they are inevitably constructed to convey "this means."

Cropping and Viewpoint

Cropping, the primary and perhaps most crucial formal manipulation used in the making of images, is often taken for granted by the viewer. In fact, the world that has not been included is essential to the construction of any photograph. Nevertheless, the carefully framed and formally simple street scenes used in the Milwaukee study appear unmediated and closed. For example, in the boulevard photograph shown above left, the planted median is tightly cropped to downplay its location in the middle of a busy six-lane arterial roadway.

Likewise, most VPS and CIS images are cropped in ways that present whole objects and places—and seemingly whole ideas. Like advertisements, they lull the viewer into abjuring critical consideration of complex physical, social and economic realities.

According to Stanley Cavell, "When a photograph is cropped, the rest of the world is cut *out*. The implied presence of the rest of the world, and its explicit rejection are as essential in the experience of a photograph as what it explicitly presents."[11]

The photographer's use of point-of-view provides another formal tool to control meaning. According to theorist Victor Burgin, "It is the position of the point-of-view, occupied in fact by the camera, which is bestowed upon the spectator.…[T]hrough the agency of the frame, the world is organized into a coherence which it actually lacks, into a parade of tableaux, a succession of decisive moments."[12]

The most effective way of achieving this coherence, Burgin argued, is to center the camera/viewer within the frame; and many survey photographs use centered framing in precisely this manner. For example, by centrally locating the viewer in a place that can hardly be occupied, the boulevard photo projects a potential for inhabitation and tranquility quite apart from the reality of noise, fumes, and concern for physical safety.

Left: +7.2 Highest-rated image in the Visual Preference Survey.
Right: +6 Pedestrian realm with outdoor café.

Tension between Reality and Formality

Consider the credence given to a photograph, as opposed to a drawing of a place. A photograph offers evidence that a place seems certain to exist. But what is one intended to "see" or believe? André Bazin has claimed that "the objective nature of photography confers on it a quality of credibility absent from all other picture making."[13] Roland Barthes has gone one step further, arguing that a photograph is essentially fused with its referent, its principal statement being "that has been."

Many contemporary artists have attempted to challenge this common belief in the denotative function of the photograph. For example, they may obscure indicators of time and place, usually structured to reinforce the authenticity of an image, to erode the sense of a specific reality.

Although the Milwaukee survey photographs are not formally constructed to disrupt the authenticity of the referent, locational clues are often obscured in order to focus the viewer on abstract qualities of place. Experience of the real place thus does not correspond with the photographic representation.

While posing as ordinary snapshots, other forms of dislocation may also be used. For example, depth perception may be reinforced, and the viewer may be drawn into photographic space through perspectival devices such as the regularly spaced, receding lampposts of the waterfront image shown here. Or else, homogenized formal elements may be used to dislocate the viewer, seducing him or her to project one city onto another, irrespective of specific production or evolution of place. Thus, red bricks and a handful of ubiquitous shops may appear to be easily transposed from Venice or Sacramento to Milwaukee—irrespective of climate, culture, or political economy.

It seems possible that physical environments could be more critically represented in public processes. Attempts have been made to use video and film to overcome some of these limitations. Current image-survey methods might also be improved by consistently pairing eye-level photographs with a plan or bird's-eye view that clarifies connections between spaces.

But there is no substitute for the real thing. Taking resident groups on visits to real places within their community—places that are culturally and historically specific to them—is the only way to truly provide a multisensory experience.

This is not an unreasonable suggestion, since the actual number of survey respondents is usually manageable, and the process could involve a series of weekend walkabouts. Participants might be asked to take photographs during the walk to bring back for reference in discussions. In this way, the photo as a referent to some larger whole might be more meaningful than a disembodied image, removed from its context. If the objective is to more fully engage residents in the design of *their* city, it is crucial that they have a deeper understanding of limitations and opportunities beyond the visual.

Acts of Translation

All photographs are made twice over—first configured by the photographer, and then more significantly reconfigured by the viewer. Even though neither of these acts can ever be "objective," image surveys rarely acknowledge these acts of translation. Nelessen has summed up the VPS process as follows:

> *The VPS™ provides the pictures/vision of what your community wants and what it does not want on its land. The images are not arbitrary; they are not unreasonable. They are a product of a public process. They represent public consensus from people who have experienced the place. They provide insight and reasoned responses. They represent a consensus vision. It is planning and design by democracy.*[14]

Widespread citizen engagement in planning processes is essential, yet one must ask whether these image surveys produce a meaningful public process, or set up a self-fulfilling prophecy.

While proponents of image-survey methodologies imply that photographs are value free, a majority of the 230 images shown in the Milwaukee survey displayed three dominant qualities: wealth, leisure and nostalgia. It was assumed that survey respondents would concentrate on the spatial and formal aspects of the places depicted. Yet, one photograph after another like that on p. 41, contains signifiers of prosperity, showing athletic white pedestrians ambling along brick streets lined with fashionable shops, cafes, and mock gas lamps.

The viewer clearly relates such photographs to personal experience, subjecting "this is" to individual interpretation and translation, thereby influencing his or her perception and survey ratings. The result is that survey participants evaluate the made-placeless photographs—casting votes for those that exhibit the most attractive trappings of contemporary consumer culture and constructed nostalgia.

Meaning

Can these survey images hold meaning? If so, to what degree of precision can this meaning be predicted? Numerous theorists have examined this question in terms of painting and photography. Perceptual psychologist and art theorist Rudolf Arnheim wrote that pictures "do not offer explicit formulations of intellectual concepts, which are the prerogatives of language."[15] He argued that we are affected by pictures, but do not know what they mean—insisting that images are poor conceptual communicators.

Victor Burgin has countered this position, insisting that "content, too, may be *produced* as deliberately as one may plan the formal composition of the photograph."[16] What is the meaning of the image shown above of the commercial entryway? Is it really that "Harley Davidson, a corporate American success story, should be more visible in Downtown," as the survey authors concluded after the ranking was complete.[17]

Furthermore, differences in the cultural background of individual viewers cannot be discounted. Might an African-American, for example, not be expected to respond differently to such an image of largely white pedestrians in an upscale commercial district?

The serial presentation of the 230 images in the Milwaukee survey also creates a cumulative effect. As each photograph is projected for a few seconds, viewers are drawn into complicity with a set of values and meanings intentionally or unintentionally structured into the survey sequence. Perhaps the survey process would be improved if participants could see all the images first to establish a context, then rank them at their own pace in a more thoughtful manner.

Left: +4.4 Harley Davidson, a corporate American success story, should be more visible in Downtown.

Right: +5.8 Police must be visible....

Text and Image

A photograph is polysemic—containing different meanings that are usually controlled by juxtaposition with the verbal text of a title or caption. Roland Barthes discussed the function of text in relation to image as either *relay* or *anchorage*. "In relay, the image and the linguistic text are in a relationship of complementarity: the linguistic message explains, develops, expands the significance of the image."[18] By contrast, text can also *anchor* one among several possible meanings, while clearly rejecting the others.[19]

The photographs in the Milwaukee Downtown Plan use captions to anchor the desired connotation. However, because the captions were not attached until after the survey, these may not be the same meanings participants had in mind. Did most viewers think, "Police must be visible…" as they rated the photograph of two horses carrying smiling uniformed men through a lovely park? Initially projected as straightforward snapshots of real urban places, the survey photographs from the Milwaukee Downtown Plan were only anchored by captions at the time of publication.

What are the alternatives? What if captions were included from the beginning? The survey's creators already know what they want each image to "embody." In relation to VPS, Nelessen's book provides a long list of positive and negative planning and design features that can be extracted from the survey images. Moreover, what if the

images were accompanied by clear captions listing location, time, and other relevant information to anchor the image more specifically to place?

Representation and Identity

The pervasive influence of language is felt even without the presence of captions. According to Bazin, "Even the uncaptioned photograph, framed and isolated on a gallery wall, is invaded by language when it is looked at: in memory, in association, snatches of words and images continually intermingle and exchange one for the other; what significant elements the subject recognizes 'in' the photograph are inescapably supplemented from elsewhere."[20]

The idea that one brings individual experience to the photograph assists our understanding of the constitution of the human subject through representation. Representation creates identity, and we form ourselves—and our places—to be like images around us.

Those who represent the culture to itself thus have the power to create identity. Representations of architecture, landscape, and urban space within real estate promotion, advertising, and public planning processes have immense power to shape formal and spatial expectations or norms. Photography, film, television, digital media, and advertising teach us not only how to "see," but how to construct our understanding of the city and our place in it. For this reason, those immersed in contemporary culture—both those who inhabit and those who design the built environment—would benefit from greater critical awareness of how to receive and construct photographic images.

Urban photographs are not simple depictions of real places, but images intentionally constructed by both maker and viewer. Urban image surveys are about what someone wants the city to "look" like—not what it *feels* like to be there, or how it works. Photography and simulated urban images are valuable but inadequate tools by themselves for understanding existing urban conditions, imagining better places, and implementing positive change in the city. A richer set of methods might support a more meaningful collaboration between designers and inhabitants, addressing complex urban realities outside the frame.

Notes

1. Andrew Saint, "Architecture as Image: How Can We Rein in this New Beast?" in Peter Rowe and William Saunders, eds., *Reflections on Architectural Practice in the Nineties* (New York: Princeton Architectural Press, 1996). Also see Jean Baudrillard, *The Ecstasy of Communication* (New York: Semiotext(e), 1988).

2. Nelessen described the history of the Visual Preference Survey in *Visions for a New American Dream: Process, Principles and an Ordinance to Plan and Design Small Communities* (Chicago: American Planning Association, 1994). Also see the A. Nelessen Associates, Inc., website at http://www.nelessen.org.

3. Nelessen, *Visions for a New American Dream*, p. 83.

4. See http://www.lgc.org for the Community Image Survey developed by the Center for Livable Communities.

5. This lineage can be traced to John Zeisel's *Inquiry by Design: Tools for Environment-Behavior Research* (Monterey, CA: Brooks/Cole Publishing Company, 1981). More recently, while building on Kevin Lynch's seminal work on mental maps, Jack Nasar has made connections to psychological theory and recent environmental behavior research. It is important, however, to distinguish between Lynch's use of the word "image" as an experientially derived cognitive image, and the Community Image Survey's use of the term to describe a photograph of an urban condition. See Jack L. Nasar, *The Evaluative Image of the City* (London: Sage Publications, 1998); and Kevin Lynch, *The Image of the City* (Cambridge, MA: MIT Press, 1960).

6. Quote from the Center for Livable Communities website at http://www.lgc.org.

7. Christian Norberg-Schulz, *Meaning in Western Architecture* (New York: Rizzoli, 1980).

8. Juhani Pallasmaa, *The Eyes of the Skin: Architecture and the Senses* (London: Academy Editions, 1996).

9. In a short essay accompanying the reissue of *The Eyes of the Skin*, Pallasmaa further developed his thoughts on vision. "A remarkable factor in the experience of enveloping spatiality, interiority, and hapticity is the deliberate suppression of sharp, focused vision. This issue has hardly entered the theoretical discourse of architecture, as architectural theorizing continues to be interested in focused vision, conscious intentionality, and perspectival representation." Juhani Pallasmaa, "Eyes of the Skin: Architecture and the Senses," *Architecture*, March 2006, pp. 28-29.

10. See http://www.georgetown.org/pdfs/williams.drive.study.2005/Civic.Design.Presentation.12-8-05.pdf. The image of Charlottesville's "corner" district appears on page 8 ranked with a 5.3.

11. Stanley Cavell, *The World Viewed* (New York: Viking Press, 1971), p. 24.

12. Victor Burgin, "Photography, Fantasy, Function," in *Thinking Photography* (London: Macmillian Press, 1982), p. 146.

13. André Bazin, "The Ontology of the Photographic Image," in *What is Cinema?* (Berkeley: University of California Press, 1967), p. 13.

14. Nelessen, *Visions for a New American Dream*, p. 97.

15. Rudolf Arnheim, "The Images of Pictures and Words," *Word and Image*, Vol. 2, No. 4 (Oct-Dec 1986), p. 310.

16. Victor Burgin, "Art, Commonsense and Photography," *Camerawork*, London, 1976, pp. 1-2.

17. Milwaukee Downtown Plan, p. 31.

18. Burgin, "Art, Commonsense and Photography," p. 5.

19. Roland Barthes, "The Third Meaning," *Image, Music, Text* (London: Fontana, 1973).

20. Burgin, "Photography, Fantasy, Function," p. 192.

All the images reproduced here were part of a Visual Preference Survey conducted between October 1997 and January 1998 and published on pages 18-35 of the Milwaukee Downtown Plan (1999). The entire document may be viewed online at http://www.mkedcd.org/downtownplan. All photos courtesy of A. Nelessen Associates