URBAN DESIGN International
https://doi.org/10.1057/s41289-024-00264-1

ORIGINAL ARTICLE

Deciphering physical disorder of urban street space in China's rust belt: Identification, perception, and interpretation through street-view images

Shuqi Gao[1] · Huimin Zhao[2] · Jingjia Chen[2] · Cao Gan[2] · Guikai Huang[2] · Ying Long[3]

Accepted: 9 November 2024

Abstract

Physical disorder can cause social disorder and urban decline. The emergence of virtual auditing via online street-view images provides a brand-new way to study physical disorder in the real world. However, most precedent studies focused on cities in developed nations, leaving cities in the developing world largely unstudied. This study aims at addressing this lacuna by taking the City of Qiqihar (Heilongjiang, China) as an example, to decipher its physical disorder status and related factors. By geo-sampling, we extracted 4852 street-view images from Tencent Maps. After training, auditors developed a checklist with 13 disorder items and independently audited the 4852 images to identify disorder items and perceive disorder of each image. We found that the number of disorder items in a street-view image is strongly correlated with the chance of its being perceived as disorder by auditors. In addition, seven out of the 13 disorder items are more strongly associated with the perceived disorder when compared with other disorder items. This study demonstrates that virtual auditing via street-view images is feasible for researching physical disorder and related issues in Chinese cities. It also contributes to the refinement of the traditional "broken windows theory".

Keywords Physical disorder · Virtual auditing · Street view · Shrinking cities · Urban decay · China

Introduction

Physical disorder is detrimental to neighborhood development, social stability, and personal health. In order to alleviate physical disorder, one top priority is auditing the physical environment and measuring the level of disorder in every place. However, it is not an easy task due to the following two issues. First, it is costly. Auditing the physical environment involves a great amount of time and financial resources (Hwang 2017; Mooney et al. 2017). Precedent studies applied community attitude surveys (Li and Wu 2013) and in-person observations (Jones et al. 2011; Marco et al. 2015) that engaged a significant number of workforce. Therefore, only a few well-funded research groups have the capability to do such studies. Second, the urban environment keeps changing. Affected by everyday human activities, the physical environment changes subtly every day, and brings uncertainty for public policy and urban planning (Abbott 2005; Krieger 2022). This is especially so in shrinking cities and declining neighborhoods that are afflicted with sabotage (Beauregard 2021; Gao and Ryan 2024; Gao et al. 2023; Rigolon et al. 2021; Stern and Lester 2021; Zhou et al.

Shuqi Gao and Huimin Zhao contributed equally to this paper.

✉ Ying Long
ylong@tsinghua.edu.cn

Shuqi Gao
shuqigao@seu.edu.cn

Huimin Zhao
zhaohm21@mails.tsinghua.edu.cn

Jingjia Chen
chen-jj18@tsinghua.org.cn

Cao Gan
ganc18@mails.tsinghua.edu.cn

Guikai Huang
guikai_huang@outlook.com

[1] School of Architecture, Southeast University, Xuanwu District, Nanjing, Jiangsu, China

[2] School of Architecture, Tsinghua University, Haidian District, Beijing, China

[3] Key Laboratory of Ecological Planning & Green Building, Ministry of Education, School of Architecture and Hang Lung Center for Real Estate, Tsinghua University, Haidian District, Beijing, China

Published online: 10 January 2025

2022). Thus, auditing the physical environment should be a routine task, which further decreases the possibility of its application.

Recently, in order to save time and financial cost in auditing physical environment, some scholars have applied new technology, ranging from remote sensing (Deng and Ma 2015), unmanned aerial vehicles (Grubesic et al. 2018; Park and Ewing 2022), and online street views (Clarke et al. 2010; Rundle et al. 2011; Marco et al. 2017; Ma et al. 2023; Bader et al. 2017; Plascak et al. 2020). While remote sensing imagery, as one of the open data sources that comprehensively covers urban areas, can be employed for virtual audits of physical disorder, it only provides an aerial perspective. This limitation, coupled with low-resolution imagery in some areas, severely restricts the amount of actionable audit data. Although drone imagery improves upon these constraints, it still requires on-site data collection and is significantly affected by weather conditions, resulting in high acquisition costs. Given the limitations of the two above approaches, online street view is the most popular in current virtual auditing literature. This popularity is due, on one hand, to its provision of data comparable to that obtained from in-person audits, with the added advantage of continuous updates, enabling researchers to track changes without visiting sites. On the other hand, the availability of open street-view data, which covers extensive urban areas, has made access highly cost-effective (Zhang et al. 2023). Although people may cast doubts on whether auditing via street-view images has enough credibility, studies have demonstrated its feasibility by comparing virtual auditing with in-person auditing, revealing that the former has good consistency with the latter (e.g., Clarke et al. 2010; Rundle et al. 2011; Mooney et al. 2017).

However, most of the precedent virtual auditing studies focused on cities in developed countries by leaving cities in developing countries largely unstudied. For instance, Mooney et al. applied virtual auditing techniques to assess nine previously validated items aimed at capturing physical disorder (e.g., litter, graffiti, and abandoned buildings) on 1826 block faces using Google Street View imagery (Google, Inc., Mountain View, CA, USA) from 2007 to 2011 across four U.S. cities: San Jose, CA; Detroit, MI; New York, NY; and Philadelphia, PA (Mooney et al. 2014). Similarly, Feuillet et al. conducted a virtual audit using Google Street View data to examine disorder elements such as graffiti and litter across 4486 street segments in urban areas of Belgium, France, Hungary, the Netherlands, and the United Kingdom (Feuillet et al. 2016).

This situation can be partially explained by the fact that online map services are much better and earlier established in most developed countries than in developing countries. Meanwhile, the physical environment varies from place to place and has vernacular characteristics, requiring more studies toward urban environments in different countries. Fortunately, in recent years, some Internet technology companies in developing countries, e.g., China, have started to offer online map services with street-view images. Therefore, now it is possible to decipher characteristics of urban physical disorder in developing countries, as a supplement to the current diverse body of literature in physical disorder and virtual auditing.

In this paper, we utilized street-view images from Tencent Maps (a China's online-map service, similar to Google Maps) to conduct virtual auditing and study physical disorder in China. Our case city is Qiqihar, located in northeast China where has a moniker as China's rust belt due to its struggling with economic stagnation and population loss (Xie et al. 2016). The structure of this paper is organized as follows. We first present a brief review of studies regarding physical disorder and its auditing. We then introduce the background and context of Qiqihar, as well as our research method. Subsequently, we present our research findings in two subsections: first, perceiving and identifying features of physical disorder; second, interpreting factors of perceived disorder. The paper concludes with a discussion of contributions, policy recommendations, and conclusive remarks.

Understanding and auditing physical disorder

Understanding physical disorder

Order and disorder are a pair of related concepts. Their definitions involve arguing for the preeminence of one set of values over the other (Body-Gendrot 2014). In general, order implies a state of peace, safety, tranquility, and observance of law, while disorder implies a state of turbulent, threatened, and lack of control (Body-Gendrot 2014; Gau and Pratt 2010; Ross and Mirowsky 1999). However, there is no universally accepted distinction between the two concepts. Some phenomena, behaviors, and activities may be treated as order from one perspective, but may be regarded as disorder from another (Skogan 1990). Additionally, human perception of the disorder may be preempted by prior beliefs, making prejudice play a significant role (Sampson and Raudenbush 2004). Therefore, it is worthwhile to conduct research on social and physical disorder in different regions, with researchers from diversified backgrounds. As noted by Skogan (2015), "research concepts are not right or wrong; rather, they are either useful or not useful."

Specifically, the concept of disorder has two distinct but interrelated branches: social disorder and physical disorder. Social disorder is a matter of events that people feel threatened and can hardly perceive it until getting to the place; while physical disorder involves enduring visual signs of

deterioration that may be discerned by people who have not even been to the place, but has somehow seen it via various ways of media (Sampson and Raudenbush 1999; Sampson 2009; Skogan 2015; Wallace and Schalliol 2015). In other words, the social disorder appears as a series of more-or-less episodic events, while physical disorder refers to ongoing conditions (Sampson 2009; Skogan 1990). That also being said, discerning social disorder requires auditors to spend a much longer period of time when compared with discerning physical disorder, which can be instantly and tangibly perceived by outliers. However, the differences between the two types of disorder are not entirely clear-cut, as many aspects of physical disorder can serve as indicators of social disorder. Similarly, physical disorder can exacerbate the degree of social disorder (Murphy 2012; Ross and Mirowsky 1999; Raudenbush and Sampson 1999).

Auditing physical disorder

Although some aspects of physical disorder may be regarded as beneficial to either society or individuals, in most cases, physical disorder is obviously detrimental. The well-known "broken windows" theory reveals that, without immediate intervention, any trivial signs of sabotage will decay and cause physical disorder (Wilson and Kelling 1982). This proposition was further demonstrated by subsequent empirical studies that physical disorder is associated with crime, hooligans, physical illness, and juvenile delinquency (Chen and Rafail 2021; Gold and Nepomnyaschy 2018; Kotus and Rzeszewski 2013; Mayne et al. 2019). Therefore, for urban planners and policymakers who want to improve quality of life and the physical environment, locating places with physical disorder features and evaluating their severities are the two prerequisites, which can be achieved by auditing the physical environment. However, neither of the two tasks is easy to complete in practice.

First, physical disorder has various manifestations, each contributing differently to the overall perception of disorder depending on the perspective. To date, in order not to be entangled with too many features in the physical environment, scholars have to select a limited amount of top-priority features to help auditors mark the times of each feature's appearance. In reference studies (Allen 2013; Franzini et al. 2008; Hoeben et al. 2018; Kotus and Rzeszewski 2013; Perkins et al. 1993; Perkins and Taylor 1996; Rohe and Burby 1988; Ross and Mirowsky 1999; Sampson and Raudenbush 1999, 2004; Skogan 1990; Wallace and Schalliol 2015), some scholars have presented their prioritized features in auditing physical disorder, most of which include the following items: vandalism, vacant and abandoned properties, graffiti, litter and junk, trashed lots, unkempt/run-down buildings and streets, loitering, and boarded-up windows. However, due to significant distinctions in the development trajectory of the physical environment in different nations, it is in doubt if precedent studies, most of which were conducted in the US, can be directly applied to auditing physical disorder elsewhere.

Second, since almost everyone may have their own specific perception of the holistic view of physical disorder, it is difficult to construct a universal benchmark on evaluation, and thus it is also difficult to reach a consensus on the priorities of subsequent interventions. To measure the severity of physical disorder, precedent studies focused on inquiring either the residents who lived in and around the target area or the auditors who were trained by a specific evaluation system. The first group of studies (Rohe and Burby 1988; Skogan 1990; Ross and Mirowsky 1999) mainly applied face-to-face/phone/mail/e-mail interviews to solicit residents' viewpoints toward the severities of the physical disorder, also known as perceived physical disorder. In order to quantify the severity, these studies also applied methods to normalize residents' viewpoints, including attitude scale, semantic analysis, grounded theory data analysis, and statistical analysis. The second group of studies (Sampson and Raudenbush 1999; Rundle et al. 2011; Grubesic et al. 2018) focused on the observed physical disorder, and utilized observation as their core strategy for data collection. The research method of observation dates back to a very early time, but it did not evolve until the recent decade when advances in technologies bring new inputs to this method. Today, it involves a variety of sub-categories, including direct observation, videotaping, analyzing satellite images or images taken by unmanned aerial systems, auditing online street-view images, and so on. Actually, previous studies have found that perceived disorder plays a greater role in the negative health effects of residents (Carbone 2020; Weden et al. 2008). However, compared with observed disorder, its measurement relies more on questionnaires or interviews, which are labor-intensive. Furthermore, neither of the two groups of studies has established a universal standard for measuring the severity of the physical disorder. Therefore, it is necessary to conduct similar studies in more regions to check and verify physical disorder's consistency and diversities, thus providing possibilities for enriching theoretical thinking.

The emerging virtual auditing techniques

With the development of an online map service, scholars can apply it in virtual auditing that largely decreases the cost and threshold of conventional auditing method of in-person or videotaping. Conventional in-person auditing studies are labor-intensive, and auditors are frequently distracted from the surrounding environment and some disorder items may react to the presence of auditors, both of which will cause bias (Spano 2005; Hoeben et al. 2018). Some auditors may

face risks when entering dangerous areas, while bad weather and poorly walkable urban streets further complicate in-person auditing. It was not until Sampson and Raudenbush (1999) developed the Systematic Social Observation (SSO), which largely depended on the videotaping technique, the aforementioned difficulties of in-person auditing were partially resolved. However, the SSO is even more costly when considering that it engages in renting or purchasing video cameras and cars. A substantial shift occurred in 2010 when Google Maps began offering online street-view images. Since then, a large portion of physical disorder studies have transitioned to virtual auditing methods or, at the very least, started combining in-person auditing with virtual auditing (Bader et al. 2015, 2017; Ben-Joseph et al. 2013; Clarke et al. 2010; Griew et al. 2013; Mooney et al. 2014; Nesoff et al 2020; Rundle et al. 2011).

As a new research method, the main issue for virtual auditing lies in its reliability when compared with conventional in-person auditing method. Many studies (e.g., Badland et al. 2010; Clarke et al. 2010; Odgers et al. 2012; Rundle et al. 2011) have demonstrated that virtual auditing has fairly high reliability and is cost-effective when compared with the conventional in-person auditing. Mooney et al. (2014) reveal that the virtual auditing's results of physical disorder were also correlated with the socioeconomic data of housing vacancy and unemployment rates, further adding credit to its feasibility. Most online map services also allow auditors to track changes of a place by making longitudinal comparisons of street-view images (Hendricks and Meyer 2021).

However, some studies also discover several drawbacks of virtual auditing: first, some physical items cannot be well audited in the virtual environment, including highly detailed objects (e.g., litter, broken glass), noises and odors (Clarke et al. 2010; Rundle et al. 2011); second, the virtual auditing is limited to areas that vehicles are accessible to capture street view images (Rundle et al. 2011); third, some sidewalks and their attached features cannot be well observed due to the fact that street-view images were shot on top of a car, and some obstructions may prevent auditors from having a full picture of the physical environment (Bader et al. 2017); fourth, the virtual auditing may be restrained by some street-view images that were not up-to-date (Clarke et al. 2010). However, compared with virtual auditing's merits, its drawbacks seem to be trivial. Thus, virtual auditing has been accepted by scholars from a variety of research fields, including public health (Griew et al. 2013; Mooney et al. 2020; Schootman et al. 2016), urban studies (Li et al. 2015; Long and Liu 2017), and criminology (Khorshidi et al. 2021; Sytsma et al. 2021).

One major advantage of virtual auditing, largely overlooked by previous scholars, is that it allows researchers to infer the socioeconomic status of a study area without relying on socioeconomic statistics. Unlike developed countries like the US, which have multiple types of statistical data (e.g., Census, American Community Survey, American Housing Survey) down to fine-scale (a typical Census block equals to 0.0028 square kilometers), developing countries like China usually have poor socioeconomic statistics that their minimum statistical unit (township) is usually larger than several square kilometers, equivalent to thousands of census blocks in the US. Since social disorder and physical disorder are highly correlated, the virtual auditing provides an alternative way to study social vitality and disorder from a much finer scale in countries with limited access to fine-level socioeconomic data.

Research approach

Study area: City of Qiqihar

In this study, we chose the City of Qiqihar, a prefectural city in northeast China, to conduct virtual auditing and further analysis. Northeast China consists of three provinces, Heilongjiang, Jilin, and Liaoning, while the City of Qiqihar is located in the western part of Heilongjiang Province, sharing its western border with Inner Mongolia Province (see Fig. 1, upper left). It was once the most developed region in China around the mid-20th century with the moniker as the "Cradle of China's Industry". However, since the opening-up era (1978 to date), due to changing regional policies (Fan and Sun 2008), unstable geopolitical status (Calder and Ye 2010), family planning policy (Attane 2002) and other factors, northeast China is lagging behind many other regions in China. Today, northeast China is well known for its economic decline, deindustrialization, and depopulation, and has a new moniker as "China's rust belt" (Chung et al. 2009).

Located in northeast China, the City of Qiqihar is also affected by the regional background. Qiqihar is famous for its industrial sector, especially the heavy equipment manufacturing industry which was originally established with the assistance of the Soviet Union, during China's first 5-year plan period (1953–1957). Subsequently, the city's predominant economic sector remained heavy industry, managed by several state-owned enterprises (SOEs). However, since the 1980s, as a result of China's economic transition and opening-up policy, most Chinese SOEs are facing various challenges when competing with private and transnational corporations. Qiqihar has a similar scenario. Although some of the city's economic indicators are growing, its economic ranking in China continues to decline. In 1952, the city's GDP (473 million RMB) accounted for nearly 7‰ of China's GDP, while in 2015, this proportion dropped to 2‰. On the other hand, the city's population has also been declining in recent years. The city's seven municipal

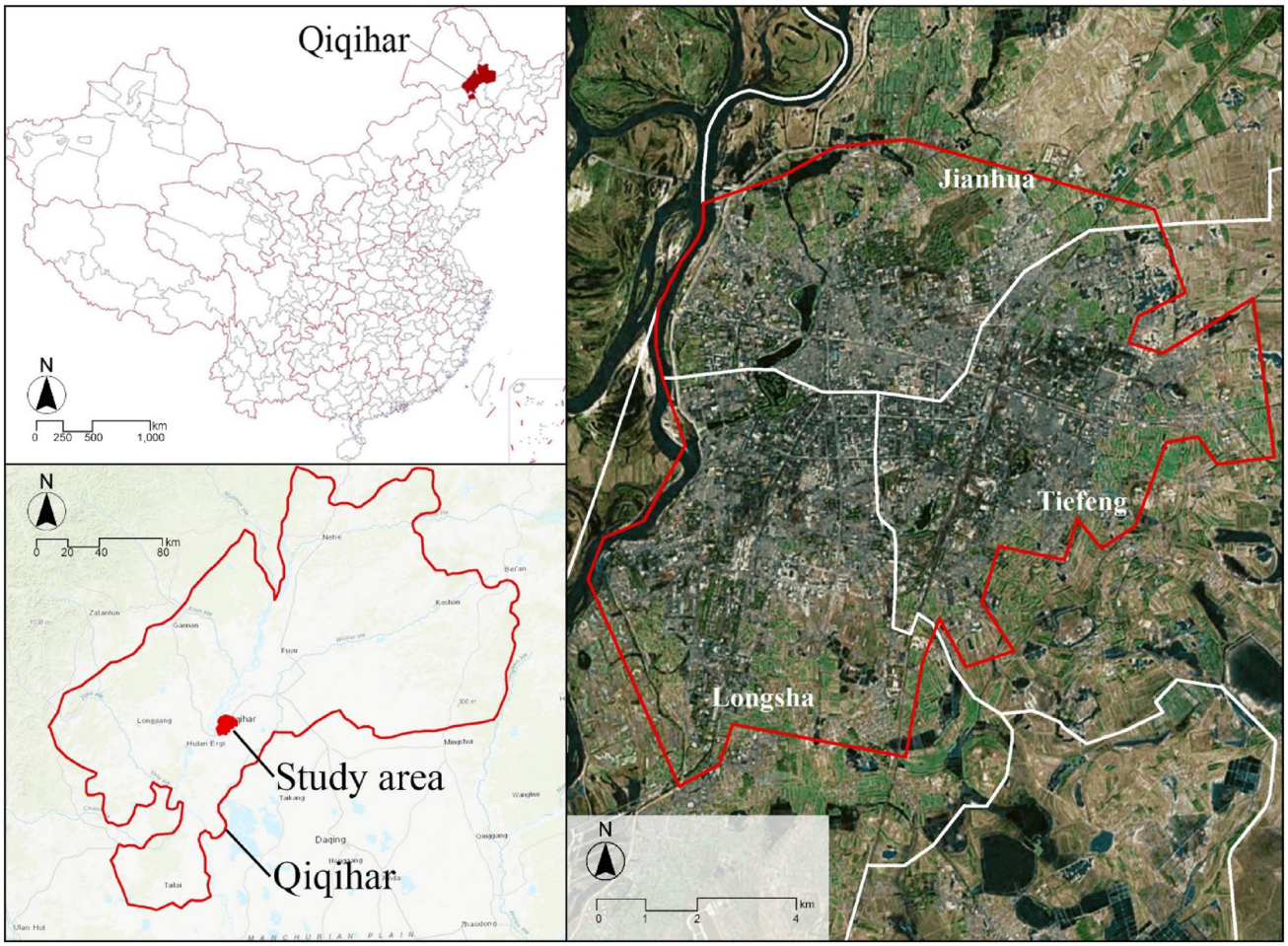


Fig. 1 Location map of Qiqihar and this study's research area. The *red polygon* within the down-left fraction is the boundary of Qiqihar's prefectural city, which contains seven municipal districts, one county-level city, and eight counties. Though the boundary of Qiqihar's prefectural city is very large, its core urban area is relatively small and encompasses only three municipal districts (Jianhua, Tiefeng, and Longsha), while other districts and counties are rural. The *red polyline* in the right fraction encloses the study area and the *white polylines* are the municipal boundary of each of the three core urban districts

districts' population declined from 1.55 million (in 2010) to 1.41 million (in 2020) (Statistical Bureau of Heilongjiang 2021). With its growing aging population and ongoing economic stagnation, the trend of depopulation is likely to worsen in the future.

This paper focuses on Qiqihar's core urban area (116 km^2, red polygon in Fig. 1 right fraction) that consists of three municipal districts (Jianhua, Tiefeng, and Longsha). According to the Statistical Bureau of Heilongjiang (2021), these three districts had 9,59,787 residents in 2020, which has decreased by 2.0% from 979,517 in 2010. The following reasons justify our selection of Qiqihar for this study. First, precedent studies indicate that shrinking and declining cities tend to have a much more severe status of physical disorder (Deng and Ma 2015; Skogan 2015); second, since physical disorder has not been well studied in China, we prefer to conduct this study in places with more severe status of physical disorder, e.g., Qiqihar; third, Qiqihar's socioeconomic pattern is similar to the regional background in northeast China, both have a similar economic structure and are afflicted with depopulation and economic stagnation. Although adding more cities for this research may help us achieve more generalized findings, physical disorder studies are rare in China, so we are unable to establish a general list of physical disorder items that can be universally applicable for auditing Chinese cities. With these considerations in mind, we decided to take Qiqihar as a pioneering case for studying physical disorder in China.

Research method and data

This paper employs a three-stage research method to address two sets of research questions. First, what are the major physical disorder features in Chinese cities, and are there

any unique features with distinct manifestations or those not reported in previous studies? Second, which physical barriers significantly affect auditors' perceptions, and among those with significant impacts, which have the strongest influence?

First, we address data extraction. Although Google's online map service does not offer street-view image service in China, several of China's IT companies are offering similar services, including Tencent, Baidu, AutoNavi, etc. Among these, China's domestic online map service Tencent Maps (http://map.qq.com/) offers more up-to-date street-view images and covers more cities than other online map services. It has several sets of Qiqihar's street-view images on different dates. The latest set of images is from November 2015 and covers most streets within our research area.

To largely decrease the workload of downloading and auditing all street-view images within the research area, we adopt a distance-sampling method with an interval distance of 100 m along streets, so as to select sample points within the study area for extracting their street-view images. Through ArcGIS, we extracted geographic coordinates of all the streets within our research area from the four-dimensional Map, a mobile-car navigation company that offers an API platform (see Fig. 2 left fraction). We then conducted the aforementioned distance sampling method that extracted 3267 points with their geographic coordinates. To evaluate the change of physical environment in future studies, we searched each of the 3267 points' 50-m radius buffer-zone in Tencent Maps, and removed points with only one historical street-view image. The search results in selecting 1213 points out of the 3267 points. The selected points have historical street-view images at two different times, November 2015 and July 2013 (see Fig. 2).

Some related studies (Rundle et al. 2011; Wilson et al. 2012; Mooney et al. 2014) applied image mosaicking that produces omnidirectional street-view images with panoramic views for auditing. Considering the possible distortion of panoramic views that may bring bias to auditors' discretion on identifying and judging, and the omnidirectional images contain too many items that may overwhelm auditors, we do not apply the image mosaicking to generate panoramic views. Instead, we segmented each point's panoramic view into four directional images (front, backward, left, and right of each point), each of which overlays the brim with adjoining images. Thus, we extracted 4852 images from the 1213 points' online street views.

Second, we address identifying physical disorder items. Before employing auditors to label physical disorder items in street-view images, we were faced with the question of how many auditors would be necessary to ensure sufficient credibility. Theoretically, the more auditors we hired and the higher inter-rater reliability (IRR) achieved by the auditors,

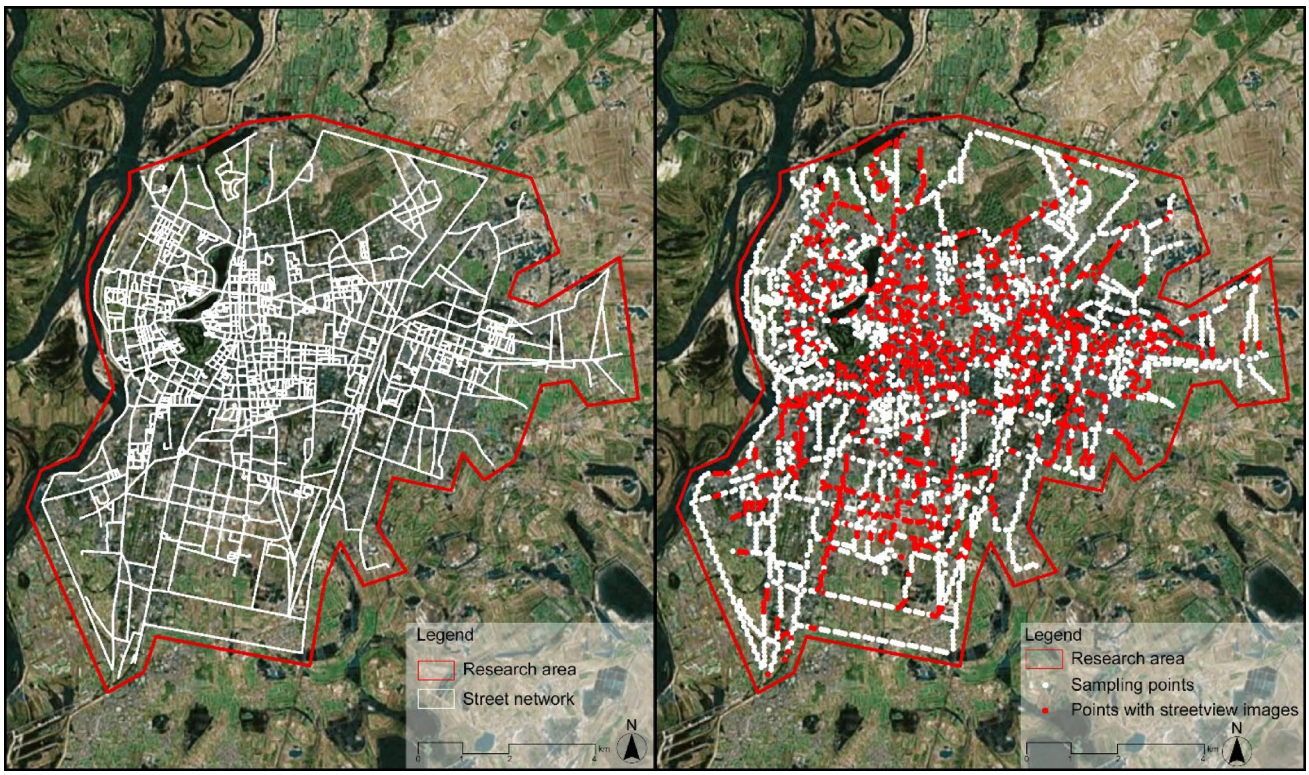


Fig. 2 Qiqihar's street network (*left*); sampling points with street-view images (*right*)

the study would have more credibility for readers. However, restrained by limited resources, we had to hire a reasonable number of auditors. We referred to precedent studies that also hired auditors in virtual auditing (Table 1). The median number of auditors in these studies is around three to four. Therefore, we decided to employ three auditors to save marginal cost of conducting virtual auditing.

Each of the three auditors specialized in either architecture or urban planning. They were trained using reference studies and a training protocol for auditing developed by Mooney et al. (2014). By referring to these precedent studies, the first author constructed a checklist of physical disorder items and sent the checklist to each of the three auditors. Then, each auditor independently checked 500 images that were randomly selected from the 4854 images to see whether the items on the checklist appeared. The auditors deleted the items that they did not find in the images, and added new items that may relate to physical disorder. The first author then collected each auditor's feedback and created a new checklist with 13 items of physical disorder. We compared the item list with the physical disorder items mentioned in precedent studies and categorized them into three groups: similar manifestations, distinct manifestations, barely mentioned. We then calculated inter-rater reliability (Fleiss kappa, a similar method with precedent studies that were summarized in Table 1) to see to what extent did the three auditors reach a consensus on identifying disorder items.

Third, we address perceiving disorder and interpreting factors. While identifying disorder items, auditors were also asked to mark a dichotomous score of whether they perceived disorder from a street-view image (0 = not perceived, 1 = perceived). We hypothesized that the number of disorder items in a street-view image is correlated with its being perceived as disorder by the auditors, and some disorder items outperformed others on affecting auditors' perception. We then applied the point-biserial test and binomial regressions to verify the aforementioned hypotheses.

Findings

Identifying physical disorder items

Our virtual auditing reveals the fact that, although the research of physical disorder is well-grounded in the Western world, a substantial revision is required when conducting related research elsewhere, especially in places where the built environment and ideology of aesthetics are distinct.

After independently auditing sample images, the three auditors revised the checklist by including 13 items that relate to physical disorder in Qiqihar. Each auditor then took the checklist and independently audited all the 4854 images by identifying disorder items that appeared in each image.

We found out that some physical disorder items outperformed others on reaching a much higher inter-rater reliability (Fleiss kappa, see Table 2). Items such as street vendors, vacant stores, potholes, and broken barrier walls all had Fleiss' kappa values exceeding 0.40. These items share a common characteristic: their names and appearances in street-view images can be more easily recognized and linked by auditors. Conversely, items with lower inter-rater reliability suggest that their definitions varied among auditors, highlighting the need for future revisions to the virtual auditing protocol in the context of Chinese cities.

Table 1 A summary of the number of physical disorder items, auditors and IRR of up-to-date virtual auditing studies

Reference	Number of items	Number of auditors	Inter-rater reliability (IRR)
Clarke et al. 2010	29	2	Agreement: 0.78 to 1.00; kappa: 0.27–0.91
Badland et al. 2010	51	1	–
Rundle et al. 2011	143	1	–
Odgers et al. 2012	17	4	Agreement: 0.63–0.94; Kappa: 0.19–0.79
Ben-Joseph et al. 2013	125	4	–
Griew et al. 2013	26	3	Agreement: 0.51–0.89; Fleiss kappa: 0.08–0.60
Kelly et al. 2013	89	4	Kappa: 0.19–1.00; PABAK: 0.60–1.00
Mooney et al. 2014	9	11	Cohen kappa: 0.34–0.80
Mooney et al. 2016	9	5	Kappa: 0.37–1.00
Mooney et al. 2017	19	7	–
Mayne et al. 2018	9	8	Agreement: 0.76–0.93; Kappa: 0.38–0.56
Mayne et al. 2019	9	12	Agreement: 0.77–0.99; Kappa: 0.44–0.67
Nesoff et al. 2020	Over 160	2	Intra-class correlation: 0.02–0.97
Plascak et al. 2020	31	11	–
Mooney et al. 2020	38	6	Kappa: − 0.01–0.92

Table 2 Physical disorder items identified by three auditors

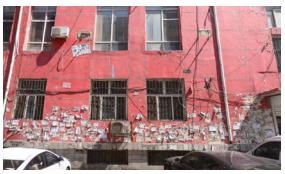
Items	Auditor A	Auditor B	Auditor C	Average	Inter-rater reliability (Fleiss kappa)	Compared with precedent studies
1. Dilapidated building 	33	23	0	19	0.014	b
2. Broken facade 	758	375	635	589	0.273	a
3. Graffiti and leftover advertisements 	244	69	260	191	0.091	b
4. Illegal or temporary structure 	193	68	68	110	0.032	b
5. Neglected storefront 	416	679	127	407	0.181	c
6. Street vendor 	105	146	84	112	0.523	c
7. Vacant store 	55	246	54	118	0.427	c

Table 2 (continued)

Items	Auditor A	Auditor B	Auditor C	Average	Inter-rater reliability (Fleiss kappa)	Compared with precedent studies
8. Trash, debris, and overgrown grass 	491	770	631	631	0.331	b
9. Construction remains 	83	103	16	67	0.232	c
10. Unpaved road 	117	240	55	137	0.198	c
11. Pothole 	53	184	914	384	0.423	a
12. Broken wire or pole 	9	88	5	34	0.108	c
13. Broken barrier wall 	157	83	224	155	0.412	c

a. Items with similar physical manifestations;

b. Items with distinct physical manifestations;

c. Items barely mentioned in former studies

We then compared the checklist with reference studies. By referring to lists of physical disorder items in precedent studies, we categorized our checklist's 13 items into three groups: items with similar physical manifestations, items with distinct physical manifestations, and items scarcely mentioned in earlier studies (see Table 2).

First, we consider items with similar physical manifestations. Compared with precedent studies, these items are not affected by people's preference for architectural style or aesthetic perception across different countries, e.g., broken facades, potholes.

Second, we address items with distinct physical manifestations. In the West, Graffiti is usually illegal writing or drawings within the public view that may be related to gang activity, but generally, it does have commercial meaning. In Qiqihar, most identified Graffiti also has no permission; however, it is either a leftover flyer or a spray-painted advertisement. Trash and debris are frequently mentioned in former studies and they have similar manifestations in Qiqihar. In contrast, when overgrown grass is a sign of badly maintained private property in former studies, it is mostly identified in Qiqihar's public space, e.g., curbside gardens. In addition, although identified illegal or temporary structures in Qiqihar are similar to mobile trailers in the US, they are usually low-end gatehouses or temporary houses for construction workers.

Third, we consider items barely mentioned in former studies. This category contains more than half of physical disorder items in the list. Some items in this category reflect China's distinct urban landscape. A barrier wall is omnipresent in Chinese cities as an important component of walled urban morphology, also a byproduct of the gated community (Xu and Yang 2009). A broken barrier wall thus acts as a sign of negligence, similar to the canonical "broken windows" theory proposed by Wilson and Kelling (1982). Construction remains are also distinct. They signify large-scale ongoing urban development, whereas they are rare in American cities characterized by physical disorder features, as such cities typically have fewer construction activities.

Some other items are categorized into the third group because previous studies primarily focused on neighborhood environments, neglecting other types of public spaces (Ross and Mirowsky 1999, 2001; Sampson and Raudenbush 1999, 2004). Most such items relate to the road (unpaved road, broken wire or poles) and are visually enlarged in street-view images, when compared with conventional disorder items. Others are neglected storefronts and vacant stores. In contrast to the spatial segregation between residential areas and commercial areas in the US, Chinese neighborhoods have much larger block sizes and more mixed land use layouts with stores usually line up on every street.

Street vendors were rarely associated with physical disorder; instead, they were often perceived as a sign of urban vitality in the United States (Ehrenfeucht 2016). However, in China, most street vendors do not have permission and are constantly stigmatized by mass media that link them with unsanitary food or swindlers. As a result, our auditors deemed such urban informality as physical disorder.

Perceived disorder and interpreting contributors

Auditors were also asked to independently give a dichotomous judgment upon their perceptions of each street-view image's level of disorder. On average, 976 out of 4852 street-view images were perceived as overall disorder by the three auditors. Unlike identifying disorder items, auditors did not receive any training before making such decisions. However, the inter-rater reliability among the three auditors turned out to be quite acceptable (Table 3). It had a higher Fleiss kappa value than most of the 13 disorder items identified (Table 2), which triggered our motivation to further interpret factors of such judgment.

In the first place, we assumed that the number of different identified disorder items in a street-view image has a positive correlation with auditors' perception of disorder. We applied the point-biserial correlation test to verify our assumption. The point-biserial test is frequently used by landscape and urban scholars to test the possible association between one dichotomous variable and one continuous variable (Rupprecht et al. 2015). The results indicated that the coefficient values for the three auditors ranged from 0.57 to 0.72. This confirms our assumption regarding the three auditors: the more disorder items identified in a street-view image, the more likely an auditor was to perceive it as disordered (Table 4).

We then assumed that some identified disorder items outperformed others on affecting auditors' perception of disorder. Since all independent variables (1 = item identified, 0 = unidentified) and the dependent variable (1 = perceived disorder, 0 = not perceived) are dichotomous, we applied binomial logistic regression models on each of the three auditors' data, respectively. For each of the three

Table 3 Number of images perceived as disorder by the three auditors

	Auditor A	Auditor B	Auditor C	Average	Inter-rater reliability (Fleiss kappa)
Number	728	1392	808	976	0.523

Table 4 Point-biserial correlation between three auditors' identified number of disorder items in a street-view image and its chance of being perceived as disorder

	Point-biserial correlation		
	Auditor A	Auditor B	Auditor C
Coefficient	0.679832**	0.57643**	0.72743**

**Denotes significance level of $p < 0.01$

regressions, most of the disorder items were statistically significant in predicting their judgment of perceiving overall disorder (Table 5). Items with weak or no statistical significance are usually least frequently identified by auditors. However, since each regression has a distinct set of disorder items with weak or no statistical significance, only seven out of the 13 disorder items have strong statistical significance among three regressions: broken facade, graffiti and leftover advertisements, neglected storefronts, trash, debris and overgrown grass, construction remains, unpaved roads, and broken barrier wall. In general, these disorder items also have higher coefficients and odds ratios than others on predicting each auditor's chance of perceiving an image as overall disorder, a verification of the aforementioned assumption.

Discussion and conclusion

Discussion

Contributions

This paper demonstrates that virtual auditing of physical disorder through the analysis of street-view images is also feasible for studying cities in China, where, to the best of our knowledge, such studies are scarce. Although the research

Table 5 Results of binomial logistic regressions predicting perceiving overall disorder of a street-view image for three auditors

Items	Auditor A			Auditor B			Auditor C		
	Coefficient	OR (95% CI)	S.E.	Coefficient	OR (95% CI)	S.E.	Coefficient	OR (95% CI)	S.E.
1. Dilapidated building	1.610***	5.002 (2.111, 11.850)	0.440	0.521	1.684 (0.506, 5.602)	0.613	–	–	–
2. Broken facade	2.655***	14.225 (11.138, 18.167)	0.125	2.334***	10.318 (7.664, 13.891)	0.152	3.047***	21.049(15.553,28.487)	0.154
3. Graffiti and leftover advertisements	2.367***	10.669 (7.096, 16.042)	0.208	2.552***	12.835 (6.801, 24.221)	0.324	2.146***	8.553(5.530,13.229)	0.223
4. Illegal or temporary structure	1.957***	7.078 (4.602, 10.885)	0.220	0.638	1.892 (0.965, 3.709)	0.343	2.308***	10.052(4.251,23.768)	0.439
5. Neglected storefront	1.562***	4.770 (3.549, 6.412)	0.151	1.960***	7.102 (5.792, 8.708)	0.104	2.153***	8.611(5.105,14.525)	0.267
6. Street vendor	1.108***	3.027 (1.791, 5.115)	0.268	1.497***	4.468 (2.792, 7.150)	0.240	-0.81**0	0.445(0.250,0.793)	0.295
7. Vacant store	.861*	2.366 (1.051, 5.323)	0.414	0.298	1.347 (0.941, 1.926)	0.183	1.412***	4.103(1.908,8.824)	0.391
8. Trash, debris and overgrown grass	2.174***	8.792 (6.562, 11.780)	0.149	2.314***	10.117 (8.248, 12.410)	0.104	3.863***	47.597(34.877,64.958)	0.159
9. Construction remains	1.286***	3.618 (1.743, 7.513)	0.373	1.066***	2.903 (1.760, 4.788)	0.255	3.261***	26.080(6.446,105.521)	0.713
10. Unpaved road	1.251***	3.495 (1.943, 6.288)	0.300	2.012***	7.477 (5.157, 10.841)	0.190	3.055***	21.232(8.279,54.449)	0.481
11. Pothole	− 0.186	0.830 (0.362, 1.904)	0.423	1.847***	6.341 (4.278, 9.399)	0.201	1.783***	5.949(4.595,7.703)	0.132
12. Broken wire or pole	2.254*	9.525 (1.464, 61.977)	0.956	2.022***	7.556 (4.392, 13.000)	0.277	2.969	19.466(0.562,674.268)	1.809
13. Broken barrier wall	1.585***	4.880 (2.979, 7.995)	0.252	1.464***	4.323 (2.417, 7.733)	0.297	1.749***	5.750(3.535,9.355)	0.248

Auditor C did not identify any dilapidated buildings within the 4852 street-view images

***Denotes significance level of $p < 0.001$

**Denotes significance level of $p < 0.01$

*Denotes significance level of $p < 0.05$

field of physical disorder is well established in Western countries, it has not been widely examined in China.

This study established a checklist of physical disorder items for virtual auditing via street-view images, which can act as a foundation for successors who conduct similar studies in China. The study also deepened the understanding of physical disorder by revealing the fact that some physical disorder items listed in precedent studies are not applicable for conducting auditing in China. In addition, we found that some previously identified physical disorder items had distinct manifestations in Chinese cities, such as graffiti, leftover advertisements, and illegal or temporary structures. We also identified unique items not reported in previous studies that contributed to auditors' perceptions of physical disorder, such as construction remains and street vendors.

The significance of this research is twofold. On a theoretical level, the findings of this study suggest that the "broken windows theory" requires further refinement. Firstly, when measuring physical disorder, it is crucial to consider the context-specific applicability and select appropriate spatial disorder elements based on the actual conditions of the study area. This approach can help reshape our understanding of spatial disorder. Secondly, in evaluating urban disorder, oversimplification must be avoided, as different disorder elements may carry varying degrees of influence on resident behavior, such as crime. Future urban improvement initiatives should prioritize addressing those elements with greater impact. Existing research supports this view. Harcourt and Ludwig (2006) empirically challenged the universal applicability of the "broken windows theory." Their study revealed that the straightforward causal link between disorder and crime rates is not always valid, particularly across different urban environments where this relationship may be overstated. Similarly, Ansfield (2020) highlighted the historical and social contexts, noting that the theory may be oversimplified or misapplied in certain regions, such as the Bronx, where structural factors like poverty and racial segregation exert a far more profound influence on disorder than the apparent surface-level issues. On a practical level, these findings also provide guidelines for local government to improve the quality of built environments by addressing major physical disorder items, as well as prioritizing urban beautification in places perceived as disorder by our auditors.

Policy recommendations

With the progress of urbanization in China, the priority of urban construction and management is gradually transferred from new-town development to urban redevelopment. Thus, improving the quality of built environments is many Chinese politicians' top agenda. To date, the common way to initiate urban redevelopment in China is either market-based, meaning developed parcels with relatively high land value attract developers to invest; or complaint-based, meaning when developed parcels are frequently complained by nearby residents, they may have a large chance to be prioritized by government-led redevelopment. Virtual auditing via street-view images provides a brand-new way to monitor changes of built environment and supports decision-making on where to start with urban redevelopment. Compared to conventional approaches to urban redevelopment, virtual auditing can identify more minor urban issues and prompt small-scale redevelopment or environmental remediation, rather than requiring costly large-scale redevelopment efforts.

Although many physical features can contribute to people's perceptions of disorder, this study suggests that addressing a few key physical items can significantly improve the quality of the built environment. Some physical disorder items, such as graffiti, leftover advertisements, trash, debris, and overgrown grass, are relatively easier to manage compared to more challenging issues like dilapidated buildings and illegal or temporary structures. However, this study demonstrates that the formers' remediation will bring a greater effect on people's perceptions of urban environment than the latter. This provides a valuable suggestion for local authorities seeking to improve the quality of the built environment at a low cost, particularly in Qiqihar and other cities in China's northeast region, where public expenditures far exceed public revenues.

Limitations

This study has several limitations. First, we used only Qiqihar as a case city, and the sample size of street-view images was relatively small, only 4325 SVIs, which may introduce bias and limit the generalizability of the findings. Future research should extend this approach to a broader range of cities. Second, while we employed virtual auditing for street-view image assessments, this remains a labor-intensive process. Fortunately, emerging tools such as deep learning models offer more efficient alternatives, which could be leveraged in future studies. Additionally, as discussed in the literature review, street-view imagery is limited to areas accessible by vehicles, meaning that some pedestrian pathways and their associated features cannot be adequately observed. Virtual audits may also be constrained by the use of outdated street-view images.

Conclusion

This study reveals that virtual auditing via street-view images is also feasible for studying physical disorder in Chinese cities. Based on precedent studies and existing training protocols, three employed auditors developed a checklist with 13 physical disorder items that were extracted from

sampled street-view images of Qiqihar. While two physical disorder items in Qiqihar exhibited similar manifestations to those described in previous studies, eleven other items either had distinct manifestations or were rarely mentioned before, thereby broadening the scope of current physical disorder research. Among the 13 disorder items, auditors reach higher inter-rater reliability on the following items: street vendor, vacant store, pothole, broken barrier wall and trash, debris, and overgrown grass. Though we did not train our auditors on how to judge overall perceived disorder of a street-view image, our three auditors reached a consensus via a dichotomous judgement of street-view images. We also found that among each of the three auditors, the number of disorder items in a street-view image was strongly correlated with its chance of being perceived as disorder by the auditors. In the three regressions that explain what items contribute to auditors' perceptions of overall disorder, seven out of the 13 disorder items have stronger statistical significance than others: broken facade, graffiti and leftover advertisements, neglected storefront, trash, debris and overgrown grass, construction remains, unpaved road, and broken barrier wall. This suggests that these items should be prioritized for intervention if local authorities aim to enhance the urban physical environment. Most of the aforementioned findings highlight the significance of this study, demonstrating that conducting virtual auditing of physical disorder in Chinese cities is valuable for being cost-effective, informative, and innovative. While this study has some limitations, the virtual auditing method paves a new pathway for studying physical disorder in China and, due to the wide availability of street-view data, could also be effectively applied to non-Chinese case studies, where it may offer comparable or even enhanced insights into urban disorder.

Acknowledgements This study was funded by the National Natural Science Foundation of China (No. 52178044, No. 52478044, No. 51778319, No. 62394331, and No. 62394335). We thank Anna Morgan for helping us in naming and phrasing physical disorder items.

Data availability The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

References

Abbott, J. 2005. Understanding and managing the unknown: The nature of uncertainty in planning. *Journal of Planning Education and Research* 24 (3): 237–251.

Allen, R. 2013. The distribution and evolution of physical neighborhood problems during the great recession. *City & Community* 12 (3): 260–279.

Ansfield, B. 2020. The broken windows of the Bronx: Putting the theory in its place. *American Quarterly* 72 (1): 103–127.

Attane, I. 2002. China's family planning policy: An overview of its past and future. *Studies in Family Planning* 33 (1): 103–113.

Bader, M.D., S.J. Mooney, Y.J. Lee, D. Sheehan, K.M. Neckerman, A.G. Rundle, and J.O. Teitler. 2015. Development and deployment of the computer assisted neighborhood visual assessment system (CANVAS) to measure health-related neighborhood conditions. *Health & Place* 31: 163–172.

Bader, M.D., S.J. Mooney, B. Bennett, and A.G. Rundle. 2017. The promise, practicalities, and perils of virtually auditing neighborhoods using Google Street View. *The ANNALS of the American Academy of Political and Social Science* 669 (1): 18–40.

Badland, H.M., S. Opit, K. Witten, R.A. Kearns, and S. Mavoa. 2010. Can virtual streetscape audits reliably replace physical streetscape audits? *Journal of Urban Health* 87 (6): 1007–1016.

Beauregard, R. 2021. The entanglements of uncertainty. *Journal of Planning Education and Research* 41 (2): 217–225.

Ben-Joseph, E., J.S. Lee, E.K. Cromley, F. Laden, and P.J. Troped. 2013. Virtual and actual: Relative accuracy of on-site and web-based instruments in auditing the environment for physical activity. *Health & Place* 19: 138–150.

Body-Gendrot, S. 2014. Public disorders: Theory and practice. *Annual Review of Law and Social Science* 10: 243–258.

Calder, K., and M. Ye. 2010. *The making of Northeast Asia*. Stanford University Press.

Carbone, J.T. 2020. The mediating effect of allostatic load on the relationship between neighborhood perceptions and depression. *SSM-Population Health* 11: 100638.

Chen, X., and P. Rafail. 2021. Physical disorder and crime revisited: New evidence from intensive longitudinal data. *Social Science Research*. https://doi.org/10.1016/j.ssresearch.2021.102637.

Chung, J.H., H. Lai, and J.H. Joo. 2009. Assessing the "Revive the Northeast"(zhenxing dongbei) programme: Origins, policies and implementation. *The China Quarterly* 197: 108–125.

Clarke, P., J. Ailshire, R. Melendez, M. Bader, and J. Morenoff. 2010. Using Google Earth to conduct a neighborhood audit: Reliability of a virtual audit instrument. *Health & Place* 16 (6): 1224–1229.

Deng, C., and J. Ma. 2015. Viewing urban decay from the sky: A multi-scale analysis of residential vacancy in a shrinking US city. *Landscape and Urban Planning* 141: 88–99.

Ehrenfeucht, R. 2016. Designing fair and effective street vending policy: It's time for a new approach. *Cityscape* 18 (1): 11–26.

Fan, C.C., and M. Sun. 2008. Regional inequality in China, 1978–2006. *Eurasian Geography and Economics* 49 (1): 1–18.

Feuillet, T., H. Charreire, C. Roda, M. Ben Rebah, J.D. Mackenbach, S. Compernolle, K. Glonti, H. Bárdos, H. Rutter, I. De Bourdeaudhuij, M. McKee, J. Brug, J. Lakerveld, and J.-M. Oppert. 2016. Neighbourhood typology based on virtual audit of environmental obesogenic characteristics. *Obesity Reviews* 17: 19–30.

Franzini, L., M.O.B. Caughy, S.M. Nettles, and P. O'Campo. 2008. Perceptions of disorder: Contributions of neighborhood characteristics to subjective perceptions of disorder. *Journal of Environmental Psychology* 28 (1): 83–93.

Gao, S., and B.D. Ryan. 2024. Can neighborhood planning in shrinking cities achieve demolition goals? A conformance and performance evaluation of neighborhood action plans in Youngstown, Ohio. *Journal of Planning Education and Research* 44 (1): 143–164.

Gao, S., H. Jansen, and B.D. Ryan. 2023. Demolition after decline: Understanding and explaining demolition patterns in US and German shrinking cities. *Cities* 134: 104185.

Gau, J.M., and T.C. Pratt. 2010. Revisiting broken windows theory: Examining the sources of the discriminant validity of perceived disorder and crime. *Journal of Criminal Justice* 38 (4): 758–766.

Gold, S., and L. Nepomnyaschy. 2018. Neighborhood physical disorder and early delinquency among urban children. *Journal of Marriage and Family* 80 (4): 919–933.

Griew, P., M. Hillsdon, C. Foster, E. Coombes, A. Jones, and P. Wilkinson. 2013. Developing and testing a street audit tool using Google Street View to measure environmental supportiveness for physical

activity. *International Journal of Behavioral Nutrition and Physical Activity* 10 (1): 1–7.

Grubesic, T.H., D. Wallace, A.W. Chamberlain, and J.R. Nelson. 2018. Using unmanned aerial systems (UAS) for remotely sensing physical disorder in neighborhoods. *Landscape and Urban Planning* 169: 148–159.

Harcourt, B.E., and J. Ludwig. 2006. Broken windows: New evidence from New York City and a five-city social experiment. *U. Chi. L. Rev.* 73: 271.

Hendricks, M.D., and M.A. Meyer. 2021. Modeling long-term housing recovery after technological disaster using a virtual audit with repeated photography. *Journal of Planning Education and Research* 44 (2): 701–719.

Hoeben, E.M., W. Steenbeek, and L.J. Pauwels. 2018. Measuring disorder: Observer bias in systematic social observations at streets and neighborhoods. *Journal of Quantitative Criminology* 34 (1): 221–249.

Hwang, J. 2017. Invited commentary: Observing neighborhood physical disorder in an age of technological innovation. *American Journal of Epidemiology* 186 (3): 274–277.

Jones, M., A.R. Pebley, and N. Sastry. 2011. Eyes on the block: Measuring urban physical disorder through in-person observation. *Social Science Research* 40 (2): 523–537.

Kelly, C.M., J.S. Wilson, E.A. Baker, D.K. Miller, and M. Schootman. 2013. Using Google Street View to audit the built environment: inter-rater reliability results. *Annals of Behavioral Medicine* 45 (suppl_1): S108–S112.

Khorshidi, S., J. Carter, G. Mohler, and G. Tita. 2021. Explaining crime diversity with Google Street View. *Journal of Quantitative Criminology* 37 (2): 361–391.

Kotus, J., and M. Rzeszewski. 2013. Between disorder and livability. Case of one street in post-socialist city. *Cities* 32: 123–134.

Krieger, M.H. 2022. Riding Uncertainty. *Journal of Planning Education and Research* 42 (3): 482–486.

Li, Z., and F. Wu. 2013. Residential satisfaction in China's informal settlements: A case study of Beijing, Shanghai, and Guangzhou. *Urban Geography* 34 (7): 923–949.

Li, X., C. Zhang, W. Li, R. Ricard, Q. Meng, and W. Zhang. 2015. Assessing street-level urban greenery using Google Street View and a modified green view index. *Urban Forestry & Urban Greening* 14 (3): 675–685.

Long, Y., and L. Liu. 2017. How green are the streets? An analysis for central areas of Chinese cities using Tencent Street View. *PLoS ONE* 12 (2): e0171110.

Ma, Y., N. Su, and T. Tu. 2023. Urban public space quality evaluation methods and practices in China. *Transactions in Urban Data, Science, and Technology* 2 (2–3): 59–80.

Marco, M., E. Gracia, J.M. Tomás, and A. López-Quílez. 2015. Assessing neighborhood disorder: Validation of a three-factor observational scale. *The European Journal of Psychology Applied to Legal Context* 7 (2): 81–89.

Marco, M., E. Gracia, M. Martín-Fernández, and A. López-Quílez. 2017. Validation of a Google Street View-based neighborhood disorder observational scale. *Journal of Urban Health: Bulletin of the New York Academy of Medicine* 94 (2): 190.

Mayne, S.L., A. Jose, A. Mo, L. Vo, S. Rachapalli, H. Ali, J. Davis, and K.N. Kershaw. 2018. Neighborhood disorder and obesity-related outcomes among women in Chicago. *International Journal of Environmental Research and Public Health* 15 (7): 1395.

Mayne, S.L., B.F. Pellissier, and K.N. Kershaw. 2019. Neighborhood physical disorder and adverse pregnancy outcomes among women in Chicago: A cross-sectional analysis of electronic health record data. *Journal of Urban Health* 96 (6): 823–834.

Mooney, S.J., M.D.M. Bader, G.S. Lovasi, K.M. Neckerman, J.O. Teitler, and A.G. Rundle. 2014. Validity of an ecometric neighborhood physical disorder measure constructed by Virtual Street Audit. *American Journal of Epidemiology* 180 (6): 626–635.

Mooney, S.J., C.J. DiMaggio, G.S. Lovasi, K.M. Neckerman, M.D.M. Bader, J.O. Teitler, D.M. Sheehan, D.W. Jack, and A.G. Rundle. 2016. Use of Google Street View to assess environmental contributions to pedestrian injury. *American Journal of Public Health* 106 (3): 462–469.

Mooney, S.J., M.D.M. Bader, G.S. Lovasi, J.O. Teitler, K.C. Koenen, A.E. Aiello, S. Galea, E. Goldmann, D.M. Sheehan, and A.G. Rundle. 2017. Street audits to measure neighborhood disorder: virtual or in-person? *American Journal of Epidemiology* 186 (3): 265–273.

Mooney, S.J., K. Wheeler-Martin, L.M. Fiedler, C.M. LaBelle, T. Lampe, A. Ratanatharathorn, N.N. Shah, A.G. Rundle, and C.J. DiMaggio. 2020. Development and validation of a Google Street View pedestrian safety audit tool. *Epidemiology (Cambridge, Mass.)* 31 (2): 301–309.

Murphy, A.K. 2012. "Litterers" how objects of physical disorder are used to construct subjects of social disorder in a suburb. *The ANNALS of the American Academy of Political and Social Science* 642 (1): 210–227.

Nesoff, E.D., A.J. Milam, C.B. Barajas, and C.D.M. Furr-Holden. 2020. Expanding tools for investigating neighborhood indicators of drug use and violence: Validation of the NIfETy for virtual street observation. *Prevention Science* 21 (2): 203–210.

Odgers, C.L., A. Caspi, C.J. Bates, R.J. Sampson, and T.E. Moffitt. 2012. Systematic social observation of children's neighborhoods using Google Street View: A reliable and cost-effective method. *Journal of Child Psychology and Psychiatry* 53 (10): 1009–1017.

Park, K., and R. Ewing. 2022. The usability of unmanned aerial vehicles (UAVs) for pedestrian observation. *Journal of Planning Education and Research* 42 (2): 206–217.

Perkins, D.D., and R.B. Taylor. 1996. Ecological assessments of community disorder: Their relationship to fear of crime and theoretical implications. *American Journal of Community Psychology* 24 (1): 63–107.

Perkins, D.D., A. Wandersman, R.C. Rich, and R.B. Taylor. 1993. The physical environment of street crime: Defensible space, territoriality and incivilities. *Journal of Environmental Psychology* 13 (1): 29–49.

Plascak, J.J., M. Schootman, A.G. Rundle, C. Xing, A.A. Llanos, A.M. Stroup, and S.J. Mooney. 2020. Spatial predictive properties of built environment characteristics assessed by drop-and-spin virtual neighborhood auditing. *International Journal of Health Geographics* 19 (1): 1–20.

Raudenbush, S.W., and R.J. Sampson. 1999. Ecometrics: Toward a science of assessing ecological settings, with application to the systematic social observation of neighborhoods. *Sociological Methodology* 29 (1): 1–41.

Rigolon, A., D. Banerjee, P. Gobster, S. Hadavi, and W. Stewart. 2021. Transferring vacant lots to private ownership improves care and empowers residents: Evidence from Chicago. *Journal of the American Planning Association.* 87 (4): 570–584.

Rohe, W.M., and R.J. Burby. 1988. Fear of crime in public housing. *Environment and Behavior* 20 (6): 700–720.

Rohe, W.M., and M.A. Stegman. 1994. The impact of home ownership on the social and political involvement of low-income people. *Urban Affairs Quarterly* 30 (1): 152–172.

Ross, C.E., and J. Mirowsky. 1999. Disorder and decay: The concept and measurement of perceived neighborhood disorder. *Urban Affairs Review* 34 (3): 412–432.

Ross, C.E., and J. Mirowsky. 2001. Neighborhood disadvantage, disorder, and health. *Journal of Health and Social Behavior*. https://doi.org/10.2307/3090214.

Ross, C.E., J. Mirowsky, and S. Pribesh. 2001. Powerlessness and the amplification of threat: Neighborhood disadvantage, disorder, and mistrust. *American Sociological Review*. https://doi.org/10.1177/000312240106600405.

Rundle, A.G., M.D. Bader, C.A. Richards, K.M. Neckerman, and J.O. Teitler. 2011. Using Google Street View to audit neighborhood environments. *American Journal of Preventive Medicine* 40 (1): 94–100.

Rupprecht, C.D., J.A. Byrne, H. Ueda, and A.Y. Lo. 2015. 'It's real, not fake like a park': Residents' perception and use of informal urban green-space in Brisbane, Australia and Sapporo, Japan. *Landscape and Urban Planning* 143: 205–218.

Sampson, R.J. 2009. Disparity and diversity in the contemporary city: Social (dis) order revisited. *The British Journal of Sociology* 60 (1): 1–31.

Sampson, R.J., and S.W. Raudenbush. 1999. Systematic social observation of public spaces: A new look at disorder in urban neighborhoods. *American Journal of Sociology* 105 (3): 603–651.

Sampson, R.J., and S.W. Raudenbush. 2004. Seeing disorder: Neighborhood stigma and the social construction of "broken windows." *Social Psychology Quarterly* 67 (4): 319–342.

Schootman, M., E.J. Nelson, K. Werner, E. Shacham, M. Elliott, K. Ratnapradipa, M. Lian, and A. McVay. 2016. Emerging technologies to measure neighborhood conditions in public health: implications for interventions and next steps. *International Journal of Health Geographics* 15 (20): 1–9.

Skogan, W.G. 1990. *Disorder and decline: Crime and the spiral of decay in American neighborhoods*. Berkeley: University of California Press.

Skogan, W. 2015. Disorder and decline: The state of research. *Journal of Research in Crime and Delinquency.* 52 (4): 464–485.

Spano, R. 2005. Potential sources of observer bias in police observational data. *Social Science Research* 34 (3): 591–617.

Statistical Bureau of Heilongjiang. (2021). District-level of residents in Heilongjiang from the 7th National Census. http://tjj.hlj.gov.cn/zxfb/202106/t20210607_87860.html (Accessed Feb 8, 2022).

Stern, M., and W. Lester. 2021. Does local ownership of vacant land reduce crime? *Journal of the American Planning Association* 87 (1): 73–84.

Sytsma, V.A., N. Connealy, and E.L. Piza. 2021. Environmental predictors of a drug offender crime script: A systematic social observation of Google Street View images and CCTV footage. *Crime & Delinquency* 67 (1): 27–57.

Taylor, R.B., and J. Covington. 1993. Community structural change and fear of crime. *Social Problems* 40 (3): 374–397.

Wallace, D., and D. Schalliol. 2015. Testing the temporal nature of social disorder through abandoned buildings and interstitial spaces. *Social Science Research* 54: 177–194.

Weden, M.M., R.M. Carpiano, and S.A. Robert. 2008. Subjective and objective neighborhood characteristics and adult health. *Social Science & Medicine* 66 (6): 1256–1270.

Wilson, J.Q., and G.L. Kelling. 1982. Broken windows. *Atlantic Monthly* 249 (3): 29–38.

Wilson, J.S., C.M. Kelly, M. Schootman, E.A. Baker, A. Banerjee, M. Clennin, and D.K. Miller. 2012. Assessing the built environment using omnidirectional imagery. *American Journal of Preventive Medicine* 42 (2): 193–199.

Xie, L., Z. Yang, J. Cai, Z. Cheng, T. Wen, and T. Song. 2016. Harbin: A rust belt city revival from its strategic position. *Cities* 58: 26–38.

Xu, M., and Z. Yang. 2009. Design history of China's gated cities and neighbourhoods: Prototype and evolution. *Urban Design International* 14 (2): 99–117.

Zhang, E., H. Xie, and Y. Long. 2023. Decoding the association between urban streetscape skeletons and urban activities: Experiments in Beijing using Dazhong Dianping data. *Transactions in Urban Data, Science, and Technology* 2 (1): 3–18.

Zhou, K., J. Koutsky, and J. Hollander. 2022. Urban shrinkage in China. *The USA and the Czech Republic: A Comparative Multilevel Governance Perspective* 46 (3): 480–496.