Landscape and Urban Planning 203 (2020) 103885

Contents lists available at ScienceDirect

Landscape and Urban Planning

journal homepage: www.elsevier.com/locate/landurbplan

The condition-care scale: A practical approach to monitoring progress in vacant lot stewardship programs

Paul H. Gobster[a,*], Alessandro Rigolon[b], Sara Hadavi[c], William P. Stewart[d]

[a] USDA Forest Service, Northern Research Station, Evanston, IL 60201, USA
[b] Department of City and Metropolitan Planning, University of Utah, Salt Lake City, UT 84112, USA
[c] Department of Landscape Architecture and Regional & Community Planning, Kansas State University, Manhattan, KS 66506, USA
[d] Department of Recreation, Sport and Tourism, University of Illinois at Urbana-Champaign, Champaign, IL 61820, USA

ARTICLE INFO

Keywords:
Urban vacancy
Cues to care
Reliability and validity
Longitudinal analysis
Cross-sectional analysis
Program monitoring and evaluation

ABSTRACT

Condition and care are key expressions of landscape stewardship and are especially important in managing vacant urban lands. In this context, visible signs of stewardship have been associated with increased neighborhood sense of place whereas signs of physical disorder reflect perceived and actual crime. To date, assessments of condition/care and disorder using neighborhood audits have shown good reliability in field and virtual assessments but are often labor-intensive, particularly when repeated over time or applied across multiple study areas. Integrating the research traditions of neighborhood audits and visual landscape quality assessment, we propose an alternative approach, the "condition-care scale," and pilot test this seven-point rating scale to assess longitudinal and cross-sectional patterns of stewardship in response to a vacant land reuse program. Lots purchased by nearby residents through the Chicago Large Lot Program were rated on the scale using Google Street View imagery, field photography, and field visits in 2014 (before purchase), 2015, 2016, and 2018 (1–4 years after purchase). Lab and field assessments showed strong intra- and inter-rater scale reliability, and independent measures of lot condition and care from parallel visual and social assessments support the scale's validity. Longitudinal analyses showed that the greatest improvements were made in the first year after purchase but that improvement levels increased steadily over five years. Cross-sectional comparisons showed significant differences between some community areas. We discuss the utility of our approach for evaluating progress in vacant land reuse programs and its generalizability to other needs of urban greening professionals.

1. Introduction

The past two decades have seen an upsurge in the use of survey instruments to inventory and assess the visual, physical features of urban neighborhoods and streetscapes (Rundle et al., 2011). These neighborhood audits employ systematic observation protocols that tie theoretically relevant environmental features and qualities to broader, latent concepts such as physical disorder and walkability (Ewing et al., 2006; Mooney et al., 2014). Implemented by trained observers, audits composed of multiple item checklists and/or rating scales of visible features are seen as more objective and reliable than single-item scales of latent concepts, which are often seen as too subjective to be reliably assessed directly via expert or public self-report measures (Bader & Ailshire, 2014).

Audits can also be cost- and labor-intensive. Although virtual audits using imagery available through sources such as Google Street View have increased the efficiency and scope of applications (Hanibuchi, Nakaya, & Inoue, 2019), the date of available imagery is not always in sync with the data needs of the investigation (Bader et al., 2017; Cândido et al., 2018). Problems with the labor intensiveness of audits and syncing of imagery are compounded in intervention research, especially when there is a need for multiple assessments such as for monitoring changes over time or comparing different areas of a large city. To date, however, most virtual audits reported in the literature have provided general assessments of neighborhood quality at a single place and point in time and few have been implemented to evaluate design or policy interventions (Rzotkiewicz et al., 2018).

Although subjective rating scales have been avoided in neighborhood audits, they have been commonly used in visual landscape quality assessments (Gobster, Ribe, & Palmer, 2019). Like audits, many visual assessments also tie latent concepts to theoretically relevant physical features, but the concepts of landscape quality studied through visual

* Corresponding author.
E-mail addresses: paul.gobster@usda.gov (P.H. Gobster), alessandro.rigolon@utah.edu (A. Rigolon), sarahadavi@ksu.edu (S. Hadavi), wstewart@illinois.edu (W.P. Stewart).

https://doi.org/10.1016/j.landurbplan.2020.103885
Received 1 March 2020; Received in revised form 16 June 2020; Accepted 17 June 2020

assessments are usually rated directly by public groups or experts (Daniel, 2001). Integrating the research traditions of neighborhood audits and visual assessments, we have developed the condition-care scale, a single, seven-point rating scale, and have pilot-tested its utility for assessing the condition and care of city-owned vacant parcels transferred to private ownership under the Chicago Large Lot Program. Although our new scale does not replace the need for more intensive audits, after five years of use across five different Chicago community areas it has proven to be an efficient way for monitoring progress of the Large Lot Program. More generally, our new scale shows significant promise for involving planners in evaluation efforts to characterize the status of urban greening in neighborhoods with land vacancy.

In this paper, we describe the condition-care scale, report the results of tests assessing its reliability and validity, and demonstrate its utility for longitudinal and cross-sectional applications. As a tool for evaluating a policy intervention and monitoring parcel-level landscape change, our findings show continued incremental progress in greening activity, and in a few cases, significant differences between community areas. Although developed in the context of a specific program, the condition-care scale is adaptable to other vacant lot programs and holds promise for wider urban greening applications.

2. Background

2.1. A convergence of two research traditions: Neighborhood audits and visual assessments

The neighborhood has long been an important scale of concern in the urban studies literature (Green, Widener, Pollock, & Pearce, 2019), and today's neighborhood audits developed by social scientists and public health researchers stem in part from approaches pioneered in the 1920s by the Chicago School of Sociology (Sampson, 2012). This research tradition emphasizes expert-based observation over public self-report surveys, the importance of the physical environment in determining behavioral outcomes, and the understanding of crime and other social problems via the measurement of latent concepts such as physical and social disorder (Sampson & Raudenbush, 1999). Contemporary audits operationalize concepts such as physical and social disorder through careful measurement of physical features, conditions, and behaviors. An often-cited example of such an approach is Sampson and Raudenbush (1999, 2004), who, in addressing Wilson and Kelling's (1982) "broken windows" theory of urban decline, developed a neighborhood auditing instrument based on systematic social observation of 126 block-level indicators inventoried by trained observers. From this work, they created a 10-item scale of physical disorder (e.g., abandoned cars, graffiti) and a 7-item scale of social disorder (e.g., public drinking, drug dealing).

The use of neighborhood audits has expanded greatly in recent years, aided by freely available digital imagery through platforms such as Google Street View (Bader et al., 2017). These virtual audits compare favorably with field assessments with respect to reliability while reducing the time and effort required. For example, to examine physical disorder and walkability, Bader et al. (2015) developed the 187-item Computer Assisted Neighborhood Visual Assessment System or CANVAS. Using physical disorder as a guiding concept to evaluate neighborhoods for health promotion and crime prevention, Reischl et al. (2016) developed the Parcel Maintenance Observation Tool or PMOT by measuring 10 features and 5 qualities of maintenance condition at the parcel (lot) level. And, with concerns to encourage outdoor physical activity in low-income neighborhoods of color, Franzini et al. (2010) used feature checklists and ratings of 161 different block-level variables to measure four indicator concepts—pleasurability, comfort, safety, and accessibility.

A parallel research tradition for understanding the effects of the environment on people has been forged by landscape architects, environmental psychologists, and others through visual landscape quality assessments (Daniel, 2001). Concern for the visual quality of landscapes grew out of the environmental movement of the 1960 s and aimed at identifying key dimensions of landscape perception and preference such as scenic beauty (Gobster, Ribe, & Palmer, 2019). This emphasis on subjective environmental qualities stands in contrast to the focus of audits on objective behaviors, as does the use of self-report ratings to measure landscape perceptions and preferences. Otherwise, the two traditions share much in common. Like audits, many visual assessments use inventories and models of physical environmental features to explain or predict landscape perceptions (Daniel & Boster, 1976; Kaplan & Kaplan, 1989). These are assessed by trained observers and public groups, with similar concerns for measurement reliability (e.g., Craik & Feimer, 1979; Palmer & Hoffman, 2001). For cost and convenience, many visual assessments also rely on virtual landscape surrogates as proxies for field ratings, with validation studies showing good equivalence (Sevenant & Antrop, 2011; Shuttleworth, 1980). As a final point of comparison between these two research traditions, while most visual assessments have been implemented in rural and wildland contexts, some of the earliest applications shared the concern with audits of a focus on the urban landscape (Lynch & Rivkin, 1959; Peterson, 1967).

Contemporary urban visual assessments address a broader set of outcomes beyond aesthetic preferences. Many have examined the visual characteristics of neighborhoods for safety purposes (Jorgensen, Hitchmough, & Dunnett, 2007; Lis et al., 2019) and of yards and gardens for ecological functioning and aesthetic purposes (e.g., Lindemann-Matthies & Marty, 2013; Nassauer, Wang, & Dayrell, 2009). Condition and care are key expressions of landscape stewardship and have been found to be important in urban settings where norms of neatness and beauty converge (Larson et al., 2016; Visscher, Nassauer, & Marshall, 2016). Although much of this work has focused on suburban residential landscapes, condition and care are also critical elements in the revitalization of high-vacancy neighborhoods (Nassauer & Raskin, 2014). In this context, condition is associated with degree of maintenance, and vacant open lots that are viewed as attractive have well-maintained vegetation and built features that convey cleanliness, order, and neatness (Morckel, 2015; Rega-Brodsky, Nilon, & Warren, 2018). Care adds an expressive component to landscape condition, and "cues to care" such as lawn ornaments, gardens, and outdoor furniture not only show how people engage with the land as owners and stewards, they communicate this intent to others (Nassauer & Raskin, 2014). Although maintenance condition and care refer to two seemingly distinct concepts, we view care as a higher-order extension of the maintenance continuum, where personal and community values are made visible through stewardship activities on the land.

2.2. Condition and care in the context of neighborhood change

Our review of these two research traditions – neighborhood audits and visual assessments – reveals a convergence of concepts and approaches. Conceptually, Nassauer's visual assessment-based cues to care framework (Nassauer, 1995, 2011) has been likened to the inverse of Wilson and Kelling's (1982) broken windows theory used in neighborhood audits by Sampson and Raudenbush (2004) and others (Troy, Nunnery, & Grove, 2016) in that the former deals with expressions of beauty, stewardship, and human presence, while the latter deals with blight, disorder, and abandonment. Methodologically, some neighborhood audits have incorporated expert-rated subjective scales of environmental qualities such as aesthetics and safety common to visual assessments in order to evaluate behavioral outcomes (e.g., Ewing et al., 2006; Gidlow et al., 2018; Gullón et al., 2015; Kamphuis et al., 2010). Likewise, some urban visual assessment have adopted expert-based observation of objective physical characteristics common to neighborhood audits to make inferences about landscape perceptions (e.g., Evans-Cowley & Akar, 2014; Li et al., 2015; Tang & Long, 2019).

In this respect, an important recognition of both research traditions is that, while it is important to know what people think about places and planning issues, in some cases their perceptions may not accurately reflect observed behavior or conditions. A commonly cited example is the difference between perceived and actual crime. Both measures are critical in understanding people–place relationships but they show low correlations with each other in terms of location and frequency of occurrence, as well as differences by gender and other demographic variables (e.g., Ambrey, Fleming, and Manning, 2014; Baldock et al., 2018; Kim, Sweithzer, & Kim, 2002). Closer to our own study, Sampson and Raudenbush (2004) found that while perceived disorder predicted observed disorder, implicit biases about race and class led residents of all races to perceive poor, black neighborhoods as having more physical disorder than actually existed. A similar disjunction has been shown in research on landscape care, where stewardship activities such as community gardening, which are intended by their participants to be expressions of care and understood by experts as such, can be perceived by non-participants as messy and neglectful if the right cues are not present (Nassauer, 1995). For these and other reasons, approaches to neighborhood assessment for issues such as urban vacancy tend to employ expert-based measures that are felt to be more ecologically valid (Sampson & Raudenbush, 2004). Thus, while each of these traditions has a distinct history and origin, their several points of convergence provide insights for our development of the condition-care scale to monitor progress in vacant lot stewardship.

This blending of traditions is evident in two studies of urban vacancy that were influential in our work. For their Detroit Sustainability Indicators Integrated Assessment, Dewar and colleagues developed a protocol to assess the dynamics of care for two Detroit neighborhoods. They first identified 95 different cues to care and neglect (Dewar, Nassauer, & Dueweke, 2013) then subsequently defined a subset of 15 essential cues to care that included the presence/absence of features such as dumping, hedges, seating, and flowers and scaled ratings for conditions of structures and turf (Dewar, Nassauer, Dueweke, & Boswell, 2015). Similarly, Krusky et al. (2015) used a subset of the 15-item physical disorder auditing tool from Reischl et al. (2016) mentioned above to examine how maintenance levels of residential yards were affected by their proximity to maintained community gardens or unmaintained vacant lots. The three-item subset included ratings for the recency of mowing, level of shrub and garden maintenance (landscaping), and amount of litter and trash. In both studies, the reduction of attributes and the inclusion of both objective features and subjective qualities indicate a refinement of approaches to address targeted questions about urban vacancy.

Building on the experience of these and similar efforts, we developed a 20-item parcel-level virtual auditing instrument to assess the condition and care of vacant lots purchased under the Chicago Large Lot Program using ground and aerial imagery from Google and Bing. From these 20 measures, we identified a subset of seven items to develop an index characterizing lot-level condition and care. Because our focus was on evaluating changes made as a result of program participation, we needed to apply the protocol before and after lots were transferred to lot purchasers. As detailed in another paper (Gobster, Hadavi, Rigolon, & Stewart, 2020a), we eventually acquired the post-purchase aerial imagery to conduct our full audit, but our work was limited and delayed by the lack of available aerial and street-level virtual imagery for dates following the program's implementation. The full audit, while providing useful information on the types of changes that occurred as a result of a program intervention, was also labor-intensive for repeated monitoring over time and application across several community areas. Because visual assessments have long used single rating scales to quantify a variety of environmental qualities, we felt a simple condition-care rating scale held high potential for use in studying urban vacancy (Schroeder, 1984), and no similar scale that we were aware of had yet been reported in the literature. Additionally, both neighborhood audits and visual assessments have long relied upon field study, and because of the uncertainty over the availability of time series imagery through Google and its equivalents (Bader et al., 2017; Rzotkiewicz et al., 2018), a combination of virtual and field imagery seemed the most expedient solution to our monitoring needs.

In the sections that follow, we detail our efforts in the development, testing, and application of the condition-care scale. The specific objectives of our paper were to:

1. Describe the scale and approach to its implementation.
2. Assess the scale's reliability and validity.
3. Pilot-test the scale over a five-year period in five different community areas to demonstrate its applicability to addressing longitudinal and cross-sectional issues in program monitoring and evaluation.
4. Evaluate the utility of the scale and implementation approach with respect to efficiency, usefulness, and generalizability to other urban vacancy and broader urban greening applications.

3. Methods

3.1. Program and pilot community areas

The Chicago Large Lot Program grew out of a 2014 Green Healthy Neighborhoods Plan developed by the City of Chicago in collaboration with local community partners. The plan addressed vacancy issues within an area on the city's south side composed of four primary community areas and portions of four other community areas, hereafter referred to as the GHN area (City of Chicago, 2014a) (Table 1; Fig. 1). Piloted later that year in the GHN area and East Garfield Park, another community area on the west side of the city also experiencing high land vacancy, the program offers qualified property owners the opportunity to purchase one or two city-owned vacant residential lots ("large lots") on their block or the adjacent block for $1 each. The purchaser must agree to maintain the lot, fence it if it is not immediately side-adjacent to their property, and pay the property taxes for five years, after which they gain full rights of ownership and sale (City of Chicago, 2014b). Although building on the lots is permitted and encouraged, it was expected that in the early years of ownership most purchasers would use their lots as private or shared green space for yard and garden expansion, along with complementary hardscape areas for play, social activity space, or off-street parking.

Successful applicants took possession of their large lots in early 2015, and later that year, we were asked by the city to conduct a program evaluation of lot purchases and lot owners in five community areas, including four in GHN and East Garfield Park. Our study included two main components: a visual assessment of changes made in lot condition and care, and a social assessment of owners' uses and perceptions of their lots and neighborhoods (Stewart et al., 2019). For the visual assessment, our principal unit of study was the purchased large lot, 424 lots across the five community areas. Because applicants could purchase one or two lots, the social assessment sampling frame was the 321 lot owners, and our mail survey resulted in a final sample size of 197 owners (71% response rate for 278 valid mailing addresses).

3.2. The condition-care scale

As shown in Table 2, the seven-point ordinal scale ranges from -1 to $+5$, with a detailed description summarizing each level. Of principal concern to many vacant land resale programs is compliance with Level 1 "basic management" provisions as stipulated under city ordinances. For Chicago, these ordinances translate to periodic mowing so that ground cover vegetation does not exceed 10 in. (25 cm) in height and that the lot be reasonably kept free of trash and other materials considered as nuisance, visually or otherwise. Below this baseline are levels labeled "unmanaged" (0) and "mismanaged" (-1) that represent substandard conditions, while above it, level 2 "regularly managed" corresponds to a broadly accepted social norm for residential yards where turf areas appear recently mown and substantially weed-free (Sisser et al., 2016). Together, these four lower levels constitute the condition

Table 1
Demographic, housing, land use and environment statistics for community areas in this study and City of Chicago.[1]

Indicator	East Garfield Park	Englewood	West Englewood	Washington Park	Woodlawn	City of Chicago (all areas)
Demographic						
Population	20,004	25,075	29,929	11,502	25,207	2,722,586
Households	6,819	9164	9,642	4,334	10,481	1,046,789
Avg. household size	2.9	3	3.1	2.7	2.3	2.6
% Black Non-Hispanic	88.7	94.6	91.2	93.5	82.8	30.1
Median Age (yrs.)	30.6	33.3	37.4	28.9	34.4	34.1
% HS Grad or higher	79.4	73.4	78.3	82.8	86.7	83.8
% College Grad or higher	14.3	7.2	8.0	18.4	30.2	37.5
Median Annual Household Income	$24,854	$21,581	$28,436	$27,423	$26,479	$52,497
% unemployed	18.2	34.2	31.3	23.9	17.2	8.3
Housing						
% Owner-occupied units	18.5	16.2	34.8	10.3	18.3	38.9
% Renter-occupied units	63.3	48.2	40.0	66.1	59.5	48.3
% Vacant units	18.2	35.6	25.0	23.6	22.2	12.8
% Occupied single family (detached + attached)	15.6	26.9	56.0	10.1	11.7	25.9
% Occupied 5 + units	30.6	27.1	7.4	63.4	57.2	40.9
Land use area						
Total land area (ac)	1,237	1,966	2,019	973	1,325	147,949
% Single family residential	4.8	12.1	24.8	1.6	4.8	19.9
% Multi-family residential	16.8	15.0	9.9	13.2	19.1	11.5
% Open space	14.8	15.6	4.1	35.9	27.0	7.1
% Vacant	13.7	21.5	13.6	17.3	9.2	5.1
Physical/social environment						
% canopy cover	17.5	23.1	25.9	18.3	21.0	17.2
% impermeable	51.7	51.1	51.2	42.4	46.1	58.0
2018 violent crimes per 100 k people	3,505	2,453	2,487	3,089	1,701	955
Vulnerability rank	15	8	9	23	28	NA
Distance from CBD (km)	6.5	12.2	12.6	10.4	11.9	NA

[1] Demographic and housing data from US Census Bureau American Community Survey 2013–2017 5-year estimates; land use data from Chicago Metropolitan Agency for Planning Parcel-Based Land Use Inventory 2013 (https://www.cmap.illinois.gov/data/land-use/inventory); canopy cover, impermeable surface, and vulnerability rank from Chicago Region Trees Initiative Priority Maps (https://mortonarb.maps.arcgis.com/); crime data from Chicago Health Atlas (https://www.chicagohealthatlas.org); distance to CBD from Google Earth. Statistics in this table for the Englewood, West Englewood, Washington Park, and Woodlawn community areas do not include data from the partial community areas of Fuller Park, Greater Grand Crossing, New City, and Grand Boulevard that are part of the Green Healthy Neighborhoods planning area (City of Chicago, 2014) and included in the community area cross-sectional comparison (see text for details).

end of the scale, where maintenance is the primary focus. Levels 3–5 relate to increasing expressions of care; including small to extensive gardens, improvements in fencing, and other social, aesthetic, or recreational signs of human presence; and a subordination of vehicles (e.g., Nassauer & Raskin, 2014).

As devised, the condition-care scale is principally concerned with the green space value of a large lot. However, we also recognized that some large lots have been and will continue to be used principally as parking lots for multiple vehicles, and for these special cases ($n = 12$) we created a parking lot subscale (see Appendix S1) with parallel descriptive criteria. While these ratings were included in the analysis, there were also a few lots developed for housing in 2016–2018, for which ratings were discontinued.

3.3. Implementation procedures

The first application of the scale was made on a comparison of large lot photo-pairs taken in 2014, the year before purchase, and 2015, one growing season after purchase. The purchased lots were identified by their street address on the Large Lot Program website (LargeLots.org, n.d.) and from there a link to the Cook County Property Tax Portal (Cook County, n.d) provided an older (circa 2007) street-level photograph of each lot and aerial image showing parcel property boundaries and location on the block. With this information, we could accurately locate the parcel to download recent imagery and identify the lot in the field.

For the pre-purchase photography, we had Google Street View imagery for 80% of the lots taken between June and November 2014, just before lots were purchased. For the remaining lots, we used lower-resolution Bing Streetside images dating from May 2014 or earlier Street View images. The post-purchase field photography was taken in September and October of 2015, after the first growing season of large lot ownership but while vegetation was still green. The purchased large lots were dispersed throughout the five community areas, and to facilitate fieldwork we developed field sheets for locating parcels and route maps dividing the areas into manageable subunits (13–36 lots each) that could be traversed by car without having to cross major roads.

Field days were scheduled for days of good weather, with most of the activity taking place from 9 AM to 3 PM on weekdays to minimize interference with traffic and harsh sun/shadow contrasts. The field crew consisted of two or three members who shared driving, navigation, and photographing duties. The photographer captured an eye-level view of the entire lot from the sidewalk or street and took extra photos to capture any distinctive features. Crew members recorded any notes next to the Street View photo of changes observed or conversations with neighbors. Using this procedure, the crew photographed all 424 large lots in the five community areas in 27 h over five days, averaging about four minutes per lot including the driving time between lots.

For rating the images, a research assistant was trained in the application of rating scale criteria (Table 1) and reviewed a training document showing examples of lots for each level of the scale (Appendix S1). Rating instructions emphasized that not all of the criteria mentioned under each value needed to be present, and if there were inconsistencies the rater should apply the rating level that best fit the description. For example, if a lot had extensive gardens (rating = 5) but did not show signs of recent care and looked weedy (rating = 1), the more permanent, infrastructural improvements would take precedence over the more temporary condition, with the inconsistency

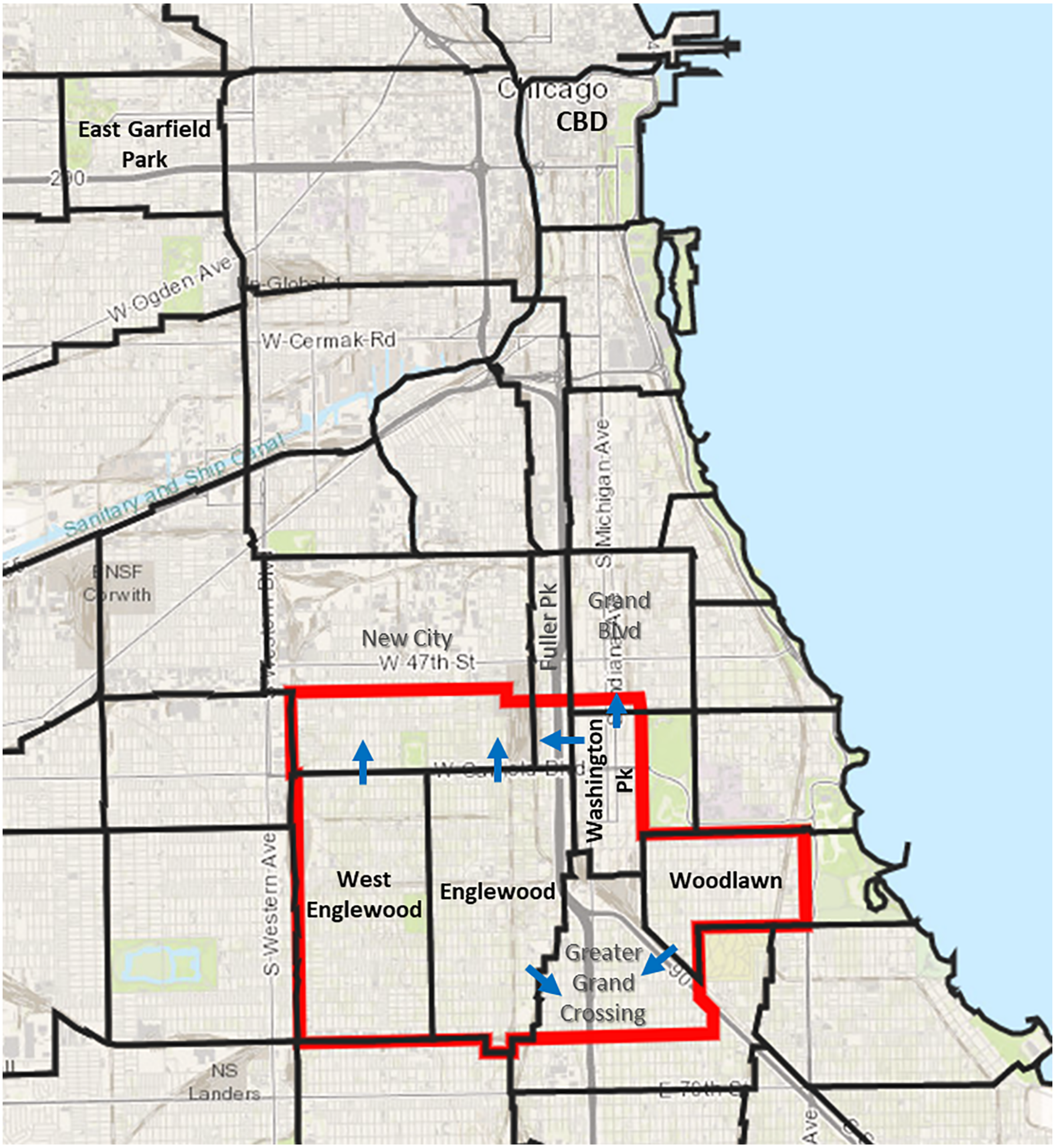

Fig. 1. Portion of the City of Chicago showing the five west (East Garfield Park) and south side (West Englewood, Englewood, Washington Park, and Woodlawn) community areas studied. Arrows point to fractional community areas merged into the four primary community areas within the Green Healthy Neighborhoods planning area (red boundary) for the cross-sectional analysis (see text Section 3.3.2 for details). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

noted in a side box on the scoring form. Informed by the criteria and examples, the rater made a single, gestalt judgment from −1 to +5. Pre- and post-purchase ratings were made at the same time, with corresponding images for a given lot displayed side-by-side on a widescreen monitor.

Field photography for the 2016 season was collected by the City of Chicago, Department of Planning and Development's Large Lot Program staff. The staff was similarly trained and used updated copies of the field sheets and route maps developed for the 2015 season. The three-person field crew collected the imagery in August and early September. The images were sent to the research team for labeling and the research assistant rated each lot on the condition-care scale, using the 2015 image for comparison.

For the 2018 season, we worked with Large Lot program staff and college juniors from their summer internship program. With five people available, we formed two crews and deployed in adjacent subunits of a community area. Along with photographing the lots, for the first time we also applied the rating scale in the field. Each crew member had a set of pre-addressed forms with columns for ratings and notes, and made independent ratings for each lot. After photographing and rating a lot, the members discussed their ratings with each other and agreed upon a consensus rating. If a consensus could not be reached, the averaged value of the ratings was used for the final score for the lot. Fieldwork for the two crews was completed in 19 h over four days in mid-July, an average of five minutes per lot, including driving between lots. In addition to the training document, Appendix S1 also provides

Table 2
The condition-care scale as applied in this study.

Rating	Descriptive criteria
−1	**Mismanaged:** Ground and tree cover is unmanaged and use by authorized or unauthorized parties is largely uncontrolled (see "unmanaged" criteria detail for more detail). In addition, there are signs where lack of control over use has resulted in abuse and safety hazards including dead trees (hazardous branches), serious erosion including potholes and large puddles from vehicle use, vehicle paths cutting through the property and off the curb into the street, junked and abandoned vehicles, dumping and large amounts of litter, vandalism and graffiti, and/or structures in obvious disrepair.
0	**Unmanaged:** Ground cover vegetation is largely weedy and not recently mown (average is > 10″ tall). Significant portions of the lot may be devoid of ground vegetation and either bare soil or broken pavement (former use). If present, shrub and tree cover is weedy and unmanaged, and there may be dense clumps of volunteer woody plants and/or malformed trees, especially along perimeter areas. Lot is unfenced, may have vehicles parked on it with areas of erosion and potholes/puddles, and litter may be prevalent. Overall there appears to be little control over use, but there are no signs of abuse or safety hazards.
1	**Periodically Managed** ("Basic Management" – minimum standard under City ordinances): Ground cover vegetation is periodically mown (1-2x/season) but may be weedy and somewhat tallish at the time photo was taken (though < 10"). May have some small volunteer trees and shrubs, especially along perimeter, but larger woody vegetation shows some signs of pruning and control. Lot may be unfenced or with bollards or older fencing, and some fencing may be in need of repair. The lot may have cars parked on it, patches of dirt from car use, and small amounts of litter, but signs of control over use are evident. For example, cars do not use the property as a cut-through between alley and street, and there may be some signs of infilling potholes and low spots to minimize erosion damage from vehicles.
2	**Regularly Managed:** Ground cover has been recently mowed and is predominantly lawn grass, and trees and shrubs if present are planted and/or pruned to have good form. The property may be unfenced or fenced, and if fencing is older it is in reasonable repair. No gardens are present, and other social, aesthetic, or recreational signs of occupancy are minimal. There may be cars parked on the property (in the rear and in some cases middle or front) but they are clearly "in their place" and do not dominate the view, and efforts are made to minimize any erosion due to vehicle use.
3	**Small-scale gardens and other cues to care:** Flowers in pots or a small garden patch is visible, usually along the front of the lot, lot corner, or around a tree (including the parkway strip). Trees and shrubs if present appear to be planted and are maintained in good form. Fencing if present is in good condition. Other social, aesthetic, or recreational signs of occupancy may be present but are small in scale. One or two vehicles may be visible, but they are in a designated space at the rear of the property.
4	**Moderate-scale gardens and other cues to care:** Flower or vegetable beds are present but are moderate in scale in relation to lot size. Trees and shrubs if present appear to be planted and are maintained in good form. Fencing if present is in good condition. Other social, aesthetic, or recreational signs of occupancy are present but are moderate in scale or number. Vehicles if present are in a designated space at the rear of the property and efforts are made to separate them from the rest of the property, e.g., screening with fencing or vegetation.
5	**Extensive gardens and other cues to care:** Flower or vegetable beds and other plantings including trees, shrubs and other ground cover dominate the lot. Fencing if present is in good condition. Numerous other social, aesthetic, or recreational signs of occupancy may be present. Vehicles are not visible.

illustrated examples of the primary implementation procedures described in this section.

3.4. Analysis and modeling

3.4.1. Reliability and validity tests

We assessed the external reliability of the condition-care scale using measures of test-retest and inter-rater reliability (American Psychological Association, n.d). For the 2014, 2015, and 2016 photo ratings, the research assistant re-rated each entire set of lots at least one year after the initial ratings were applied. Test-retest reliability was assessed with the Intraclass Correlation Coefficient (ICC) using a single measure, 2-way random-effects model with absolute agreement (Howell, 2011). We also compared the 2014 and 2015 ratings from the research assistant with one of the research principal investigators for a subset of images (n = 149), and the 2018 ratings between field team members. This test of inter-rater reliability was assessed with the ICC using a single measure, 2-way mixed-effects model with absolute agreement (Koo & Li, 2016).

We used two independently-derived measures from our broader study to assess the convergent validity of the condition-care scale (Chin & Yao, 2014). From the full visual assessment auditing instrument (Gobster et al., 2020a), we used the condition-care index, a summated scale of seven binary-coded variables indicating the presence/absence or condition of street-level features, which were assessed independently of the condition-care scale. The seven items in the index included the condition of lot pavement, shrubs and small trees, mature trees, and fencing (0 = absent or not in good condition, 1 = good condition); and the presence of cues to care including gardens, yard ornamentation, and social/recreational uses (0 = absent, 1 = present). The index was assessed for the same images/parcels used in the 2014 and 2015 assessments described above. The seven-item index showed acceptable internal consistency (Cronbach's alpha = 0.753 before purchase and 0.698 after purchase).

From the mail survey used for the social assessment (Stewart et al., 2019), we used a summated index of nine, binary-choice items from a question that asked large lot owners what they had done with their lot since they purchased it. Items in the index included mowing, cleaning up, removing trees or shrubs, installing fencing, filling sunken areas, planting trees and shrubs, planting flowers, planting vegetables, and making a sitting or play area (Cronbach's alpha = 0.697). The survey was implemented in summer 2016, and respondents who had purchased two large lots were instructed to answer the questions with respect to the lot they had improved and cared for the most. Because they did not provide the address of their most improved lot, in these instances (n = 56, 29%) we chose the lot closest to their previously owned lot, or in cases of a tie, flipped a coin.

Because each of these indexes measured similar aspects of lot-level condition and care, we expected that moderate-to-high correlations would be evidence of convergent validity (Gregory, 2007). Furthermore, we expected that values of the condition-care scale would correlate most highly with values of the alternative measures assessed during the same year. Comparisons were made using Pearson correlations (two-tailed significance test), and all validity and reliability tests were conducted using IBM SPSS Version 26.

3.4.2. Longitudinal and cross-sectional applications

To examine longitudinal changes in large lot condition and care, we used hierarchical linear modeling (HLM) to account for repeated measures of the condition-care scale (2014, 2015, 2016, and 2018) and because of HLM's flexible assumptions of normality (Raudenbush & Bryk, 2002). To measure this temporal sequence, we created binary variables for each assessment period (time 1, time 2…) and used the 2014 period (time 1) as the reference condition. In the HLM model, the time variables were treated as fixed effects, and the lot ID number was treated as a grouping factor (random-effect intercept). In this mixed-effects model, time represents the independent "treatment" variable, and lot condition and care the dependent variable. The model was estimated using maximum likelihood ratio and t-tests of independent variable significance used Satterthwaite's formula. Model fit was examined using the Akaike Information Criterion (AIC). Post-hoc tests were conducted to test for significant differences between paired estimated marginal means for the different time periods. Significance was estimated using Tukey-adjusted p values. The HLM model was estimated with RStudio, using the lmer package, and post-hoc tests were estimated with the emmeans package (R Core Team, 2013).

We also pilot-tested the scale cross-sectionally to examine differences in condition and care across community areas (see Table 1, Fig. 1) at single points in time. Specifically, we compared five community areas, including East Garfield Park and the four primary community areas of Englewood, West Englewood, Washington Park, and Woodlawn included within the GHN area. Chicago's 77 community areas are defined by natural geographic divisions and were developed in the 1920s by the same Chicago School of Sociology mentioned earlier that pioneered the neighborhood audit. Community areas were used by the Chicago School to describe and compare community social conditions and continue to be used by the city and other institutions for planning purposes, as they tend to have an identifiable sense of place and are more stable than other politically drawn boundaries such as aldermanic wards (Seligman, 2005).

The GHN area also includes adjacent fringe portions of four other community areas, and because these areas had a substantial number of large lots in our sample (n = 60 or 15% of all lots and 22% of GHN lots), we merged them with the nearest primary GHN community areas as shown by the arrows in Fig. 1. In nearly all cases, these fringe lots were within 3–4 blocks (less than 0.5 km) of the closest boundary of the primary community area. Except for one pair at one time period, none of the independent samples t-tests we conducted showed significant differences between the pairs. Thus, we were confident that merging the data would not obscure cross-sectional comparisons between the five community areas.

With this recoding, we created a five-category community area variable and ran one-way ANOVAs in SPSS for each assessment year, using pairwise post-hoc tests to look for significant differences in the estimated marginal means of condition-care scale ratings between the five community areas.

Finally, we pilot-tested the scale to examine joint longitudinal-cross sectional effects of time and community area. For this analysis, we used HLM, following the same steps described above, and included binary-coded variables for the community areas as fixed effects in the model, using one community area (West Englewood) as the reference variable. We also fitted an additional HLM where we added interaction terms between time and community areas, with the primary purpose of creating a plot of estimated marginal means across time and place.

4. Results

4.1. Reliability and validity

ICC values for test–retest reliability were 0.914 for the 2014 sampling period, 0.922 for 2015, and 0.930 for 2016, an average of 0.922. ICC values of assistant/investigator inter-rater reliability were 0.865 for the 2014 set and 0.884 for the 2015 set. For the 2018 field ratings, ICC values for inter-rater reliability were 0.967 and 0.968 for the two-person teams and 0.977 and 0.982 for the three-person teams. In all cases, ICC values indicated "excellent" agreement in application of the scale (Gullón et al., 2015; Landis & Koch, 1977).

For the validity assessment, Table 3 shows the correlation matrix, where the matches in the year the measures were assessed are highlighted in bold. Except for 2016, the highest correlations between measures were for the matching years as expected, and overall, moderate correlations in the r = 0.5–0.6 range support the validity of the scale.

Table 3
Validity matrix for the condition-care scale.

Alternative condition-care measures	Condition-care scale (r)			
	2014	2015	2016	2018
7-item audit index, 2014 (N = 419)	**0.610*****	0.544***	0.493***	0.363***
7-item audit index, 2015 (N = 391)	0.345**	**0.552*****	0.665***	0.540***
9-item mail survey index 2016 (N = 188)	0.182*	0.432***	**0.518*****	0.482***

* p < .05; *** p < .001

Table 4
HLM model of large lot condition-care for longitudinal application.

	Estimate	95% Conf. Int.	df	t	p
Fixed effects					
(Intercept)	1.120	0.994–1.247	793.479	17.403f	< 0.001
Time2 (2015)	0.521	0.410–0.632	1246.535	9.205	< 0.001
Time3 (2016)	0.705	0.591–0.818	1250.533	12.211	< 0.001
Time4 (2018)	0.925	0.813–1.036	1246.862	16.302	< 0.001
Random effects					
Intercept variance (lot ID)	1.077				

N = 424. AIC = 4945.4, marginal R^2 = 0.063. Tukey adjusted post-hoc tests were significant for all pairwise comparisons between time variables (Time1 (2014) reference condition), p < 0.001 except for Time2 - Time3, p = .008.

4.2. Pilot testing

Table 4 presents the HLM model for the longitudinal application. Each treatment variable for time is significant (p < .001) and post-hoc pairwise tests show that the estimated marginal means for each year are all significantly different from each other. The estimated marginal means plot (Fig. 2) shows that the greatest increase in lot condition and care comes in the first year after purchase (2015), but also that improvements as assessed with the condition-care scale continue in successive years, albeit at a more modest pace. Fig. 3 provides an example of a large lot from the study that illustrates this pattern of longitudinal improvement over the five-year assessment period.

For the cross-sectional application, one-way ANOVAs and post-hoc comparisons for individual years show that the West Englewood community area consistently had significantly higher condition-care ratings than Englewood and East Garfield Park (Table 5, Fig. 4). This pattern persists in the joint longitudinal and cross-sectional HLM model (Table 6), and adding variables describing community area leads to better model fit than the HLM model with time only, as indicated by the decreased AIC (compare Tables 4 and 6). Post-hoc pairwise tests confirm the findings from the cross-sectional ANOVAs while controlling for the effect of time. The estimated marginal means plot of the joint effect of time and community area for a model with time-community area interaction terms (Fig. 5) also shows that, except for one time period (2018) for one community area (Washington Park), the patterns of change appear consistent. These patterns support the stability of the scale's application in comparing community areas across time. Furthermore, the plots in Figs. 4 and 5 show that while the West Englewood community area consistently had significantly higher condition-care ratings than Englewood and East Garfield Park, the lots in each of

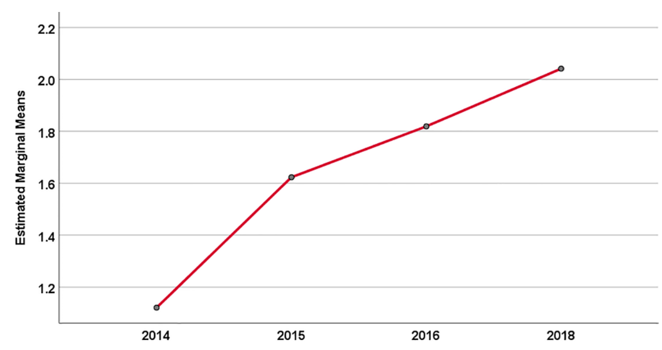

Fig. 2. Plot of estimated marginal means for condition-care scale ratings as a function of time (longitudinal application).

Fig. 3. Example of lot from the West Englewood community area exhibiting a "typical" pattern of continued improvement over time. Upper left, 2014 before purchase, rating of "1" corresponding to "basic management"—lot is mowed but shows areas of bare ground from unmanaged parking (photo: Google Street View). Upper right, 2015 one year after purchase, rating of "2," owners have begun fencing lot and have removed volunteer weed trees from alongside house (photo: US Forest Service). Lower left, 2016 two years after purchase, rating of "4," owners have completed fencing and made major infrastructural improvements including a gazebo and parking space in rear of the lot (photo: City of Chicago). Lower right, 2018 four years after purchase, owners continue to improve lot a vegetable garden, lawn ornaments, children's play equipment, and other social and recreational cues to care (photo: US Forest Service). For additional examples, see Supplementary Appendix S1.

the community areas continued to improve over time as measured by the condition-care scale.

A sampler of large lot images provided in Appendix S2 shows a range in the types and patterns of changes made to individual parcels in the five community areas over the five-year study period. While the means plot for the longitudinal assessment and example lot shown above (Figs. 2 and 3) illustrate the "typical" pattern of an increase in condition and care over time as averaged over all lots, other sub-patterns were noticeable and are reflected in the sampler. These sub-patterns included lots where changes occurred in the first year only, where no improvements were made until the later monitoring periods, where no changes occurred at all, and a pattern where declines in condition-care occurred over time.

5. Discussion

5.1. Reliability and validity

We formally assessed the reliability and validity of the condition-care scale, and our implementation also informally examined other

Table 5
One-way ANOVAs of condition-care values for the five community areas in the cross-sectional application.

		Sum of Squares	df	Mean Square	*F*	*p*
2014 before purchase	Between Groups	23.522	4	5.881	3.748	0.005
	Within Groups	657.343	419	1.569		
	Total	680.866	423			
2015, 1 year after purchase	Between Groups	36.840	4	9.210	5.400	< 0.001
	Within Groups	714.669	419	1.706		
	Total	751.509	423			
2016, 2 years after purchase	Between Groups	23.749	4	5.937	3.554	0.007
	Within Groups	659.929	395	1.671		
	Total	683.678	399			
2018, 4 years after purchase	Between Groups	22.474	4	5.618	2.925	0.021
	Within Groups	800.875	417	1.921		
	Total	823.349	421			

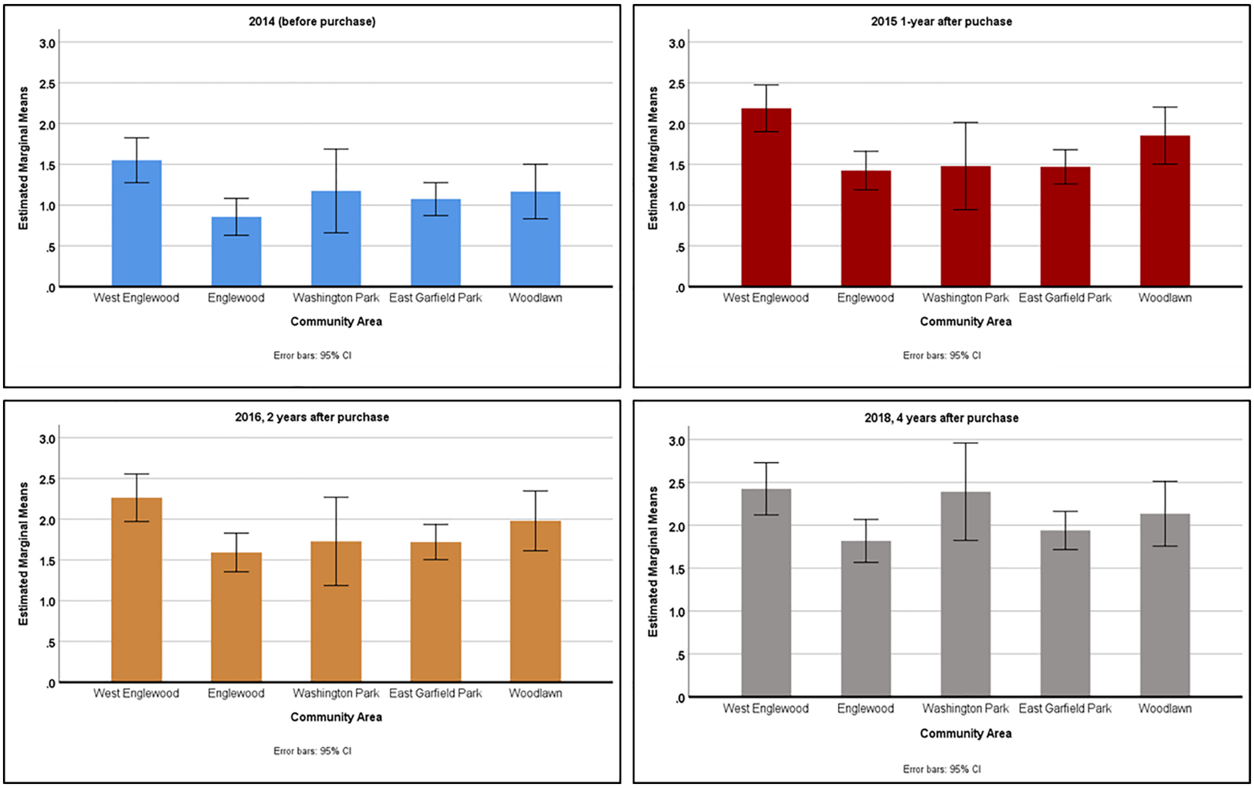

Fig. 4. Estimated marginal means for one-way ANOVAs of condition-care for the five community areas (cross-sectional application). The plots for each sampling period show the relative magnitude of improvement over time and the 95% confidence interval error bars show the significance of differences between community areas. West Englewood consistently received significantly higher scores on the condition-care scale than Englewood or East Garfield Park, $p < .05$ except for 2018, where the difference between West Englewood and East Garfield Park was $p = .087$. None of the other pairwise comparisons were statistically different.

important performance criteria including sensitivity and utility of application. Single-rater test-retest reliability and two- and three-rater inter-rater reliability values were all very high, providing evidence that the scale can be reliably applied in both lab and field situations by an individual or a small team of raters (Palmer & Hoffman, 2001). While the condition-care scale addresses a complex, subjective concept, the training document (Appendix S1) and descriptive criteria (Table 1) provide concrete examples for each level of the scale, helping raters to distinguish differences between each level and assign values for lots to the proper category, strengthening the scale's sensitivity to account for visible changes in condition and care (Daniel & Vining, 1983). In our laboratory application, it may have also aided reliability to apply the scale in a pairwise setup with images from the previous year to visually inspect if changes were present before making the rating. We did not have the advantage of a comparison photo in the field, but we did have the ability to discuss our independent ratings with other team members to arrive at a consensus score, which also strengthened the reliability of the scale's field application.

Our primary effort to measure validity was directed at the scale itself, and our previous work (Gobster et al., 2020a; Stewart et al., 2019) provided similar, independently-derived measures of condition and care. That the scale did not correlate more highly with our 7-item audit index may indicate that the two measures (as well as the 9-item mail survey index) capture somewhat different elements of condition and care, consistent with ideas of convergent validity (Chin & Yao, 2014; Gregory, 2007). Notably, the scale emphasizes turf condition in the lower, "condition" levels of the scale, while for performance reasons we had to remove our turf condition measure from the audit index (Gobster

Table 6
HLM model of large lot condition-care with joint longitudinal and cross-sectional effects.

	Estimate	Conf. Int.	df	*t*	*p*
Fixed effects					
(Intercept)	1.570	1.321–1.819	493.869	12.381	< 0.001
Time2 (2015)	0.521	0.410–0.632	1246.506	9.205	< 0.001
Time3 (2016)	0.705	0.592–0.818	1250.636	12.216	< 0.001
Time4 (2018)	0.925	0.813–1.036	1246.837	16.304	< 0.001
Nhood2 (Englewood)	−0.684	−0.994 — −0.374	423.619	−4.329	< 0.001
Nhood3 (Washington Park)	−0.413	−0.920–0.094	423.796	−1.601	0.110
Nhood4 (East Garfield Park)	−0.554	−0.851 — −0.257	424.207	−3.664	< 0.001
Nhood5 (Woodlawn)	−0.332	−0.710–0.046	426.188	−1.723	0.087
Random effects					
Intercept variance (lot ID)	1.081				

$N = 424$, AIC = 4932.9, marginal $R^2 = 094$. Tukey adjusted post-hoc tests were significant for all pairwise comparisons between time variables (Time1 (2014) reference condition), $p < 0.001$ except for Time2 - Time3, $p = .008$; significant pairwise comparisons between community areas were Nhood1 (West Englewood (reference condition) – Nhood2 (Englewood) $p < .001$ and Nhood1 – Nhood4 (East Garfield Park) $p = .002$.

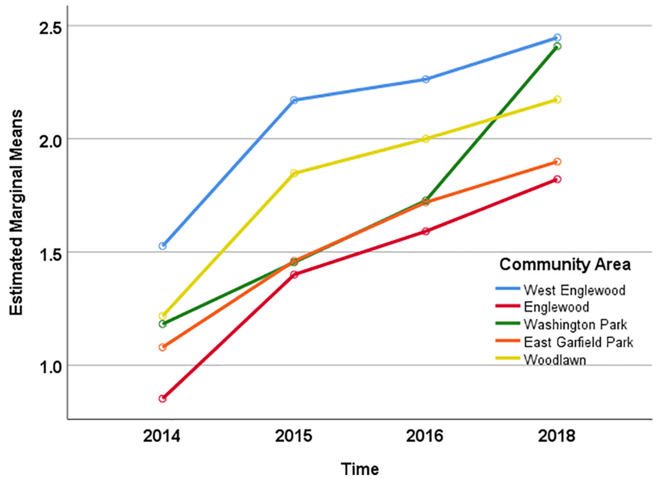

Fig. 5. Estimated marginal means for condition-care scale ratings as a function of joint effects of time and community area (integrated longitudinal and cross-sectional application).

et al., 2020a). It may also be that as a gestalt measure, when applying the condition-care scale a rater might (consciously or unconsciously) emphasize particular aspects of a given lot's condition and care, whereas the audit index forces the assessor to systematically evaluate each element that composes the index. This more qualitative type of judgment is a potential limitation of the scale, but it might also provide the rater greater flexibility in resolving inconsistencies in the elements of condition and care present and incorporating important ad hoc elements not included in the index, such as symbols (e.g., a Black Lives Matter sign or Pan-African flag) that hold special meaning to a community. Thus, what would be considered a limitation from one perspective could be viewed as a benefit from another.

As for method validity, a limitation in our field sampling was the seasonal inconsistency when we entered the field. Field imagery was also collected in 2017 but many photos were taken too late in the season and some aspects of the scale that dealt with vegetation could not be reliably assessed, so we did not complete ratings for that year. In all other years, vegetation was still largely green and growing but sampling in 2014 (pre-purchase) was summer-fall, 2015 was early fall, 2016 was late summer, and 2018 was mid-summer. Each of these seasons in the Chicago region can be somewhat different, and while we did not feel the variation significantly influenced ratings, future applications of this approach for monitoring over time should standardize the sampling to a consistent period. Given our experience, we feel mid-summer for the Chicago region is ideal, as seasonal yard and garden vegetation is fully established yet not overgrown or gone by, giving the opportunity for the best potential representation of a landowner's efforts.

5.2. Utility

Another key criterion in the evaluation of assessment methods is utility, which is usually discussed in terms of efficiency, usefulness, and generalizability of application (Daniel & Vining, 1983). Regarding efficiency, while a motivation for developing the condition-care scale was the lack of available aerial imagery to complete our post-intervention audit, the scale was also very much designed to provide a rapid assessment for parcels. In this case, "rapid" is a relative term; the actual rating is fairly quick (usually less than one minute), but pinpointing parcel locations for field or Street View sampling is quite involved the first time around. Still, this pre-work is necessary no matter whether a single scale or full audit is used, and even in our study sites where large lots were dispersed, we completed photographing and rating tasks at an average of four to five minutes per lot, including the time spent driving between lots. We did not time the application of the audit protocol, but it was substantially longer. And while the value of that information should not be diminished, if the objective is to provide a relatively easy-to-use approach to monitoring changes in vacant lot stewardship, the condition-care scale is an efficient way to do so.

Another aspect of efficiency is a method's ease to be implemented and understood. In employing the scale in both field and lab situations, we gained an appreciation for the different benefits and costs of the scale's application. The time and labor savings of lab-based virtual audits using Street View imagery and its equivalents are a major benefit that has been previously documented (e.g., Clarke et al, 2010). We also found that Street View's slightly higher (~3m) elevation of view had advantages over our field photography in seeing changes to lots that were fronted by wooden privacy fences. In field situations, there were a few lots for which we had to raise our camera over the fence to capture an image reflecting the lot's condition. Capturing images via a Google survey camera has the added advantages of anonymity and perhaps safety, though we never felt uncomfortable in the neighborhoods and were frequently welcomed by owners and neighbors. In fact, interacting with residents was a major benefit from our field-based assessments, and provided an important opportunity for gaining information from residents' experiences and for sharing information about the program (Kaplan & Basu, 2015). Involving the city's planners and interns was another benefit of the field assessment, and they felt the experience brought them new insights into the program and management issues faced by lot owners. City staff also felt that their involvement in the assessment increased their understanding of the value of the research evaluation, and in summer 2019 Large Lot Program staff and interns used the scale in a virtual evaluation of all 1,248 lots currently in the program. The combination of field assessment and virtual audits complement one another and build synergies that become greater than their sum.

Our applications of the scale both longitudinally and cross-sectionally also demonstrate its usefulness to addressing planning issues related to vacant land greening. As a tool for monitoring change over time, the scale is sensitive to changes made to lots between annual sampling periods and provides an ordinal-level metric of the magnitude of change. In the case of our application, we found that the biggest improvements to lots occurred in the first year after purchase, but that continued progress was made in the second and fourth years. As shown in the qualitative presentation of individual lots in the Supplementary Appendix S2, there may also be important sub-patterns of change including major changes such as housing development that occurred in the fourth year after purchase, which further demonstrates the utility of periodic monitoring.

The cross-sectional application also demonstrates the potential usefulness of the scale to planners wishing to understand variability in vacant lot stewardship as a function of location within a city or between cities. In our cross-sectional application, one thing that distinguishes the community area with the highest large lot condition and care, West Englewood, from the others is the much higher rate of owner-occupancy and single-family homes than the other community areas (see Table 1). While further research is needed to confirm this association, it generally agrees with research that owner-occupants are more likely to maintain their residence than absentee owners or renters (Garvin et al., 2013; Goldstein, Jensen, & Reiskin, 2001). In related work using this same data, we also found that owner-occupants who purchased large lots directly adjacent to their residence extended higher levels of care to them than lots purchased further away (Gobster, Rigolon, Hadavi, & Stewart, 2020b). Because West Englewood has higher shares of owner-occupants and single-family homes than other community areas, there might have been more cases in which property owners were able to purchase a large lot directly adjacent to their home.

As a neighborhood planning tool, the scale's cross-sectional application also shows that longitudinal data is not necessary to understand important patterns of lot care. By using simple *t*-tests or one-way

ANOVAs, planners can compare the mean scores from the scale across a sample of lots from different neighborhoods to get a snapshot of their status at any given point in time. These cross-sectional comparisons can also help planners identify differences in condition and care between neighborhoods with different demographic and built environment characteristics. Moreover, the joint HLM model (Table 5) shows that the pattern of differences between community areas remains stable across time, giving further confidence in its usefulness as a comparative tool. The only community area that showed some deviation from the overall pattern (compare Fig. 2 with Fig. 4) was Washington Park, which had the smallest sample size ($n = 23$). While we feel confident that the scale provided a reliable comparison across our community areas, we suggest that, optimally, planners should aim for samples of 50 lots or more per neighborhood unit.

As for generalizability, our application was limited to the five community areas of residential parcels purchased under the Chicago Large Lot Program, and further work is needed to test how the scale might serve other locations and issues facing urban greening program managers. Among the many vacant lot leasing and resale programs in the United States and globally, most aim at improving the condition and care of lots (Ganning & Tighe, 2015; Nassauer & Raskin, 2014). If applied to other programs, some revision of criteria for different levels of the scale may be needed, as well as adjustment of the deployment approach in order to measure changes. For the Large Lot Program, we were fortunate that all the lots in their pilot areas were transferred around the same time, but for other programs transfers are usually done on a lot-by-lot basis. This increases the challenge of assessing change, though program managers could ask residents or the granting agency to include a photo of the lot as part of the application process, and then managers could use field or online imagery when it is available to assess change.

Similar adaptations of the scale and its application could make it useful for evaluation for community gardening, easement greening programs, and related efforts. Although similar efforts have been documented (Rupprecht & Byrne, 2014), ours is the first that monitors change before and after a policy intervention in which a commitment to stewardship is made. In addition, our condition-care scale can be used cross-sectionally to identify neighborhoods where lots have not substantially improved across time in which public agencies and nongovernmental organizations could develop support programs to help owners improve the condition of their lots. Finally, while cities can use the scale by itself in longitudinal and cross-sectional applications to assess condition and care as an outcome (dependent variable), the scale could also be applied the way neighborhood audits of disorder are often used, that is as a predictor (independent variable) to assess broader social outcomes such as reductions in crime and improvements in community health (Sampson & Raudenbush, 2004).

6. Conclusions

The maintenance and care of places are central ways in which people perceive and carry out landscape stewardship (Nassauer, 1997). Cultural norms help define both minimally acceptable and desired conditions, and maintenance at a basic level can often happen without a deep dedication to personal or community values. In Chicago and other cities experiencing vacancy, many vacant lots are maintained at such a level when residents call the city to request that they mow an overgrown vacant lot. Getting beyond this basic level of maintenance requires an investment in care, which is expressed in the normative aesthetic and functional conventions people place on the land such as planting colorful flowers and edible plants or constructing a patio or play area for family and friends. Building on the parallel research traditions of neighborhood audits and visual assessments, we developed the condition-care scale to measure this continuum of vacant lot stewardship. While our new scale does not substitute for understanding the full dimensionality of condition and care that a detailed audit can provide, our development and application of it to the Chicago Large Lot Program supports its value and potential. We are enthusiastic about the potential for further application and improvement upon the approach described here and encourage planners and other researchers to adapt the condition-care scale to their particular needs.

CRediT authorship contribution statement

Paul H. Gobster: Conceptualization, Project administration, Writing - original draft. **Alessandro Rigolon:** Formal analysis, Writing - original draft. **Sara Hadavi:** Visualization, Writing - original draft. **William P. Stewart:** Conceptualization, Project administration, Validation, Writing - original draft.

Acknowledgements

This work was supported in part through USDA Forest Service Northern Research Station Cooperative Research Agreement 15-JV-11242309-075 with the University of Illinois-Urbana-Champaign. We thank Rose Grenen, Northwestern University, for research assistance, and field team members Cherie Fisher, Drew Hart, Vassia Heneghan, Sarah McCaffrey, and Sonya Sachdeva, USDA Forest Service Northern Research Station; Joyce Beal, Kim Harrison, and Danny Wong, City of Chicago Department of Planning and Development; Katya Obrez, University of Michigan; and Jonathan Weissman, University of Illinois at Urbana-Champaign. We also thank Jim Palmer and three anonymous reviewers for their helpful comments on earlier versions of this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.landurbplan.2020.103885.

References

Ambrey, C. L., Fleming, C. M., & Manning, M. (2014). Perception or reality, what matters most when it comes to crime in your neighbourhood? *Social Indicators Research, 119*, 877–896. https://doi.org/10.1007/s11205-013-0521-6.

American Psychological Association (n.d.). External reliability. Dictionary of Psychology. https://dictionary.apa.org/external-reliability.

Bader, M. D. M., & Ailshire, J. A. (2014). Creating measures of theoretically relevant neighborhood attributes at multiple spatial scales. *Sociological Methodology, 44*, 322–368. https://doi.org/10.1177/0081175013516749.

Bader, M. D. M., Mooney, S. J., Bennett, B., & Rundle, A. G. (2017). The promise, practicalities, and perils of virtually auditing neighborhoods using Google Street View. *Annals of the American Academy of Political and Social Science, 669*, 18–40. https://doi.org/10.1177/0002716216681488.

Bader, M. D. M., Mooney, S. J., Lee, Y. J., Sheehan, D., Neckerman, K. M., Rundle, A. G., & Teitler, J. O. (2015). Development and deployment of the Computer Assisted Neighborhood Visual Assessment System (CANVAS) to measure health-related neighborhood conditions. *Health and Place, 31*, 163–172. https://doi.org/10.1016/j.healthplace.2014.10.012.

Baldock, K. L., Paquet, C., Howard, N. J., Coffee, N. T., Taylor, A. W., & Daniel, M. (2018). Gender-specific associations between perceived and objective neighbourhood crime and metabolic syndrome. *PLoS ONE, 13*(7), Article e0201336. https://doi.org/10.1371/journal.pone.0201336.

Cândido, R. L., Steinmetz-Wood, M., Morency, P., & Kestens, Y. (2018). Reassessing urban health interventions: Back to the future with Google Street View Time Machine. *American Journal of Preventive Medicine, 55*, 662–669. https://doi.org/10.1016/j.amepre.2018.04.047.

Chin, C.-L., & Yao, G. (2014). Convergent validity. In A. C. Michalos (Ed.). *Encyclopedia of quality of life and well-being research* (pp. 1275–1276). Dordrecht: Springer.

City of Chicago (2014a). *Large lot program.* Chicago: City of Chicago Department of Planning and Development https://www.cityofchicago.org/content/dam/city/depts/dol/rulesandregs/DPDLARGELOTSPROGRAM.pdf.

City of Chicago (2014b). *Green healthy neighborhoods.* Chicago: City of Chicago Department of Planning and Development https://www.cityofchicago.org/city/en/depts/dcd/supp_info/green-healthyneighborhoods.Html.

Clarke, P., Ailshire, J., Melendez, R., Bader, M., & Moore, J. (2010). Using Google Earth to conduct a neighborhood audit: Reliability of a virtual audit instrument. *Health & Place, 16*, 1224–1229. https://doi.org/10.1016/j.healthplace.2010.08.007.

Cook County, Illinois. (n.d.). Cook County property tax portal. http://www.cookcountypropertyinfo.com/.

Craik, K. H., & Feimer, N. R. (1979). Setting technical standards for visual assessment procedures. In Elsner, G. H., and Smardon, R. C. (Eds.) Proceedings of our national

landscape: A conference on applied techniques for analysis and management of the visual resource. General Technical Report PSW-35. Berkeley, CA: USDA Forest Service, Pacific Southwest Forest and Range Experiment Station, 93-100.
Daniel, T. C. (2001). Whither scenic beauty? Visual landscape quality assessment in the 21st century. *Landscape and Urban Planning, 54*, 267–281. https://doi.org/10.1016/S0169-2046(01)00141-4.
Daniel, T. C., & Boster, R. S. (1976). Measuring landscape esthetics: The scenic beauty estimation method. Research Paper RM-167. Fort Collins, CO: USDA Forest Service, Rocky Mountain Forest and Range Experiment Station.
Daniel, T. C., & Vining, J. I. (1983). Methodological issues in the assessment of landscape quality. In I. Altman, & J. F. Wohlwill (Vol. Eds.), *Human behavior and environment: vol. 6*, (pp. 39–84). New York: Plenum.
Dewar, M., Nassauer, J., & Dueweke, E. (2013). Cues to care database-codebook. Unpublished codebook for Detroit Sustainability Indicators Integrated Assessment. Ann Arbor: University of Michigan Graham Environmental Sustainability Institute.
Dewar, M., Nassauer, J., Dueweke, E., & Boswell, R. (2015). Cues to care database-codebook: Essential cues to care. Unpublished codebook for Detroit Sustainability Indicators Integrated Assessment. Ann Arbor: University of Michigan Graham Environmental Sustainability Institute.
Evans-Cowley, J. S., & Akar, G. (2014). StreetSeen visual survey tool for determining factors that make a street attractive for bicycling. *Transportation Research Record, 2468*, 19–27. https://doi.org/10.3141/2468-03.
Ewing, R., Handy, S., Brownson, R. C., Clemente, O., & Winston, E. (2006). Identifying and measuring urban design qualities related to walkability. *Journal of Physical Activity and Health, 3*(Suppl 1), S223–S240.
Franzini, L., Taylor, W., Elliott, M. N., Cuccaro, P., Tortolero, S. R., Janice Gilliland, M., Grunbaum, J., & Schuster, M. A. (2010). Neighborhood characteristics favorable to outdoor physical activity: Disparities by socioeconomic and racial/ethnic composition. *Health and Place, 16*, 267–274. https://doi.org/10.1016/j.healthplace.2009.10.009.
Ganning, J. P., & Tighe, J. R. (2015). Assessing the feasibility of side yard programs as a solution to land vacancy in U.S. shrinking cities. *Urban Affairs Review, 51*(5), 708–725. https://doi.org/10.1177/1078087414555435.
Garvin, E., Branas, C., Keddem, S., Sellman, J., & Cannuscio, C. (2013). More than just an eyesore: Local insights and solutions on vacant land and urban health. *Journal of Urban Health, 90*, 412–426. https://doi.org/10.1007/s11524-012-9782-7.
Gidlow, C., van Kempen, E., Smith, G., Triguero-Mas, M., Kruize, H., Gražulevičienė, R., ... Nieuwenhuijsen, M. J. (2018). Development of the natural environment scoring tool (NEST). *Urban Forestry and Urban Greening, 29*, 322–333. https://doi.org/10.1016/j.ufug.2017.12.007.
Gobster, P. H., Hadavi, S., Rigolon, A., & Stewart, W. P. (2020a). Measuring landscape change, lot by lot: Greening activity in response to a vacant land reuse program. *Landscape and Urban Planning, 196*, 103729. https://doi.org/10.1016/j.landurbplan.2019.103729.
Gobster, P. H., Ribe, R. G., & Palmer, J. F. (2019). Themes and trends in visual assessment research: Introduction to the *Landscape and Urban Planning* special collection on the visual assessment of landscapes. *Landscape and Urban Planning, 191*, Article 103635. https://doi.org/10.1016/j.landurbplan.2019.103635.
Gobster, P. H., Rigolon, A., Hadavi, S., & Stewart, W. P. (2020b). Beyond proximity: Extending the "greening hypothesis" in the context of vacant lot stewardship. *Landscape and Urban Planning, 197*, 103773. https://doi.org/10.1016/10.1016/j.landurbplan.2020.103773.
Goldstein, J., Jensen, M., & Reiskin, E. (2001). *Urban vacant land redevelopment: Challenges and progress (Working paper WP01JG1)*. Cambridge, MA: Lincoln: Institute of Land Policy.
Green, M. A., Widener, M., Pollock, F. D., & Pearce, J. (2019). The evolution of *Health & Place*: Text mining papers published between 1995 and 2018. *Health & Place*. https://doi.org/10.1016/j.healthplace.2019.102207.
Gregory, R. J. (2007). *Psychology testing: History, principles, and applications* (5th ed.). Boston: Ally & Bacon.
Gullón, P., Badland, H. M., Alfayate, S., Bilal, U., Escobar, F., ... Franco, M. (2015). Assessing walking and cycling environments in the streets of Madrid: Comparing on-field and virtual audits. *Journal of Urban Health, 92*, 923–939. https://doi.org/10.1007/s11524-015-9982-z.
Hanibuchi, T., Nakaya, T., & Inoue, S. (2019). Virtual audits of streetscapes by crowdworkers. *Health and Place, 59*, Article 102203. https://doi.org/10.1016/j.healthplace.2019.102203.
Howell, D. C. (2011). *Statistical methods for psychology* (8th ed.). Belmont, CA: Wadsworth Publishing.
Jorgensen, A., Hitchmough, J., & Dunnett, N. (2007). Woodland as a setting for housing-appreciation and fear and the contribution to residential satisfaction and place identity in Warrington New Town, UK. *Landscape and Urban Planning, 79*, 273–287. https://doi.org/10.1016/j.landurbplan.2006.02.015.
Kamphuis, C. B. M., Mackenbach, J. P., Giskes, K., Huisman, M., Brug, J., & van Lenthe, F. J. (2010). Why do poor people perceive poor neighbourhoods? The role of objective neighbourhood features and psychosocial factors. *Health and Place, 16*, 744–754. https://doi.org/10.1016/j.healthplace.2010.03.006.
Kaplan, R., & Basu, A. (Eds.). (2015). *Fostering reasonableness: Supportive environments for bringing out our best*. Ann Arbor, MI: Maize Books.
Kaplan, R., & Kaplan, S. (1989). *The experience of nature: A psychological perspective*. New York: Cambridge University Press.
Kim, J. W., Schweitzer, J. H., & Kim, M.-C. (2002). The relationship of community characteristics to actual and perceived crime in urban neighborhoods. *International Journal of Urban Sciences, 6*, 119–131. https://doi.org/10.1080/12265934.2002.9693512.
Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*, 155–163. https://doi.org/10.1016/j.jcm.2016.02.012.
Krusky, A. M., Heinze, J. E., Reischl, T. M., Aiyer, S. M., Franzen, S. P., & Zimmerman, M. A. (2015). The effects of produce gardens on neighborhoods: A test of the greening hypothesis in a post-industrial city. *Landscape and Urban Planning, 136*, 68–75. https://doi.org/10.1016/j.landurbplan.2014.12.003.
Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159–174. https://doi.org/10.2307/2529310.
LargeLot.org. (n.d.). Large lots: The Chicago large lot program web portal. https://largelots.org.
Larson, K. L., Nelson, K. C., Samples, S. R., ... Trammell, T. L. E. (2016). Ecosystem services in managing residential landscapes: Priorities, value dimensions, and cross-regional patterns. *Urban Ecosystems, 19*, 95–113. https://doi.org/10.1007/s11252-015-0477-1.
Li, X., Zhang, C., Li, W., Ricard, R., Meng, Q., & Zhang, W. (2015). Assessing street-level urban greenery using Google Street View and a modified green view index. *Urban Forestry and Urban Greening, 14*(3), 675–685. https://doi.org/10.1016/j.ufug.2015.06.006.
Lindemann-Matthies, P., & Marty, T. (2013). Does ecological gardening increase species richness and aesthetic quality of a garden? *Biological Conservation, 159*, 37–44. https://doi.org/10.1016/j.biocon.2012.12.011.
Lis, A., Pardela, Ł., Can, W., Katlapa, A., & Rabalski, Ł. (2019). Perceived danger and landscape preferences of walking paths with trees and shrubs by women. *Sustainability, 11*(17), 4565. https://doi.org/10.3390/su11174565.
Lynch, K., & Rivkin, M. (1959). A walk around the block. *Landscape, 8*, 24–34.
Mooney, S. J., Bader, M. D. M., Lovasi, G. S., Neckerman, K. M., Teitler, J. O., & Rundle, A. G. (2014). Validity of an ecometric neighborhood physical disorder measure constructed by virtual street audit. *American Journal of Epidemiology, 180*, 626–635. https://doi.org/10.1093/aje/kwu180.
Morckel, V. (2015). Community gardens or vacant lots? Rethinking the attractiveness and seasonality of green land uses in distressed neighborhoods. *Urban Forestry & Urban Greening, 14*, 714–721. https://doi.org/10.1016/j.ufug.2015.07.001.
Nassauer, J. I. (1995). Messy ecosystems, orderly frames. *Landscape Journal, 14*(2), 161–170. https://doi.org/10.3368/lj.14.2.161.
Nassauer, J. I. (1997). Cultural sustainability: Aligning aesthetics and ecology. In J. I. Nassauer (Ed.). *Placing nature: Culture and landscape ecology* (pp. 65–83). Washington, DC: Island Press.
Nassauer, J. I. (2011). Care and stewardship: From home to planet. *Landscape and Urban Planning, 100*, 321–323. https://doi.org/10.1016/j.landurbplan.2011.02.022.
Nassauer, J. I., & Raskin, J. (2014). Urban vacancy and land use legacies: A frontier for urban ecological research, design, and planning. *Landscape and Urban Planning, 125*, 245–253. https://doi.org/10.1016/j.landurbplan.2013.10.008.
Nassauer, J. I., Wang, Z., & Dayrell, E. (2009). What will the neighbors think? Cultural norms and ecological design. *Landscape and Urban Planning, 92*(3–4), 282–292. https://doi.org/10.1016/j.landurbplan.2009.05.010.
Palmer, J. F., & Hoffman, R. E. (2001). Rating reliability and representation validity in scenic landscape assessments. *Landscape and Urban Planning, 54*, 149–161. https://doi.org/10.1016/S0169-2046(01)00133-5.
Peterson, G. L. (1967). A model of preference: Quantitative analysis of the perception of the visual appearance of residential neighborhoods. *Journal of Regional Science, 7*, 19–31. https://doi.org/10.1111/j.1467-9787.1967.tb01420.x.
R Core Team (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: Sage.
Rega-Brodsky, C. C., Nilon, C. H., & Warren, P. S. (2018). Balancing urban biodiversity needs and resident preferences for vacant lot management. *Sustainability, 10*, 1679. https://doi.org/10.3390/su10051679.
Reischl, T. M., Krusky, A. M., Zimmerman, M. A., Aiyer, S. M., Kruger, D. J., Franzen, S. P., & Lipton, R. (2016). Assessing physical disorder in urban neighborhoods: Reliability and validity of the parcel maintenance observation tool. *Journal of Community Psychology, 44*, 7–21. https://doi.org/10.1002/jcop.21730.
Rundle, A. G., Bader, M. D. M., Richards, C. A., Neckerman, K. M., & Teitler, J. O. (2011). Using Google Street View to audit neighborhood environments. *American Journal of Preventive Medicine, 40*, 94–100. https://doi.org/10.1016/j.amepre.2010.09.034.
Rupprecht, C., & Byrne, J. (2014). Informal urban green space: Comparison of quantity and characteristics in Brisbane, Australia and Sapporo, Japan. *PloS ONE, 9*(6), Article e99784. https://doi.org/10.1371/journal.pone.0099784.
Rzotkiewicz, A., Pearson, A. L., Dougherty, B. V., Shortridge, A., & Wilson, N. (2018). Systematic review of the use of Google Street View in health research: Major themes, strengths, weaknesses and possibilities for future research. *Health and Place, 52*, 240–246. https://doi.org/10.1016/j.healthplace.2018.07.001.
Sampson, R. J. (2012). *Great American City: Chicago and the enduring neighborhood effect*. Chicago: University of Chicago Press.
Sampson, R. J., & Raudenbush, S. W. (1999). Systematic social observation of public spaces: A new look at disorder in urban neighborhoods. *American Journal of Sociology, 105*, 603–651. https://doi.org/10.1086/210356.
Sampson, R. J., & Raudenbush, S. W. (2004). Seeing disorder: Neighborhood stigma and the social construction of "broken windows". *Social Psychology Quarterly, 67*(4), 319–342. https://doi.org/10.1177/019027250406700401.
Schroeder, H. W. (1984). Environmental perception rating scales: A case for simple methods of analysis. *Environment and Behavior, 16*, 573–598. https://doi.org/10.1177/0013916584165003.
Seligman, A. (2005). Community areas. Encyclopedia of Chicago. Chicago: Chicago Historical Society. http://www.encyclopedia.chicagohistory.org/pages/319.html.
Sevenant, M., & Antrop, M. (2011). Landscape representation validity: A comparison

between on-site observations and photographs with different angles of view. *Landscape Research, 36*, 363–385. https://doi.org/10.1080/01426397.2011.564858.
Shuttleworth, S. (1980). The use of photographs as an environmental presentation medium in landscape studies. *Journal of Environmental Management, 11*, 61–76.
Sisser, J. M., Nelson, K. C., Larson, K. L., Ogden, L. A., Polsky, C., & Chowdhury, R. R. (2016). Lawn enforcement: How municipal policies and neighborhood norms influence homeowner residential landscape management. *Landscape and Urban Planning, 150*, 16–25. https://doi.org/10.1016/j.landurbplan.2016.02.011.
Stewart, W. P., Gobster, P. H., Rigolon, A., Strauser, J., Williams, D. A., & van Riper, C. J. (2019). Resident-led beautification of vacant lots that connects place to community. *Landscape and Urban Planning, 185*, 200–209. https://doi.org/10.1016/j.landurbplan.2019.02.011.
Tang, J., & Long, Y. (2019). Measuring visual quality of street space and its temporal variation: Methodology and its application in the Hutong area in Beijing. *Landscape and Urban Planning, 191*, Article 103436. https://doi.org/10.1016/j.landurbplan.2018.09.015.
Troy, A., Nunery, A., & Grove, J. M. (2016). The relationship between residential yard management and neighborhood crime: An analysis from Baltimore City and County. *Landscape and Urban Planning, 147*, 78–87. https://doi.org/10.1016/j.landurbplan.2015.11.004.
Visscher, R. S., Nassauer, J. I., & Marshall, L. L. (2016). Homeowner preferences for wooded front yards and backyards: Implications for carbon storage. *Landscape and Urban Planning, 146*. https://doi.org/10.1016/j.landurbplan.2015.09.001.
Wilson, J. Q., & Kelling, G. L. (1982). The police and neighborhood safety: Broken windows. *Atlantic Monthly, 249*(March), 29–38.