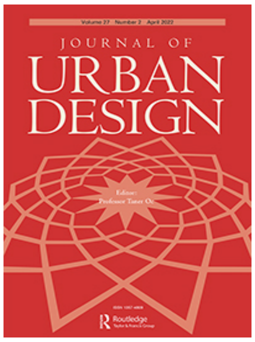

Journal of Urban Design

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/cjud20

Street design preference: an on-line survey

Emily Talen, Kyoung Whan Choe, Gaby N. Akcelik, Marc G. Berman & Kim L. Meidenbauer

To cite this article: Emily Talen, Kyoung Whan Choe, Gaby N. Akcelik, Marc G. Berman & Kim L. Meidenbauer (2022): Street design preference: an on-line survey, Journal of Urban Design, DOI: 10.1080/13574809.2022.2066512

To link to this article: https://doi.org/10.1080/13574809.2022.2066512

Published online: 06 May 2022.

JOURNAL OF URBAN DESIGN
https://doi.org/10.1080/13574809.2022.2066512

Routledge
Taylor & Francis Group

Check for updates

Street design preference: an on-line survey

Emily Talen[a], Kyoung Whan Choe[b], Gaby N. Akcelik[b], Marc G. Berman[b] and Kim L. Meidenbauer[b]

[a]Urbanism Lab, Social Sciences Division, University of Chicago, Chicago; [b]Environmental Neuroscience Lab, Department of Psychology, University of Chicago, Chicago

ABSTRACT
This paper is a methodological and empirical contribution that reports on the results of an innovative on-line survey of streetscape preference. The rating of experts and non-experts were compared to gauge the reliability of preference for views about streetscape quality. The goal was to evaluate the degree to which a set of streetscape design characteristics were similar among experts and non-experts and what dimensions of streetscape quality had the most agreement. Results show relatively high correlation between measures of streetscape quality and respondent preference, although there were some differences as well. Interestingly, among the six streetscape qualities studied, the quality of 'walkability' showed the least agreement among respondents.

KEYWORDS
Streetscape; design preferennces; walkability

Urban designers need to be responsive to community needs and preferences, and this extends to the qualities deemed important for good street design. To help in this effort, urban designers should make use of methods that help gauge design preferences. To that end, this paper reports on the results of an innovative on-line survey of streetscape preference. The paper compares the rating of experts and non-experts to gauge preference for streetscape qualities that were drawn from a field manual of streetscape design. Three urban design experts and a sample of novice participants were recruited from Amazon's Mechanical Turk (MTurk) to respond to a survey ranking preferences on dimensions of streetscape design. For the MTurk survey, sample size was determined by obtaining a large enough sample of ratings per image. Images of streetscapes from around the City of Chicago were used, where each image shows varying qualities related to pedestrian experience.

The research questions are two-fold. First, what is the degree of alignment between experts and non-experts when it comes to streetscape design? Relatedly, what dimensions have the most and least agreement within each group (experts and non-experts)? The goal was to help understand perceptions of streetscape quality and the degree to which these perceptions vary. Results show relatively high correlation between measures of streetscape quality and respondent preference, although there were some differences as well. Interestingly, among the six streetscape qualities studied, the quality of 'walkability' showed the least agreement among respondents.

CONTACT Emily Talen talen@uchicago.edu

The survey employed a set of qualities of streetscape design drawn from a field manual, using a novel approach to survey research to measure preference. Two online surveys using two sets of images were conducted. One set of images was obtained using Google Street View images, which were selected based on a GIS analysis of block quality. The second set used 1119 Chicago street images taken in the field by a team of project assistants. Three experts and 1028 Amazon MTurk workers completed that study. The first study was completed in October 2019 and the second study was conducted in April, 2020.

Background

Streets and sidewalks are essential components of the urban environment, not only because they constitute a major source of publicly owned land, but because they serve multiple functions: servicing car, bike, and pedestrian travel modes while also providing a setting for commerce, physical activity, social interaction, and communication (Jacobs 1995; Carmona et al. 2018; Park, Tian, and Larsen 2019). In the past decade, significant research has been undertaken to assess streetscape qualities and their effects. This has been motivated by, first, a perceived need to improve place quality, which many in the urban planning and design field, especially in the U.S. but also globally, believe is too car-dominated and lacking in qualities amenable to pedestrians. Second, these place qualities have been extended further by research devoted to assessing the effects of place quality on a range of other possible outcomes, particularly health, active living, crime and safety perception, social interaction, and climate change (Talen and Koschinsky 2014; Talen 2015). The urban design qualities of streetscapes are believed to have a significant impact on the ability of streets to be successful in social and economic terms (Jones, Al-Shaheen, and Dunse 2016).

Streetscape preference studies are but one category of environmental preference research (e.g., Kaplan and Kaplan 1989). Many studies use images as visual representations of the physical environment and predict human perceptions of these images (Ibarra et al. 2017; Kardan et al. 2015). Researchers have explored the differences in these perceptions among population groups – experts vs. non-experts, or people from varying cultural backgrounds. Herzog (1992) observed that environmental preferences were determined by judging images of urban spaces that fell into four categories: 'open-undefined, well-structured, enclosed settings, and blocked views'. Well-structured spaces were viewed the most favourably due to their complexity and coherence. Nasar (1994) found that pleasant urban spaces most often correlated with order or compatibility between design elements and moderate complexity, whereas exciting urban spaces were associated with higher levels of complexity and low order. Research by Krempen (1974) underscored the importance of complexity and variation in the built environment in creating positive environmental perceptions and behaviours.

Understanding how people perceive the built environment has been accomplished using surveys, interviews, and field observation (McGinn et al. 2007). These approaches have been critiqued as being 'subjective' and introducing 'response bias' (Zhou et al. 2019), but also for being time-consuming and impractical. An alternative is to use computational methods to measure streetscape qualities. With the development of computer vision algorithms and high-performance computing systems, semantic scene segmentation techniques have been used to extract features from images to predict people's perceptions of the built environment. Zhang et al. (2018) and Tomás et al. (2019) used data from MIT Place

Pulse dataset (http://pulse.media.mit.edu/) to measure perceptual qualities from six dimensions: 'safety', 'beauty', 'depressing', 'lively', 'wealthy' and 'boring'. Researchers have been exploiting the new availability of street-level images like those provided by Google Street View to more efficiently extract urban features and measure the spatial distribution of elements like street greenery (e.g., Li, Cai, and Ratti 2018). Xiangyuan et al. (2020) extracted information from millions of panoramic images to create an index of streetscape perception. They classified street segments in Shenzhen, China according to five criteria to create a 'streetscape perception map'.

One research area implicating streetscape quality involves the linkage between architectural form and cognitive health, constituting the field of 'cognitive architecture' (Sussman and Justin 2021; Ellard 2015). Relatedly, psychologists and neuroscientists have been investigating the science of environmental impact on psychological well-being – in particular that the built environment can have a significant impact on mood, behaviour, and cognitive health (Adams 2014; Cooper and Burton 2014; Hartig 2008; Joye 2007; Kaplan 1987). The bi-directional and dynamic relationship between the environment and brain functioning has been laid out by the field of environmental neuroscience (Kaplan and Berman 2010).

Another motivation for understanding streetscape design preference is walkability: what kinds of streets provide better pedestrian environments and thus support walkability (Xiangyuan et al. 2020)? Streets are implicated in walkability research since streets are the primary settings for pedestrian activity (Clemente et al. 2005; Adkins et al. 2012). Often this research has been directed at finding the 'morphological and spatial structures' associated with walking behaviour – such as building enclosure and frontage quality (Kashef 2011, 39; Speck 2012). The level and quality of experience is believed to be dependent on the street's ability to offer comfort, safety, and visual interest (Southworth 2005). Walking is also believed to contribute to mental health by improving cognitive functioning and preventing cognitive decline (Weuve et al. 2004).

Analytical framework

To understand streetscape design preference – specifically, how expert and non-expert preferences compare, and how qualities are similarly ranked – a useful approach is to start with a set of design qualities to evaluate. For this study, the streetscape qualities included in *Measuring Urban Design Qualities – An Illustrated Field Manual* were selected (Clemente et al. 2005). The manual is the result of research aimed at identifying the physical characteristics that support pedestrian activity (Ewing et al. 2006; Ewing and Handy 2009; Clemente et al. 2005). The research examined a long list of perceptual qualities, such as adaptability, ambiguity, centrality, clarity and compatibility of the urban environment which were most frequently discussed in previous literature and shown as important qualities for pedestrian experience by empirical evidence. The study involved a 'visual assessment survey' to measure these qualities for 48 video clips of streetscape samples and identify physical features related to these qualities such as number of courtyards/plazas/parks, number of major landscape features and proportion of historical buildings.

The study consolidated expert opinion about urban design quality in a measurable way, qualities that, because of their measurability, could be empirically validated, for example by associating measured characteristics with observed pedestrian life or

Figure 1. Street walls with high (left) vs. low (right) permeability. Source: *Talen, Emily*

measures of active living. To settle on these urban design qualities, researchers interviewed experts over an extended period to arrive at a set of five urban design qualities that could be objectively measured. The selection process included inter-rater reliability of scene ratings, relationships between physical features and urban design qualities, and correlations with overall walkability. After measuring the performance on the selection criteria and validation with field tests, imageability, enclosure, human scale, transparency and complexity were selected and the protocols for measuring these five qualities were included in a field manual. Ewing and Handy (2009) further analysed the results from Ewing et al. (2006) and discussed qualitative and operational definitions for the five perceptual qualities. The descriptions and measurable qualities associated with each dimension are listed in Table 1 and summarized below. Figure 1 shows an example of what is often considered high vs. poor quality in streetscape design.

Imageability relates to the degree to which a place is memorable and distinctive. Imageability might be strongest where there are visual contrasts in terms of surface treatment and building height and form, or where there is a richness of architectural quality, such as landmark buildings. Imageability might be enhanced by street furniture, signs and symbols of various kinds, and terminated vistas. Appleyard (1979) and Lynch (1981) produced pioneering work on imageability as an integral component of an individual's identity, belonging and place-making processes, and that positive place-based associations lead to increased community involvement and investment. Timms and Tight (2010) found that introducing more varied, dynamic and aesthetically pleasing streetscapes and frontages increased walkability and biking and other sustainable modes of transit.

Complexity is about visual richness and diversity of activities. Complexity might involve patterns and designs that mimic nature (Hollander and Anderson 2020), and the presence of these patterns may confer psychological benefits (Alexander 2002; Joye 2007; Salingaros 2007; Berman, Jonides, and Kaplan 2008). Ellard (2015) found that complex facades create positive effects; monotonous, simplistic facades impact people negatively (Salingaros 1998). While surveys have revealed strong preference for 'strongly articulated patterns of fenestration, clear and graphic silhouette outlines and similarities of surface texture' that create a sense of order, 'highly varied' streetscapes with 'no underlying sense of order' are chaotic rather than complex, and tend to be low-rated (Gjerde 2011, 160).

Table 1. Urban design qualities and their measurement.

Imageability	Measures
The quality of a place that makes it distinct, recognizable, and memorable. A place has high imageability when specific physical elements and their arrangement capture attention, evoke feelings, and create a lasting impression.	number of courtyards, plazas, and parks number of major landscape features proportion historic building frontage number of buildings with identifiers number of buildings with non-rectangular shapes presence of outdoor dining number of people noise level
Enclosure	**Measures**
The degree to which streets and other public spaces are visually defined by buildings, walls, trees, and other elements. Spaces where the height of vertical elements is proportionally related to the width of the space between them have a room-like quality.	number of long sight lines proportion street wall proportion sky
Human Scale	**Measures**
The size, texture, and articulation of physical elements that match the size and proportions of humans and, equally important, correspond to the speed at which humans walk. Building details, pavement texture, street trees, and street furniture are all physical elements contributing to human scale.	number of long sight lines proportion windows at street level average building heights number of small planters number of pieces of street furniture and other street items
Transparency	**Measures**
The degree to which people can see or perceive what lies beyond the edge of a street or other public space and, more specifically, the degree to which people can see or perceive human activity beyond the edge of a street or other public space. Physical elements that influence transparency include walls, windows, doors, fences, landscaping, and openings into midblock spaces.	proportion windows at street level proportion street wall proportion active uses
Complexity	**Measures**
The visual richness of a place. The complexity of a place depends on the variety of the physical environment, specifically the numbers and kinds of buildings, architectural diversity and ornamentation, landscape elements, street furniture, signage, and human activity.	number of buildings number of basic building colours number of accent colours presence of outdoor dining number of pieces of public art number of people

Source: Based on Clemente et al. (2005).

Transparency concerns permeability and interactive frontages – buildings with windows on the ground floor and the ability to see inside a building – features considered essential for a high quality pedestrian experience (Jacobs 1961; Mantho 2015; Kickert 2016). Transparency has been associated with large storefront windows where item display is easily seen (Askari and Soltani 2018). Researchers have empirically demonstrated the 'far-reaching benefits' of active frontages in terms of 'safety, comfort, sociability and liveliness' (Heffernan, Heffernan, and Pan 2014, 92). Jan Gehl defined this interactivity as buildings being engaged in 'meaningful conversation' with pedestrians (Gehl 2011, 47). The large, uniform walls of formula ('big box') retailers and buildings with opaque glass frontage produce negative psychological responses (Muhlebach and Muhlebach 2013; Ellard 2015). The presence of ground-floor windows have been shown to be strongly correlated with perceived walkability (Oreskovic et al. 2014).

Enclosure refers to the degree to which streets and sidewalks are visually enclosed by surrounding buildings or other vertical elements (such as street trees), where the context of the street contributes spatial definition and provides a sense of intimacy for the pedestrian (Yin and Wang 2016; Naik et al. 2014; Wang et al. 2019). Parking lots and undefined open space are thought to reduce sense of enclosure on the street (the feeling of being in an 'outdoor room'), and therefore pedestrian quality; a continuous façade is preferred (Park, Tian, and Larsen 2019). Part of this sense of enclosure is finding the proper balance between street width and building height (Jacobs 1995). Soltani, Hosseinpour, and Zare (2018) assessed urban street qualities and found that a sense of enclosure had the greatest effect on feelings of safety and walkability. Al-Homoud and Tassinary (2004) investigated the potential moderating effect of space enclosure on social interactions, finding that enclosure facilitates interaction and increase sociability and sense of belonging (see also Gillem 2009).

Human scale relates to the relationship between the human body and built form (Gehl 2011). Streets designed solely in function of cars tend to not have a human scale in which spaces are defined by human measurements and capabilities (Shen et al. 2018). Monumental spaces tend to associate with long travel distances and thus are not considered to be at a human scale. Monumental buildings, either with long horizontal or high vertical emphasis, also dwarf human size. Human scale also pertains to the fine-grained elements of the streetscape, where human scale means that humans can directly interact with surfaces and openings rather than engaging these elements from afar (Ye et al. 2019).

Besides the above five qualities and walkability, 'disorder' was initially used (in the pilot study) as an additional aspect of streetscape quality to gauge. Order should not be thought of as the converse of complexity; disorder, however, has been shown to trigger certain negative perceptions and anti-social behaviours. One theory is that disorder inspires negative feelings about place quality and signals a lack of caring about particular places. This 'broken windows' theory, first introduced by Wilson and Kelling (1982), posits that an unrepaired broken window or other associated incivility like litter could engender a perception that a place is unsafe and tolerates disorderly behaviour (Braga et al. 1999; Braga and Bond 2008; Chae and Zhu 2014). On the basis of these studies, Kotabe, Kardan, and Berman (2016) found that visual disorder alone (apart from demographic factors or policing, or from semantic cues that rules had been broken, such as litter and graffiti) had a measurable impact on rule-breaking behaviours. It should be noted that broken windows theory, though widely accepted and often used to guide policing practice, has also been refuted, with some studies showing that the impact of physical disorder on crime has been over-stated (e.g., O'Brien, Farrell, and Welsh 2019).

Research questions

There are two main research questions. First, what is the degree of alignment between experts and non-experts when it comes to streetscape design preference? Relatedly, what dimensions have the most and least agreement within each group (experts and non-experts)? The goal is to help understand perceptions of streetscape quality and the degree of rating alignment. Knowing the answers to these questions will facilitate understanding how streetscape quality can be measured, how perception varies and along what dimensions, and levels of agreement.

Methodology

Pilot study

Utilizing the concepts discussed above, the initial pilot study included eight perceptual qualities: the urban design qualities associated with walkability (imageability, transparency, complexity, enclosure and human scale), plus three additional dimensions: walkability, disorder, and preference. Preference was included, in part, because novices may prefer a street, but not be able to articulate what it is about the streetscape that they like. Therefore, having a preference measure provided a way to gauge if these design qualities were related to preference or not for the novice raters. In addition, preference for a scene has been shown to be highly related to, but not completely overlapping with other scene attributes, such as naturalness or orderliness in previous work (Kardan et al. 2015; Ibarra et al. 2017). As image preference has been shown to be an important predictor of the positive emotional effects of a given environment (Meidenbauer et al. 2020), and may provide additional information about scene quality in street scenes, it was included in this pilot study.

To measure these perceptions, a large sample of street images was obtained from Google Street View using their API (https://developers.google.com/streetview). The images were selected from blocks in the City of Chicago, narrowed to only those blocks that satisfied the following conditions: the block had a sidewalk view; it had at least one commercial use (i.e., the block was not entirely residential); it had at least one neighbourhood amenity (e.g., a grocery store, library or day care facility); and it did not have an obviously anti-pedestrian land use such as a parking lot. The last three conditions were determined using GIS data obtained from the City of Chicago. For each geo-coordinate in the selection of blocks, 4 images were extracted to be used in the pilot study.

A total of 552 Chicago street images were obtained from Google Street View by sampling two sidewalk images from 278 geo-coordinates in Chicago. A total of 588 Amazon Mechanical Turk (MTurk) workers completed the multi-image rating task across ten different dimensions/questions. Amazon Mechanical Turk (MTurk) is an online marketplace where ‘employers’ (usually social scientists and consumer researchers) pay Mturk ‘workers’ (any individual who signs up on the MTurk site) to complete surveys and tasks, referred to as HITs or Human Intelligence Tasks, for money. There are approximately 100,000 Mturk workers, with at least 2,000 active at any given moment, and the population pool is thought to refresh every 12 to 18 months, allowing for a relatively large and diverse group of workers over time (Difallah, Filatova, and Ipeirotis 2018). The majority of workers (75%) are located in the US, and all of the workers in the current study were US-based. In the US, on average, the gender distribution is relatively even (55% female, 45% male), the age distribution is slightly younger than the US working population (average age is approximately 35 years on Mturk vs. 45 years in US working population), and the average income is somewhat lower than a typical individual from the US (Difallah, Filatova, and Ipeirotis 2018). However, Mturk workers are a much more diverse sample than adult participants in social science research, which are often predominately college students.

In our sample, the mean age of the participants was 38.5 years (SD = 10.9; range = 20 to 73 years). Of the 588 total participants, 375 identified as male, 209 identified as female, and 3 identified as other/nonbinary. Each participant rated all 552 images on one design

Figure 2. Screenshot of one of the survey questions given to participants. Source: MTurk survey.

Table 2. Questions used for pilot study.

Perceptual quality	Survey instructions per perceptual quality
Walkability	*Select 4 streets you would most want to walk down.*
Preference	*Select 4 street images you like.*
Imageability	*Select the 4 streets that have the most character.*
Complexity	*Select 4 streets with the most visual richness and diversity of activity.*
Transparency	*Select 4 streets where you have the sense that there is human activity going on inside of the buildings.*
Disorder	*Select the 4 streets that seem the most disorderly.*
Enclosure	*Select the 4 streets that most make you feel like you are in an outdoor room.*
Human Scale	*Select the 4 streets with the most human scale (e.g., small buildings and narrow streets).*

quality (e.g., complexity or enclosure), and each image was rated by 58.8 (SD = 2.8) participants on average. Participants rated images in a 20-minute online session. On each trial, participants are asked to select the four images that they evaluated most highly on one design quality (e.g., complexity or enclosure) among 12 images presented in a 4×3 grid, as shown in Figure 2. Each participant saw all images once and clicked the 4 images that best represented the dimension. One line of instruction was used to explain the participants' task, e.g., 'Select 4 streets you would most want to walk down', according to which the participant could click on 4 images in response. The online instructions used in the survey are listed in Table 2. A highly-rated image would be selected by many (or all) participants; a lower-rated image would be selected by few (or no) participants. The probability of selecting each image across participants, i.e., the choice probability, was used to quantify how much that image represented that dimension. To ensure that participants were paying attention to the task, they were asked to identify a blurry image (e.g., the leftmost image in the second row in Figure 2) that was presented along with the street images and drag it to the trash can at the bottom.[1] In the pilot experiment, 61 Mturk workers completed the task in one day.

Participants were placed in groups where they were presented with the same images. Theoretically, a highly-liked image would be selected by a large number of participants and a less-liked image would be selected by few participants. The probability of an image being selected in the survey for a certain perceptual quality was used as the score of the image on each perceptual quality. Formally, the score is calculated as follows:

Score (perceptual quality i, image j) = number of clicks on image j in the survey for perceptual quality i/ total number of participants in a group

After collecting the ratings from the pilot survey, split-half correlations were calculated to check if ratings were consistent across participants. The test provides an indication of the validity of the one-line questions that were used.

Study #2

The pilot study was used to verify the basic approach and check rating consistency. Based on the results of the pilot study, a second experiment was implemented to resolve two issues. First, it was determined that the quality of the images from the Google Streetview API was sometimes poor, particularly where the image included obstructions like trash bins, signage, and transit infrastructure. Another problem with Google Streetview imagery is that it is not recorded at eye-level. Images are shot from above, and include a wide angle perspective that distorts the distance relationships between the viewer and surrounding streetscape objects.

To resolve these issues, a team of student photographers was employed to randomly select and photograph streets in a variety of neighbourhoods throughout Chicago. Students were instructed to take photographs with their phones at eye level, from the sidewalk, angled to capture both sides of the street, and only on streets with commercial uses. A mix of residential and commercial uses was permitted, but students were told to avoid streets that were only residential. In total, 1119 suitable street images were obtained. About one-half of the images were mapped (i.e., those that had GPS information), as shown in Figure 3.

Second, based on the results of the pilot study, the procedure was refined to slightly change the wording for the 'imageability', 'transparency' and 'enclosure' questions, as listed in Table 3. Two questions were excluded from the pilot study, 'preference' and 'disorder'. Preference was omitted because of the overlap with the other dimensions, especially imageability. Disorder was deleted because the results showed weak correlation among participants, indicating ambiguity.

A total of 440 MTurk workers completed the image ratings. 438 participants completed the demographics questionnaire. The mean age of these participants was 38.1 years (SD = 11.5; range = 19 to 74 years). Of the 438 participants, 272 identified as male, 160 identified as female, 1 identified as other/nonbinary, and 5 preferred not to disclose their gender Each participant rated a randomly-selected set of images (708 out of 1119 images) and each image was rated by 34.8 (SD = 1.5) participants.

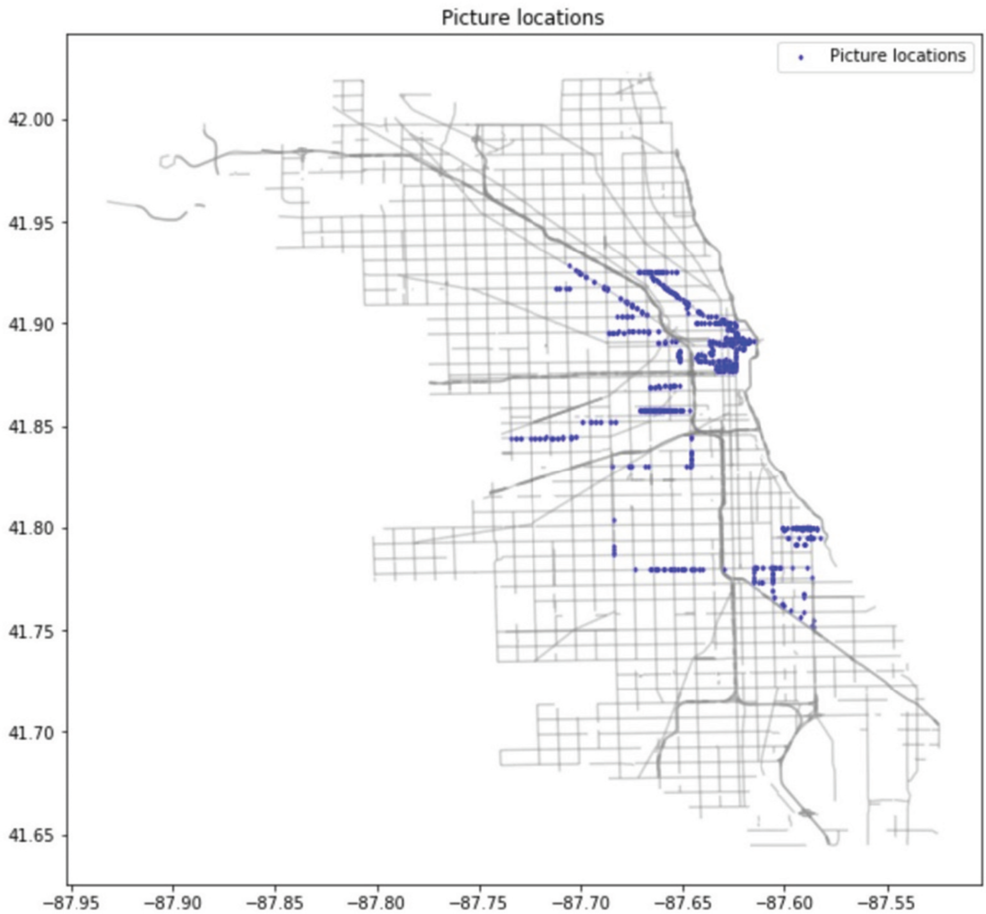

Figure 3. Image locations for approximately 50% of locations (second study). Source: Kyoung Whan Choe.

Table 3. Questions and participants for second study.

Perceptual quality	Survey instructions per perceptual quality	# participants	Average # of raters per image
Walkability	*Select 4 streets you would most want to walk down.*	52	32.9
Imageability	*Select the 4 streets that have the most character (i.e., that capture your attention)*	54	34.2
Complexity	*Select 4 streets with the most visual richness and diversity of activity.*	56	35.4
Transparency	*Select 4 streets where you can see or perceive what's going on inside of the buildings.*	52	32.9
Enclosure	*Select 4 streets that feel enclosed and room-like, rather than wide open.*	59	37.3
Human Scale	*Select 4 streets with the most human scale (e.g., small buildings and narrow streets).*	56	35.4

In addition to MTurk respondents, the second experiment was also completed by a panel of three experts. These experts are all employed in professional positions in the urban design field, with advanced degrees in planning and urban design. One has a master's degree in architecture, one has a PhD in urban design, and one has a PhD in planning. All three have teaching and research interests in urban design and its application in streetscape contexts. They were compensated for their participation and completed the same image rating task as the MTurk respondents. Each expert rated six data points for each image/expert (one expert saw each image six times).

Choice probability was used to indicate perceptual quality and evaluate results. In that method, the probability of selecting each image across participants is used to quantify how much an image represents a given dimension. However, because there were fewer data points for the experts, a second ranking called 'TrueSkill' was employed which is based on a statistical machine learning algorithm used to rank players (for example in Xbox Live).[2] Choice probability ratings are more intuitive, but they can be skewed if there are few raters. The TrueSkill rating is calculated by transforming clicks into 'winner' (clicked) and 'loser' (not clicked) pairs and then feeding those into the TrueSkill algorithm. The advantage of the TrueSkill score is that it takes into account whether an image is chosen over trivially low rated images or competitive images. The algorithm can make the score higher if the image was clicked among similarly high rated images because the chosen image 'won' the click over others (choice probability does not take into account whether an image is chosen over other high or low-rated images, so the score is somewhat coarse). Note that when there are many respondents (e.g., 60), choice probability is highly correlated (.9) with the Trueskill score.

Results

The distribution of preference ratings obtained in the pilot study is shown in Figure 4 (choice probability, on the horizontal axis, is used as a measurement of preference). The most and least liked street images are shown in Figures 5 and 6. The images reveal one of the problems identified with using the Google Street View images: some images were clearly residential streets (despite using GIS variables to constrain the selection) and thus were not accurately capturing the streetscape qualities

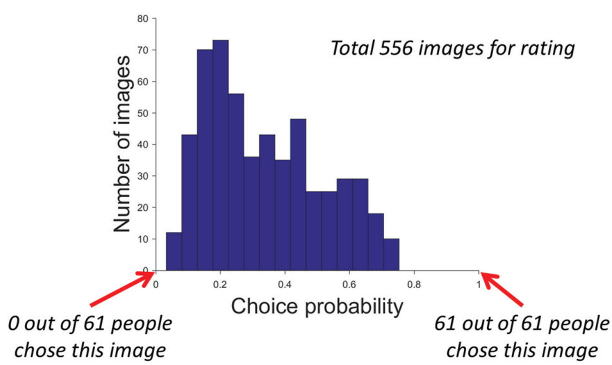

Figure 4. The distribution of preference ratings (pilot study).

Figure 5. The 12 most-liked street images from the pilot study. Note: Source: Mturk survey; CP refers to the choice probability in Figure 4.

Figure 6. The 12 least-liked street images from the pilot study. Note: CP refers to the choice probability in Figure 4. Source**: Mturk survey**

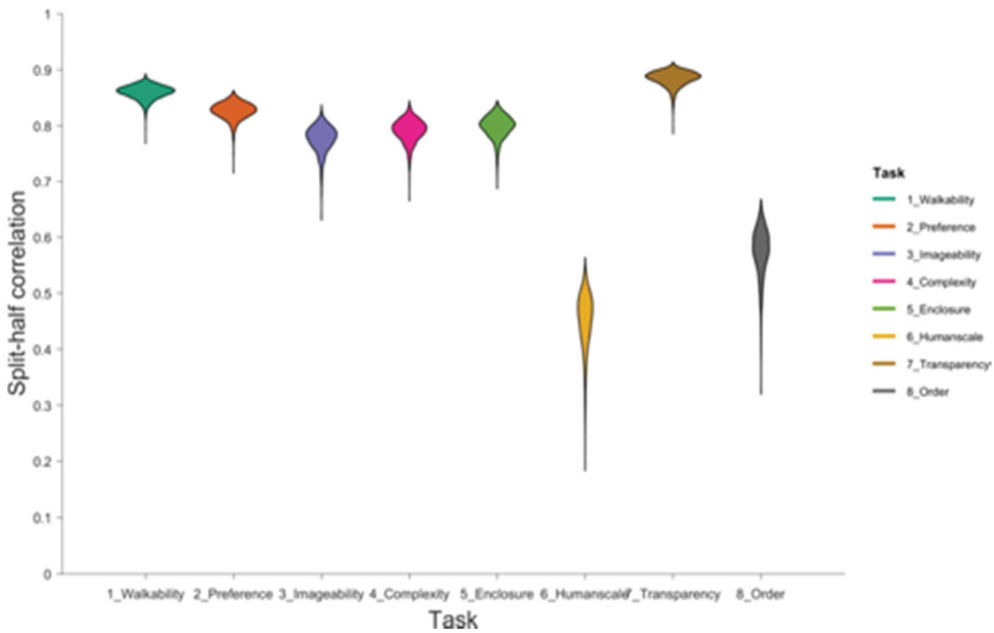

Figure 7. Split-half correlations for each of the eight perceptual qualities (pilot study).

intended to be the focus on this study. It is interesting that in the pilot study, the difference between most and least liked streetviews seemed to be almost entirely explained by the presence or absence of greenery.

Figure 7 shows the split-half correlation for the eight tasks, which was used to test inter-rater reliability. Overall, high split-half correlations were observed. The graph shows that ratings for walkability, preference, imageability, complexity, enclosure, and transparency are highly consistent across participants. However, it was also found that ‘human scale’ and ‘order’ yielded lower reliability, and the question wording in the follow-up study for human scale was revised. But for walkability, preference, imageability, complexity, enclosure, and transparency, ratings were highly consistent across participants, suggesting that these one-line questions are interpreted in a similar manner by ordinary people. Also of note is that walkability was positively correlated with imageability and enclosure, but not with human scale and transparency.

The remaining results, presented below, are from the second study, using a revised set of images and questions and employing a panel of three experts.

First, the degree to which respondents agreed with each other was evaluated – a measure of rater reliability. Agreement among experts was relatively high. Comparing each pair of experts on each dimension (resulting in 18 correlations), only one correlation was below .5 (between expert 1 and expert 2 on walkability, with a correlation of .453), and 13 out of the remaining 17 pairs were above .6 correlation.

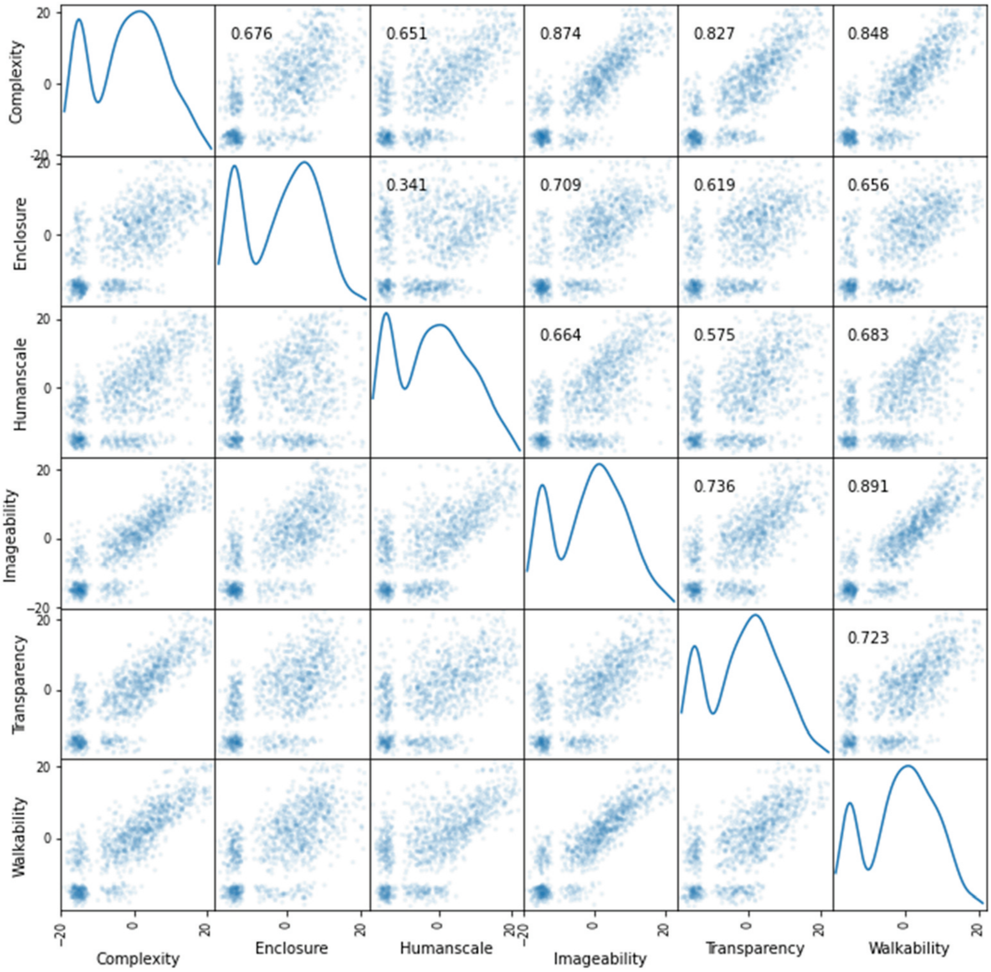

Figure 8. Expert ranking correlations.

The level of agreement between Mturk participants was obtained by calculating 2000 rank correlations between the ratings from two groups of 24 randomly-selected participants. Correlations were between .6 and .8 for all qualities except for human scale and walkability, which were around .5.

Next, the degree to which street qualities are rated similarly was examined, i.e., to what extent are ratings for each street quality correlated with other street qualities? This involved looking at the correlation among experts and among Mturk respondents separately.

Mostly, image ratings among experts were much more highly correlated than among non-experts. Figure 8 is a scatterplot matrix of the correlation among experts. The bi-modal distribution of the correlations indicates that there is dimension agreement on the low and high ends of the ratings. The clustered dots on the lower left corner represent images that received no clicks in 6 appearances. The figure shows that walkability was highly correlated with complexity and imageability, and complexity was highly correlated

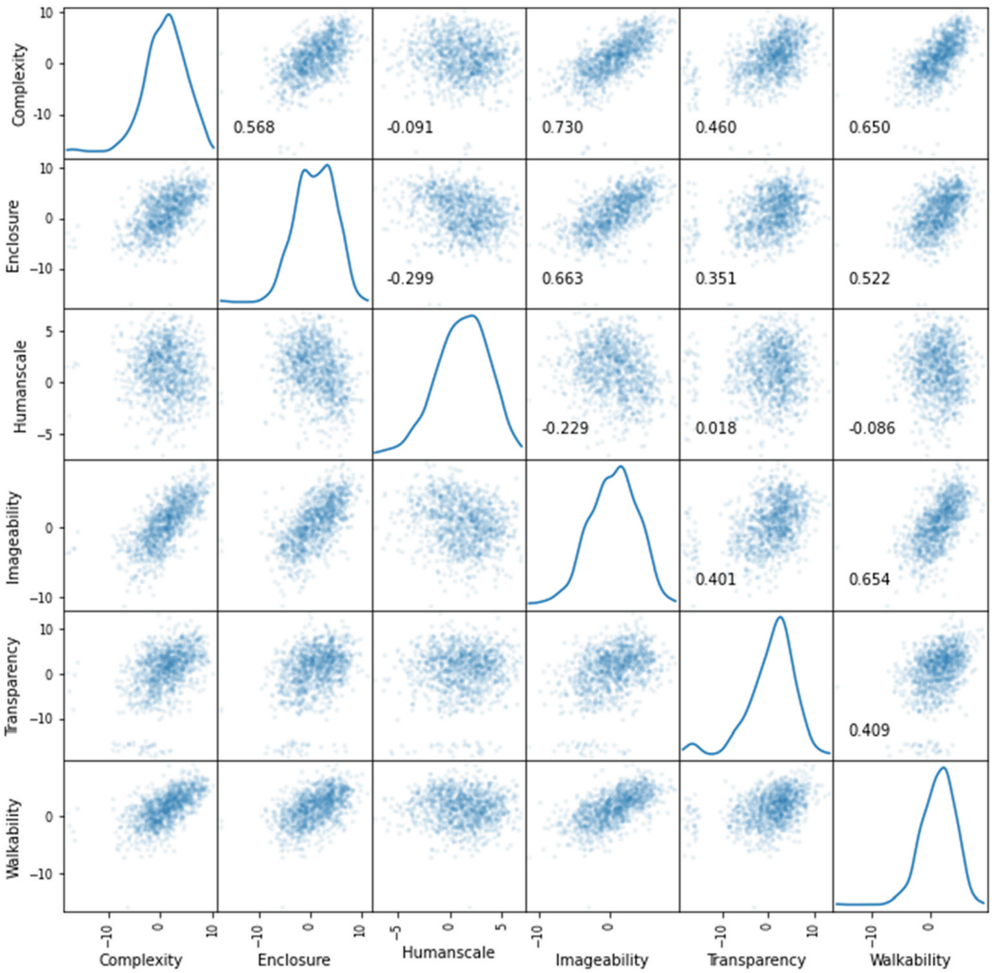

Figure 9. Mturk ranking correlations.

with transparency. Interestingly, enclosure and human scale were not as highly correlated. This may be due to the fact that street views with tall buildings in the downtown area might receive high ranking for enclosure, but experts would likely rank them lower on the human scale dimension.

Figure 9 is a scatterplot matrix of the correlations among Mturk respondents, and it confirms that overall, the degree of correlation was much lower for Mturk respondents than for experts. In particular, correlations between human scale and other qualities were low, which might indicate that the concept is less familiar . Human scale was inversely correlated with imageability, enclosure, and complexity. On the other hand, Mturk ratings for walkability were correlated with complexity and imageability, similar to expert ratings (although the correlation was not as high). The highest correlation for Mturk respondents was between complexity and imageability (.730), although again, not as high as the experts (.874).

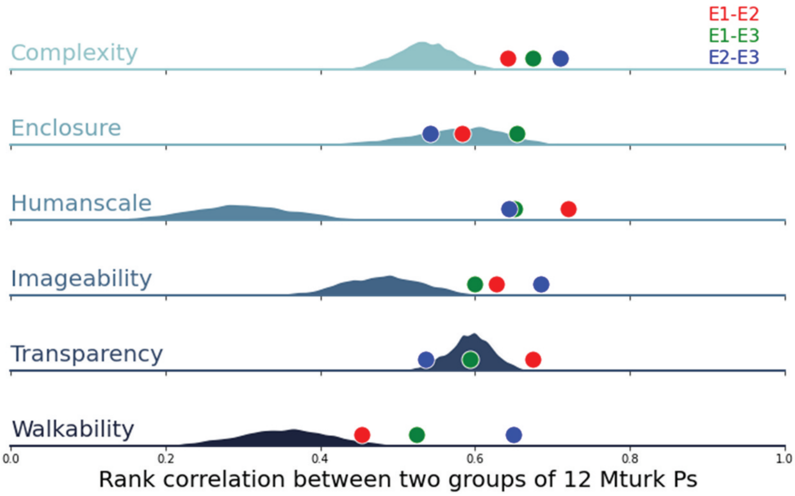

Figure 10. Rank correlation, Mturk and expert.

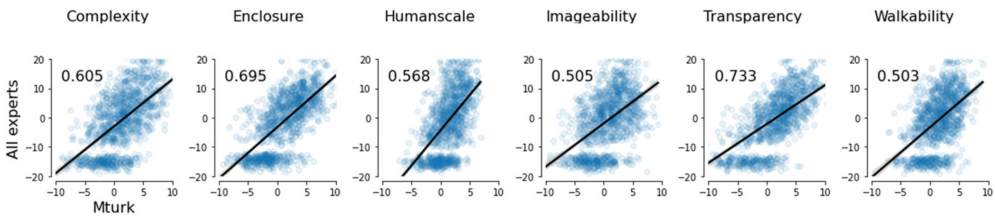

Figure 11. Expert vs. Mturk correlation.

Next the level of agreement among experts was compared, as well as the level of agreement among Mturk respondents more directly. Figure 10 compares ratings by showing how rank correlations among the two groups differ. The correlations between each pair of experts are shown together with the rank correlations of randomly selected Mturk respondents. Note that the level of agreement between Mturk participants was calculated using 12 randomly-selected respondents instead of the 24 reported above. This number was chosen in order to make the calculation more comparable to the expert ratings – because each Mturk participant saw a subset of images (approximately 700 from 1119), aggregating 12 participants would yield about eight ratings per image, which is relatively close to the expert ratings in which each image was rated by each expert six times.

Figure 10 shows that expert agreement on image ranking was higher for four out of six qualities. For complexity, imageability, and walkability, agreement between experts was higher than between Mturk participants, and for human scale, agreement was significantly higher among experts. For the dimensions of enclosure and transparency, agreement among experts and agreement among Mturk respondents were similar.

What about the agreement between experts and non-experts on rating streetscape dimensions? Figure 11 is a graphic illustration that helps interpret the level of agreement on image rating: a scatterplot matrix of the correlation between all experts

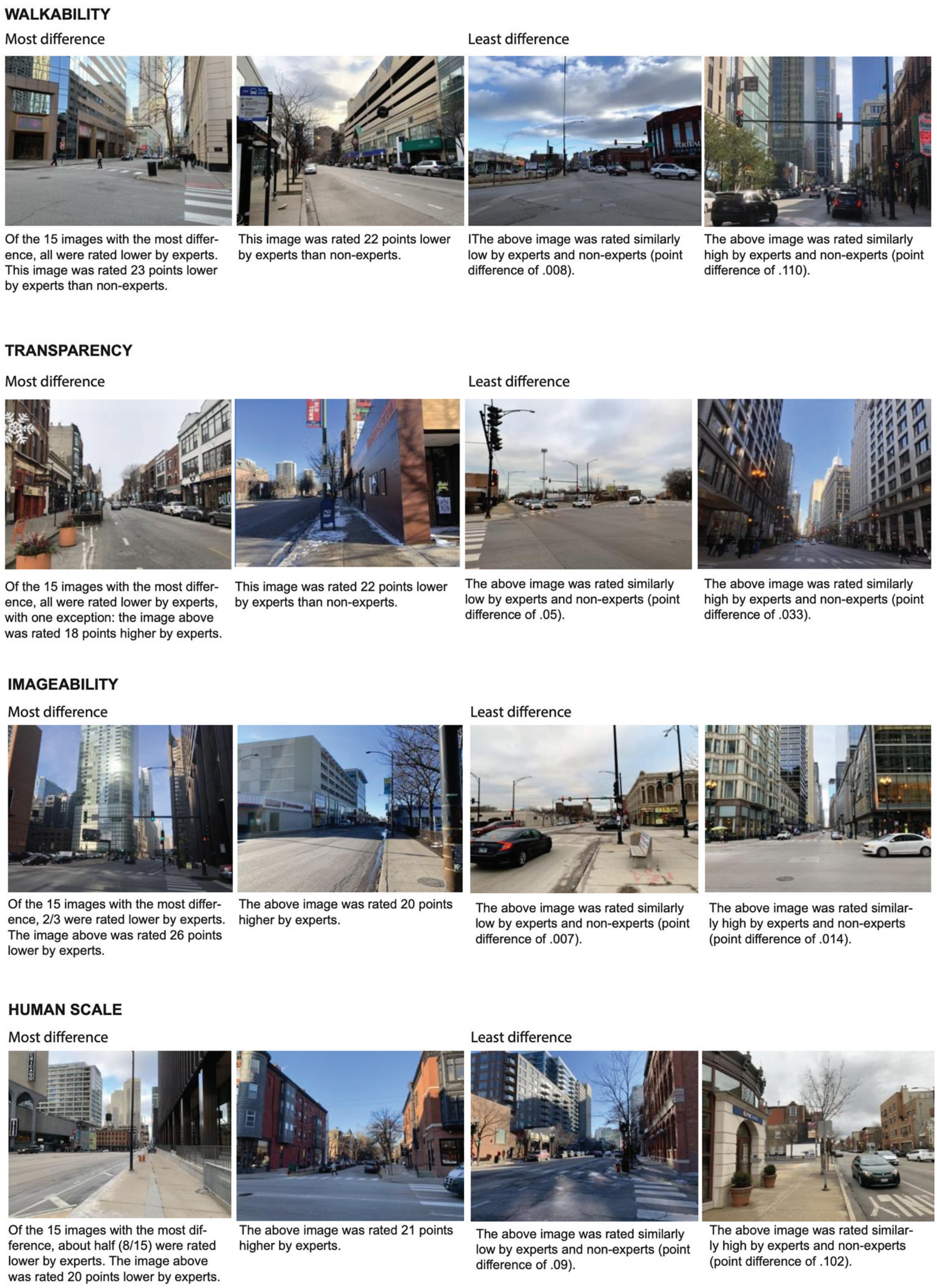

Figure12. Images with the most and least rating difference between experts and non-experts.

ENCLOSURE

Most difference

Of the 15 images with the most difference, 4/5 were rated lower by experts. The image above was rated 22 points lower by experts.

The above image was rated 19 points higher by experts.

Least difference

The above image was rated similarly low by experts and non-experts (point difference of .05).

The above image was rated similarly high by experts and non-experts (point difference of .014).

COMPLEXITY

Most difference

Of the 15 images with the most difference, all but 2 were rated lower by experts. The image above was rated 21 points lower by experts.

The above image was rated 20 points higher by experts.

Least difference

The above image was rated similarly low by experts and non-experts (point difference of .03).

The above image was rated similarly high by experts and non-experts (point difference of .107).

Figure12. Continued.

combined vs. Mturk respondents for each dimension. The results show that there was highest agreement on enclosure and transparency, and lowest agreement on walkability and imageability. Overall, there was fairly robust agreement between experts and novices even though the relationship between the features seemed to differ between the two groups.

Finally, the images that captured the largest rating differences and the greatest similarities between experts and non-experts for each dimension can be examined. Figure 12 is a series of images that show, for each dimension, images that had the most and least difference between experts and non-experts (Mturk respondents). The images can be interpreted in many different ways, but several interesting comparisons stand out.

First, images with wide differences in ratings between experts and non-experts tended to be images where experts rated images lower. Only in the case of 'human scale' was the ratings difference split evenly, where experts rated images lower about half the time. For walkability, transparency, and complexity, ratings differences were mostly a matter of experts rating images lower, and imageability and enclosure also produce lower expert ratings.

Second, images that experts rated much higher than non-experts would likely surprise the experts. For example, the images shown for enclosure, complexity and human scale show the hallmarks of what urban designers often associate with good streetscape design – spatial definition by buildings, architectural diversity and a variety of physical elements, and qualities that match the proportion of humans. And yet, an image with much higher rating for enclosure by non-experts had no building frontage, and an image showing good element diversity was rated much lower by non-experts. As another example, an image with much higher rating among non-experts for transparency shows only one side of the street, and the building seems to have fairly small windows.

One explanation for these differences might be that novices tend to be more literal in their interpretation. For example, an image showing a bridge overhead might be interpreted by non-experts as showing enclosure, whereas experts would view it as an obstruction that does not contribute to enclosure in a positive way. For complexity, a scene showing construction materials and heavy traffic might be interpreted as being 'complex' by novices, but experts would tend to reserve complexity as relating to feature diversity in a more positive way.

However, a few commonalities among ratings can also be seen. Low-rated mages with the least rating difference among experts and non-experts tended to be dominated by pavement and, with the exception of human scale, low-scale buildings. Images with similarly high ratings tended to be streetscapes with tall buildings, again with the exception of human scale.

Conclusion

This study is both a methodological and an empirical contribution. The approach demonstrated an efficient method for streetscape evaluation via crowdsourcing, showing how Mturk respondent ratings can be a viable, crowd-sourced approach to gaining an understanding of urban design qualities. The research also provided an empirical contribution, showing how the qualities included in an urban design field manual are held among experts and non-experts, and providing evidence that there is a relatively high correlation between that particular set of streetscape quality measures and respondent preference.

Another contribution of the study is that it showed how experts and non-experts vary in how they rank elements of streetscape design. Experts, with their background and training, tended to have higher agreement on streetscape quality. Enclosure and transparency produced higher levels of agreement between experts and non-experts overall, although some images produced interesting rating variation in individual cases. The qualities of walkability and imageability had lower agreement, perhaps signalling that these concepts invoke a higher degree of personal interpretation.

The use of static photography to record design preference is a limitation. One issue is that it is difficult to control for streetscape quality variation. Some images show street trees and other don't; some images illuminate the foreground or middle ground, while others show a distant perspective. In person recording of streetscape preference has the potential to overcome the limitations of static imagery, but it also carries its own logistical limitations, and large sample sizes involving a large number of respondents are often infeasible. The large sample size used in this study goes some way towards resolving the problems associated with static imagery, since image variability is widely distributed.

Another potential limitation to be acknowledged is the lack of pedestrians on the street in most of the images, in part due to the fact that many photos were taken during winter months. However, there is an argument to be made that a consistent lack of visible pedestrians in the survey images puts the focus squarely on the streetscape qualities being evaluated.

Despite these caveats and limitations, the research raised a number of interesting questions that need further study. What would account for ratings differences between experts and non-experts? Although there was relatively high correlation between theorized measures of streetscape quality and respondent preference overall, there were significant differences as well. Why did the most differences entail experts rating images so much lower? Why were there more differences in ratings on the dimensions of imageability and walkability and higher correlation with transparency?

Further study is needed to tease out the factors that might be involved. The urban design qualities used in the study were defined by researchers that specialize in urban design (Ewing et al. 2006; Ewing and Handy 2009); the meaning of the selected qualities may not be accessible to those who are not trained in urban design. Perhaps making the definitions and meanings of each quality more accessible to non-experts would yield more accurate results. In addition, expanding the participant pool of urban designers beyond three might yield different ratings outcomes.

Of particular interest in future studies would be the role of particular elements captured in the selection of photographs – things like trees, pavement, cars and pedestrians. Might the differences between experts and non-experts be somehow dependent on urbanistic scale and its representation? Or are the differences evident at any scale, from large format buildings and wide streets to small buildings and narrow streets? Are experts more attuned to triggers like parking garages, or knowledge that a given element is actually a fake frontage that lacks 'real' transparency or enclosure? What is the impact of what is being shown in the foreground and background of an image? Are experts looking more closely at the full range of building frontage, assessing top and bottom, in contrast to non-experts who are pursuing a different set of building elements? What is the impact of litter or graffiti? Are experts better able to tune such elements out and focus more on frontage quality or building form? What nuances of streetscape are experts picking up that non-experts are not?

Ultimately, research results from studies like this should inform policy related to streetscape design, especially zoning or urban design regulations. One justification for this kind of research is that it might be possible to show that certain qualities of streetscape design actually improve occupants' mood and cognitive functioning (see, for example, Sussman and Justin 2021). If so, then understanding how to measure the sensory characteristics of the pedestrian realm could prove to be a powerful tool for enhancing mental health outcomes on a large scale, by designing built structure to increase human well-being (Ibarra et al. 2017). If mental health benefits are driven by predictable visual patterns, it may be possible to optimize the visual properties of the built environment to create more restorative spaces for human inhabitation.

Notes

1. A working demo of the survey is available at: https://users.rcc.uchicago.edu/~kywch/FIREst_201908_pilot/rating_preference.html.
2. More information on TrueSkill is available here: https://www.moserware.com/2010/03/computing-your-skill.html.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the University of Chicago [Social Sciences Research Center Faculty Seed Grant].

References

Adams, M. 2014. "Quality of Urban Spaces and Wellbeing." In *Wellbeing and the Environment*, edited by R. Cooper, E. Burton, and C. Cooper, 249–270. Vol. 2. West Sussex, England: John Wiley & Sons.

Adkins, A., J. Dill, G. Luhr, and M. Neal. 2012. "Unpacking Walkability: Testing the Influence of Urban Design Features on Perceptions of Walking Environment Attractiveness." *Journal of Urban Design* 17 (4): 499–510. doi:10.1080/13574809.2012.706365.

Al-Homoud, M., and L. Tassinary. 2004. "Social Interactions At The Neighborhood-Level As A Function Of External Space Enclosure." *Journal of Architectural and Planning Research* 21 (1): 10–23. Retrieved from: http://www.jstor.org/stable/43031056

Alexander, C. 2002. *The Phenomenon of Life: An Essay on the Art of Building and the Nature of the Universe*. Berkeley, Calif: Center for Environmental Structure.

Appleyard, D. 1979. "The Environment as a Social Symbol." *Ekistics* 46 (278): 272–281. Retrieved from: http://www.jstor.org/stable/43619666

Askari, A. H., and S. Soltani. 2018. "Contribution of Building Façades to Attractive Streetscapes: Study of Two Main Streets in Kuala Lumpur City." *Journal of Design and Built Environment* 18 (1): 29–40. doi:10.22452/jdbe.vol18no1.4.

Benfield, K. (2014). "What Makes a Good Main Street Work?" *What Makes a Good Main Street Work?* May 27. http://kaidbenfieldarchive.com/20140527-what-makes-a-good-main-street-work.html

Berman, M. G., J. Jonides, and S. Kaplan. 2008. "The Cognitive Benefits of Interacting with Nature." *Psychological Science* 19 (12): 1207–1212. doi:10.1111/j.1467-9280.2008.02225.x.

Braga, A. A., and B. J. Bond. 2008. "Policing Crime and Disorder Hot Spots: A Randomized Controlled Trial." *Criminology* 46 (3): 577–607. doi:10.1111/j.1745-9125.2008.00124.x.

Braga, A. A., D. L. Weisburd, E. J. Waring, L. G. Mazerolle, W. Spelman, and F. Gajewski. 1999. "Problem-oriented Policing in Violent Crime Places: A Randomized Controlled Experiment." *Criminology* 37 (3): 541–580. doi:10.1111/j.1745-9125.1999.tb00496.x.

Carmona, M., T. Gabrieli, R. Hickman, T. Laopoulou, and N. Livingstone. 2018. "Street Appeal: The Value of Street Improvements." *Progress in Planning* 126: 1–51. doi:10.1016/j.progress.2017.09.001.

Chae, B. G., and R. J. Zhu. 2014. "Environmental Disorder Leads to Self-regulatory Failure." *Journal of Consumer Research* 40 (6): 1203–1218. doi:10.1086/674547.

Clemente, O., R. Ewing, S. Handy, and R. Brownson (2005). *Measuring Urban Design Qualities—An Illustrated Field Manual*. Princeton, NJ: Robert Wood Johnson Foundation. https://activelivingresearch.org/sites/activelivingresearch.org/files/FieldManual_071605.pdf

Cooper, R., and E. Burton. 2014. "Wellbeing and the Environmental Implications for Design." In *Wellbeing and the Environment*, edited by R. Cooper, E. Burton, and C. Cooper, 653–668. Vol. 2. West Sussex, England: John Wiley & Sons.

Difallah, D., E. Filatova, and P. Ipeirotis (2018). "Demographics and Dynamics of Mechanical Turk Workers." *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*, 135–143. New York: Association for Computing Machinery (ACM).

Ellard, C. 2015. *Places of the Heart: The Psychogeography of Everyday Life*. Toronto: Bellevue Literary Press.

Ewing, R., and S. Handy. 2009. "Measuring the Unmeasurable: Urban Design Qualities Related to Walkability." *Journal of Urban Design* 14 (1): 65–84. doi:10.1080/13574800802451155.

Ewing, R., S. Handy, R. C. Brownson, O. Clemente, and E. Winston. 2006. "Identifying and Measuring Urban Design Qualities Related to Walkability." *Journal of Physical Activity and Health* 3 (Suppl 1): 223–240. doi:10.1123/jpah.3.s1.s223.

Gehl, J. 2011. *Life Between Buildings: Using Public Space*. New York: Van Nostrand Reinhold.

Gillem, M. 2009. "Make-Believe Main Streets: Hyperreality and the Lifestyle Center." *Traditional Dwellings and Settlements Review* 20 (2): 13–26. Retrieved from: http://www.jstor.org/stable/41758698

Gjerde, M. 2011. "Visual Evaluation of Urban Streetscapes: How Do Public Preferences Reconcile with Those Held by Experts?" *Urban Design International* 16 (3): 153–161. doi:10.1057/udi.2011.10.

Hartig, T. 2008. "Green Space, Psychological Restoration, and Health Inequality." *The Lancet* 372 (9650): 1614–1615. doi:10.1016/S0140-6736(08)61669-4.

Heffernan, E., T. Heffernan, and W. Pan. 2014. "The Relationship between the Quality of Active Frontages and Public Perceptions of Public Spaces." *Urban Design International* 19 (1): 92–102. doi:10.1057/udi.2013.16.

Herzog, T. R. 1992. "A Cognitive Analysis of Preference for Urban Spaces." *Journal of Environmental Psychology* 12 (3): 237–248. doi:10.1016/S0272-4944(05)80138-0.

Hollander, J. B., and E. C. Anderson. 2020. "The Impact of Urban Façade Quality on Affective Feelings." *Archnet-IJAR: International Journal of Architectural Research* 14 (2): 219–232. doi:10.1108/ARCH-07-2019-0181.

Ibarra, F. F., O. Kardan, M. R. Hunter, H. P. Kotabe, F. A. C. Meyer, and M. G. Berman. 2017. "Image Feature Types and Their Predictions of Aesthetic Preference and Naturalness." *Frontiers in Psychology* 8: 8. doi:10.3389/fpsyg.2017.00632.

Jacobs, A. 1995. *Great Streets*. Cambridge, MA: MIT Press.

Jacobs, J. 1961. *The Death and Life of Great American Cities*. New York: Random House.

Jones, C., Q. Al-Shaheen, and N. Dunse. 2016. "Anatomy of a Successful High Street Shopping Centre." *Journal of Urban Design* 21 (4): 495–511. doi:10.1080/13574809.2016.1192947.

Joye, Y. 2007. "Architectural Lessons from Environmental Psychology: The Case of Biophilic Architecture." *Review of General Psychology* 11 (4): 305–328. doi:10.1037/1089-2680.11.4.305.

Kaplan, R., and S. Kaplan. 1989. *The Experience of Nature: A Psychological Perspective*. Cambridge, UK: CUP Archive.

Kaplan, S. 1987. "Aesthetics, Affect, and Cognition: Environmental Preference from an Evolutionary Perspective." *Environment and Behavior* 19 (1): 3–32. doi:10.1177/0013916587191001.

Kaplan, S., and M. G. Berman. 2010. "Directed Attention as a Common Resource for Executive Functioning and Self-Regulation." *Perspectives on Psychological Science* 5 (1): 43–57. doi:10.1177/1745691609356784.

Kardan, O., E. Demiralp, M. C. Hout, M. R. Hunter, H. Karimi, T. Hanayik, and M. G. Berman. 2015. "Is the Preference of Natural versus Man-made Scenes Driven by Bottom–up Processing of the Visual Features of Nature?" *Frontiers in Psychology* 6. doi:10.3389/fpsyg.2015.00471.

Kashef, M. 2011. "Walkability and Residential Suburbs: A Multidisciplinary Perspective." *Journal of Urbanism* 4 (1): 39–56.

Kickert, C. 2016. "Active Centers – Interactive Edges: The Rise and Fall of Ground Floor Frontages." *Urban Design International* 21 (1): 55–77. doi:10.1057/udi.2015.27.

Kotabe, H. P., O. Kardan, and M. G. Berman. 2016. "The Order of Disorder: Deconstructing Visual Disorder and Its Effect on Rule-breaking." *Journal of Experimental Psychology. General* 145 (12): 1713–1727. doi:10.1037/xge0000240.

Krempen, M. 1974. "A Possible Analogy between (Psycho-linguistic and Architectural Measure—The Type-Token Ratio (TTR)." In *Psychology and the Built Environment*, edited by D. Canter and T. Lee, 87–95. New York, NY: Halsted Press.

Li, X., B. Y. Cai, and C. Ratti. 2018. "Using Street-level Images and Deep Learning for Urban Landscape Studies." *Landscape Architecture Frontiers* 6 (2): 20–29. doi:10.15302/J-LAF-20180203.

Lynch, K. 1981. *A Theory of Good Urban Form*. Cambridge: MIT Press.

Mantho, R. 2015. *The Urban Section: An Analytical Tool for Cities and Streets*. London: Taylor and Francis Group.

McGinn, A. P., K. R. Evenson, A. H. Herring, S. L. Huston, and D. A. Rodriguez. 2007. "Exploring Associations between Physical Activity and Perceived and Objective Measures of the Built Environment." *Journal of Urban Health* 84 (2): 162–184. doi:10.1007/s11524-006-9136-4.

Meidenbauer, K. L., C. U. Stenfors, G. N. Bratman, J. J. Gross, K. E. Schertz, K. W. Choe, and M. G. Berman. 2020. "The Affective Benefits of Nature Exposure: What's Nature Got to Do with It?" *Journal of Environmental Psychology* 101498.

Muhlebach, E., and R. Muhlebach. 2013. "The 'Malling' of American Retail and How to Re-Discover Street Retailing." *Journal of Property Management* 78 (3): 22–25.

Naik, N., J. Philipoom, R. Raskar, and C. Hidalgo (2014). "Streetscore – Predicting the Perceived Safety of One Million Streetscapes." *IEEE Conference on Computer Vision & Pattern Recognition Workshops 2014*, 793–799. doi:10.1109/CVPRW.2014.121.

Nasar, J. L. 1994. "Urban Design Aesthetics: The Evaluative Qualities of Building Exteriors." *Environment and Behavior* 26 (3): 377–401. doi:10.1177/001391659402600305.

O'Brien, D. T., C. Farrell, and B. C. Welsh. 2019. "Looking through Broken Windows: The Impact of Neighborhood Disorder on Aggression and Fear of Crime Is an Artifact of Research Design." *Annual Review of Criminology* 2 (1): 53–71. doi:10.1146/annurev-criminol-011518-024638.

Oreskovic, N., P. Roth, S. Charles, D. Tsigaridi, K. Shepherd, K. Nelson, and M. Bar. 2014. "Attributes Of Form In The Built Environment That Influence Perceived Walkability." *Journal of Architectural and Planning Research* 31 (3): 218–232. Retrieved from: http://www.jstor.org/stable/44114605

Park, K., G. Tian, and S. S. Larsen. 2019. "Street Life and the Built Environment in an Auto- Oriented US Region." *Cities* 88: 243–251. doi:10.1016/j.cities.2018.11.005.

Salingaros, N. A. 1998. "A Scientific Basis for Creating Architectural Forms." *Journal of Architectural and Planning Research* 15 (4): 283–293.

Salingaros, N. A. 2007. *A Theory of Architecture*. Solingen: ISI Distributed Titles.

Shen, Q., W. Zeng, Y. Ye, S. Müller Arisona, S. Schubiger, R. Burkhard, and Q. Huamin. 2018. "StreetVizor: Visual Exploration of Human-Scale Urban Forms Based on Street Views." *IEEE Transactions on Visualization and Computer Graphics* 24 (1): 1004–1013. doi:10.1109/TVCG.2017.2744159. January. 2018.

Soltani, A., M. Hosseinpour, and P. Zare. 2018. "The Development And Assessment Of Environmental Features Associated With Walkability Of Urban Streets." *Theoretical and Empirical Researches in Urban Management* 13 (1): 22–36. Retrieved from: http://www.jstor.org/stable/26302734

Southworth, M. 2005. "Designing the Walkable City." *Journal of Urban Planning and Development* 131 (4): 246–257. doi:10.1061/(ASCE)0733-9488(2005)131:4(246).

Speck, J. 2012. *Walkable City: How Downtown Can Save America, One Step at a Time*. New York: Farrar, Straus and Giroux.

Sussman, A., and B. H. Justin. 2021. *Cognitive Architecture: Designing for How We Respond to the Built Environment*. New York: Routledge.

Talen, E., Ed. 2015. *Retrofitting Sprawl: Addressing Seventy Years of Failed Urban Form*. Athens, GA: University of Georgia Press.

Talen, E., and J. Koschinsky. 2014. "Compact, Walkable, Diverse Neighborhoods: Assessing Effects on Residents." *Housing Policy Debate* 24 (4): 717–750. doi:10.1080/10511482.2014.900102.

Timms, P., and M. Tight. 2010. "Aesthetic Aspects of Walking and Cycling." *Built Environment (1978)* 36 (4): 487–503. doi:10.2148/benv.36.4.487.

Tomás, R., L. Hans, R. Víctor, and H. Ricardo. 2019. "Explaining Subjective Perceptions of Public Spaces as A Function of the Built Environment: A Massive Data Approach." *Landscape and Urban Planning* 181: 169–178.

Wang, R., Y. Lu, J. Zhang, P. Liu, Y. Yao, and Y. Liu. 2019. "The Relationship between Visual Enclosure for Neighbourhood Street Walkability and Elders' Mental Health in China: Using Street View Images." *Journal of Transport & Health* 13: 90–102. doi:10.1016/j.jth.2019.02.009.

Weuve, J., J. Kang, J. Manson, M. Breteler, J. Ware, and F. Grodstein. 2004. "Physical Activity, Including Walking, and Cognitive Function in Older Women." *Journal of the American Medical Association* 29212 (12): 1454–1461. doi:10.1001/jama.292.12.1454.

Wilson, J. Q., and G. L. Kelling (1982). "Broken Windows: The Police and Neighborhood Safety." *Atlantic Monthly* 249 (3), 29–38.

Xiangyuan, M., M. Chenyan, W. Chao, X. Yuliang, R. Yang, N. Peng, C. Zhang, and F. Ren. 2020. "Measuring Human Perceptions of Streetscapes to Better Inform Urban Renewal: A Perspective of Scene Semantic Parsing." *Cities* 110: 103086. doi:10.1016/j.cities.2020.103086.

Ye, Y., D. Richards, Y. Lu, X. Song, Y. Zhuang, W. Zeng, and T. Zhong (2019). "Measuring Daily Accessed Street Greenery: A Human-scale Approach for Informing Better Urban Planning Practices." *Landscape and Urban Planning*. https://doi.org/10.1016/j.

Yin, L., and Z. Wang. 2016. "Measuring Visual Enclosure for Street Walkability: Using Machine Learning Algorithms and Google Street View Imagery." *Applied Geography* 76: 147–153. doi:10.1016/j.apgeog.2016.09.024.

Zhang, F., B. Zhou, L. Liu, Y. Liu, H. H. Fung, H. Lin, C. Ratti, et al. 2018. "Measuring Human Perceptions of a Large-scale Urban Region Using Machine Learning". *Landscape and Urban Planning* 180: 148–160. doi:10.1016/j.landurbplan.2018.08.020.

Zhou, H., S. He, Y. Cai, M. Wang, and S. Su. 2019. "Social Inequalities in Neighborhood Visual Walkability: Using Street View Imagery and Deep Learning Technologies to Facilitate Healthy City Planning". *Sustainable Cities and Society* 50: 101605. doi:10.1016/j.scs.2019.101605.