Article

Urban Analytics and City Science

Re-examining Jane Jacobs' doctrine using new urban data in Hong Kong

EPB: Urban Analytics and City Science
2023, Vol. 50(1) 76–93

Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/23998083221106186
journals.sagepub.com/home/epb
SAGE

Jianxiang Huang, D.Des.
Sustainable High Density Cities Lab, HKUrban Lab, The University of Hong Kong, Pokfolam Rd. Hong Kong SAR, China;

The University of Hong Kong Shenzhen Institute of Research and Innovation, Shenzhen Virtual University Park, Shenzhen, China

Yuming Cui
Department of Architecture, 7/F Knowles Building, The University of Hong Kong, Pokfulam Rd., Hong Kong, China

Lishuai Li
School of Data Science, City University of Hong Kong, 16-201, 16/F, Lau Ming Wai Academic Building, 83 Tat Chee Avenue, Hong Kong SAR, China

Mengdi Guo, Ph.D.
The University of Hong Kong Shenzhen Institute of Research and Innovation, Shenzhen Virtual University Park, Shenzhen, China;

Department of Architecture, School of Architecture, Tianjin University, Tianjin, China

Hung Chak Ho, Ph.D.
Department of Urban Planning and Design, Faculty of Architecture, The University of Hong Kong, Hong Kong SAR, China; Department of Anaesthesiology, School of Clinical Medicine, LKS Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China

Yi Lu, Ph.D.
Department of Architecture and Civil Engineering, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong SAR, China

Chris Webster, Ph.D.
Faculty of Architecture, The University of Hong Kong, Hong Kong SAR, China;

Healthy High Density Cities Lab, HKUrban Lab, The University of Hong Kong, Pokfulam Rd., Hong Kong SAR, China

Abstract

Jane Jacobs (1961) theorized that four urban form conditions, namely, mixed use, short block, aged buildings and density, are indispensable for the 'exuberant diversity' and conducive to, or perhaps

Corresponding author:
Jianxiang Huang, Department of Urban Planning and Design, The University of Hong Kong, 8/F Knowles Building, Pokfulam Rd, Hong Kong SAR, China.
Email: jxhuang@hku.hk

even determinant of, the success of a city district. Jacobs' theory has been used widely as a reference point in case study research and policy and design prescriptions. We found five studies that attempted to test it more formally, using various performance indicators such as mobile phone activities, walking, crime and mortality. Their findings were inconsistent and unable to settle theoretical controversies. Questions remained as what performance indicators are most strongly associated with her urban form conditions? Are these conditions independently associated with desired outcomes or in combination and what are the interaction effects? Our study aimed to test Jacobs' theory that urban form conditions contribute to the vitality and success of city districts. Jacobs' urban form conditions were measured using GIS data for each of Hong Kong's Tertiary Planning Unit. Performance outcomes were gauged using a combination of 'new urban data', comprising Twitter activities, sentiment tones and Point-Of-Interest (POI), and 'traditional data', comprising walking commute, employment and mortality. Urban context, income and demographic indicators were used as controls in fitting spatial regression models to predict measures of performances based on urban conditions. Results showed that Jacobs' urban form conditions contribute positively to 'vitality' indicators such as the density of tweets, walking trips and POI, but not with 'failure and success' indicators such as expressed sentiment on Twitter, employment, or mortality. Out findings suggest that her theory largely hold for Hong Kong, except that dwelling density should be substituted by building density, whilst tall buildings associated positively with desirable outcomes, contrary to Jacobs' observation in the American context. More generally, we demonstrate how new urban data can be used to evaluate classic planning theories at scale.

Keywords
Jane Jacobs, urban form, social media, walking activities, vitality

Introduction

The 'Death and Life of Great American Cities' (Jacobs, 1961) provided a landmark urban planning theory with abiding influences. Jacobs' key message was that certain urban form conditions, notably mixed use, short block, aged buildings and dwelling density, are conducive to the success of a city district by realizing 'its best potentials', producing 'exuberant diversity' and giving 'best chances to city life' (ibid, p.150–151). Whilst many refer to Jacobs' insights as 'theory', their original formulation is clearly more accurately described as powerfully narrated observation than a system of well-tested hypotheses. But her observations have become established wisdom and working doctrine for two generations of urban designers and planners, from the new urbanist' disregard for low density, car-dependent suburbs (Katz, 1993) to Montgomery's narratives (2013) that densely populated cities with a mixture of land use, housing types incentivize walking, social interactions and happiness. Since many would claim that Jacobs' doctrine have been tested in numerous design and planning schemes, we can perhaps accept them also as theory, on the assumption that the case study (a design scheme) can be a powerful research design for falsifying hypotheses (Flyvbjerg, 2006). We also note with caution, however, that it is hard to imagine how any failed urban design scheme could be used to falsify Jane Jacobs' theory in any meaningful way. In this sense, it is probably more correct to view it as doctrine or what Anglo-Austrian philosopher of science Karl Popper called a 'pseudo-science' theory (Webster, 2015).

However, there remains dissenting voices that demand more evidence in support of her assertions (Hill, 1988; Marshall, 2012). Louis Mumford and Herbert Gans criticised her work as 'anecdotal' and 'lack of rigor' (Harris, 2011, p.85). Contemporary urban economist Edward Glaeser argued that Jacobs' idea of preserving aged buildings in order to maintain neighbourhood affordability 'is not how supply and demand work' (Glaeser, 2010). He further contested her view of tall buildings, which he notes, 'gave factory owners and workers space that was both more humane and more efficient' (Glaeser, 2011). Some note that contemporary cities have very different demographic and economic conditions than in Jacobs' time (Kotkin, 2000). Recently, Trenkner (2011) raised the question, heretical for some, of whether it is time to 'retire' Jane Jacobs' vision of the city.

Empirical testing of Jacobs' theory remained thin in research literature. A survey of academic papers returned five studies in this category (De Nadai et al., 2016a; Fowler, 1987; Schmidt, 1977; Sung et al., 2015; Weicher, 1973). The results, two supporting versus three disproving to some degree, were insufficient to settle the theoretical controversies, not to mention these five studies varied in the choice of performance indicators, statistical rigour and geographical locations. Questions remained as to whether Jacobs' key message hold under a comprehensive research framework? What performance outcomes are most strongly associated with her urban form conditions? Are these conditions independently associated with desired outcomes or in interaction with confounders?

The aim of the study was to test the hypothesis that Jacobs' urban conditions are conducive to urban vitality. We translated Jacobs' observational-doctrine into the following generalised empirical model in Formula (1), based on a version first developed by Weicher (1973) and subsequently adapted by others.

$$P = f(M, B, A, D, H, C, e) \tag{1}$$

P = performance of a city district, including 'vitality' or 'failure and success'. The former was measured using social media activities, sentiment and point-of-interests, commuting mode choice, whereas the latter using employment, income, mortality and deprivation indices.
M = Land use mixture, measured by an entropy formula
B = Block configuration, such as block size or road junction density A = Aged building
D = Density (people, buildings, jobs, etc.,)
H = Building height
C = Confounding variables that should have an independent association with or effect on P and should be controlled for in an empirical test.
e = An error term capturing unmodelled influences on P

Relevant works

Empirical literature testing Jacobs' theory remained thin. Recent developments in data analytics supports the revisiting of Jacobs. Her performance outcomes have been captured by evidence from 'new urban data', such as social media and mobile phone, whilst her inputs have been operationalized by a recent wave of literature measuring urban morphology and performance, with or without referencing her work.

Testing Jane Jacobs' theory

A survey of academic publications yielded five studies in this category, a scarcity compared with the bulk of literature citing Jacobs more broadly. Most researchers accepted her premise as a given without critical scrutiny. Three out of the five studies refuted Jacobs to some degree; two supported her partially. Weicher (1973) measured the 'failure and success' of neighbourhoods in Chicago, USA using death, delinquency, mental health records and overall diversity. He rejected Jacobs' idea after failing to find supporting evidence, suggesting that it may be 'only applicable to Eastern cities of higher density'. A follow up study conducted by Schmidt (1977) in Denver, USA, using similar outcomes indicators with the Weicher study except for substituting mortality with disease and crime, provided little supporting evidence for Jacobs. Fowler (1987) studied neighbourhood 'failure and success' in Toronto, Canada, using crime incidence obtained from police and self-reported friendship and safety as outcome indicators. His findings were mixed, and the study lent no support for Jacobs. More supportive evidence came from two recent papers: Sung et al. (2015) used walking activities in Seoul, South Korea as the outcome indicator. They found walking behaviours to be positively associated with short block, aged building and job density, but not with mixed use nor dwelling density. De Nadai et al., (2016a) studied mobile phone activities and urban form conditions in Bologna, Florence, Milan, Palermo, Rome and Turin in Italy. They reported that mobile phone activities are positively associated with short block, work density and (the lack of) border vacuum of highways; but not to mixed use nor aged building, nor dwelling density. They partially supported Jacobs whilst acknowledging that super-blocks rarely exist in European cities, where aged buildings are ubiquitous. The five studies above contradicted with each other in findings. They are also inconsistent in technical terms, from the choice of performance indicators to the use of controls, to say nothing of the differences in sample size, unit of analysis and study locations (Supplemental Table S1). These inconsistencies beg the question whether their inconclusiveness in findings is due to the varying robustness of research design, or Jacobs' theory in explaining evidence across geographical regions.

Measuring the outcomes

Jacobs (1961) used a variety of constructs to describe the performances of a city district, which can be categorized as 'vitality' or 'failure and success'. The former category is often viewed as a collective outcome of a place. In its simplest measurable form, 'vitality' reduces to pedestrian density or walking, and it has been extensively captured by transport and travel behaviour studies (Ewing and Cervero, 2010) often without referencing Jacobs. Walking has also been studied in another sub-field that has recently flourished in the domain of public health, active travel and urban greenness. For example a study linked street greenery and lower density with recreational physical activities in Hong Kong (Lu, 2019). An alternative conceptualisation is to consider 'vitality' a complex multi-factor outcome influenced by economic, social and diversity, which can be best captured by 'new urban data', a consortium of data collected from mobile devices, social media platforms and online information. For example researchers measured urban vitality using the density of POI (Kang et al., 2020) or small catering business (Ye et al., 2018), a sub-category of the POI dataset which depend on and attract pedestrian foot traffic. Others used the density of Twitter or Instagram activities (Huang et al., 2021, 2022) as proxies for vitality.

The latter category consists of the 'failure and success' of a city district, alternatively referred to as 'crime', 'death', 'deprivation' and 'happiness' in Jacobs' descriptions. They are of the same order as, and independent from 'vitality'. 'Crime' was often measured using crime incident statistics (Traunmueller et al., 2014). By 'death' Jacobs probably meant 'deprivation', the negative outcomes of a city district, which has been captured by social studies as part of a long tradition dating back to 19th century social-spatial surveys (Booth, 1895). Although 'death' was occasionally interpreted literally as the death of residents, such as in Weicher's study (1973) using neighbourhood mortality data. We come back to this at the 'methods' section of the paper, when we select our own indicators on 'failure and success'. 'Happiness' is often captured by survey instruments, and recently, by sentiment tones expressed in social media messages. For instance strong sentiments, happy or sad, have been found near parks, transport hubs and polluted areas in New York based on geo-coded Twitter analysis (Bertrand et al., 2013). Another approach to crowd-sourcing sentiment data is user-generated votes collected online. With this method, quantity of greenery has been shown to be positively associated with pleasant perceptions, whilst wide streets, fortress-like buildings and public housing tend to be associated with the opposite (Quercia et al., 2014).

Measuring the inputs

The four Jacobs' conditions considered essential for urban vitality, namely land use mixture, short blocks, density and aged buildings, have been captured by a recent wave of studies on urban morphology and performances, thanks to the advancement of urban spatial data analytics. Land use mixture are often measured using entropy, a mathematical measure of the degree of disorder of a system. A two-dimensional land use mixture was computed by Bouyer et al. (2011), whilst its three-dimensional equivalence was proposed by Yue et al. (2017) using categorical POI data in dense cities where housing on top of shops are common. Interestingly, land use diversity was regarded as both an outcome and an input, based on the belief that the diversity of urban functions is a measure of vitality on its own. Examples can be found in a recent paper (Xia et al., 2020). Land use diversity and block size, often measured using the density of street intersections, were found to associate positively with walking and cycling. Block size, measured using GIS polygon database by Long and Haung (2017), was found to have associated negatively with urban economic vitality. Similar observations were made in transport literature, in which block size was often measured using the density of street intersections, and small block was found to correlate positively with behaviours such as walking and cycling (Ewing and Cervero 2010). Density was measured as the concentration of people, jobs or buildings, and it is one of the most studied input variables in literature. A higher density was found to associate positively with social interactions and economic activities (Ye et al., 2018). Aged buildings, which have been empirically linked to increased social interactions and housing affordability (Montgomery, 2013), can be efficiently analysed at scale using municipal property databases.

Jacobs also referred to 'tall building' and 'border vacuum' as negative urban conditions. The former was associated with crime, hindered child development, restrained social relations, poor physical and mental health (Newman, 1975; Oda et al., 1989) , although more recent evidence tend to suggest otherwise (Yeung and Wong, 2003). 'Border vacuums', meaning the hard edges of a city district such as railroads, highways or dead-end alleys, was regarded as barriers of pedestrian movement and sources of neighbourhood plight by Jacobs (1961). Yet 'border vacuum' was less studied compared with others, possibly because there is less of it in sleek, post-industrial metropolis compared with those in Jacobs' time.

Research gap

We note that so far, there have been no studies using social media, POI or related geospatial data analysis, to evaluate Jacobs. There is a need for a research design derived systematically from Jacobs' theory which brings together 'new data', geospatial data analysis, and adding to the five earlier studies. Such a study allows one to test the strength of association between a range of Jacobian performance outcomes and urban form conditions, and to evaluate whether such association is independent or in interaction with social-demographic confounders.

Methods

We developed an empirical research design in accordance with Jacobs' theory and key constructs (Supplemental Figure S1). Urban form conditions, performance indicators and confounding variables were computed using a combination of geo-coded social media data, building environment data and census data. Hypothesis-testing was conducting using statistical models.

Location

The study is in Hong Kong (HK), a city of 7.29 million with a range of urban form conditions, including the compact historical core, large housing estates in new towns, and 'rural' village houses. HK has a density spectrum close to Jane Jacobs' observation on the compact North American cities of the 1950's (Supplemental Figure S2). The city keeps a comprehensive database on its population and the built environment at the level of the Tertiary Planning Unit (TPU), equivalent to the census tract. It ranks high globally on social media penetration, internet connection speed and smart phone usage (Go-Globe, 2015), making it an ideal location for both 'new data' and 'old data' studies. We also note that Jacobs' ideas have been influential in HK's urban planners' circle (Ng, 2000).

Urban form conditions

We developed a total of five urban form indicators in accordance with Jacobs' theory. The first four responded directly to Jacobs' remark '… by deliberately inducing these four conditions, planning can induce city vitality …' (1961, p.14). We added tall building as the fifth, due to Jacobs' apparent dislike of it and its abundance in HK and increasingly elsewhere. Each was operationalized in reference to literature and computed using various data sources.

- Mixed Use was computed for both Land Use Mix (M_l) and Housing Mix (M_h), in accordance with Jacobs' doctrine that 'the district … must serve more than one primary function' (1961, p. 152). The former was computed using an entropy index based on primary land use categories extracted from government digital maps (HKLD, 2016). M_l for each TPU can be expressed in Formula Equation (2) below

$$M_l = -\sum_{i=1}^{n} \frac{P_{ij} * \ln P_{ij}}{\ln(n)} [\text{for all j}] \tag{2}$$

where P_{ij} is the area of land use category i as a proportion of the TPU area j; n is the number of land use categories (n=7i.e., residential, commercial, institutional, industrial, open space, village and other). Similarly, Housing Mixture (M_h) can computed using the same formula, substituting land use categories with housing types (private, public, public ownership schemes, etc.).

- **Short Block:** The density of street intersections (B_{den}^{int}) for each TPU was used as a proxy for the average block size, in reference to literature (De Nadai et al., 2016a; Sung et al., 2015). B_{den}^{int} was calculated using the total number of street intersections divided by the area of the TPU.
- **Aged Building** was measured using the Mean Building Age ($A_{bldg\_age}^{mean}$) and the Standard Deviation of Building Age ($A_{bldg\_age}^{stdv}$) in reference to Jacobs that 'the district must mingle buildings that vary in age and condition, including a good proportion of old ones'. (1961, p.187). Information on building age were acquired from government sources (HKHS, 2016).
- **Density** was measured in accordance with Jacobs' description of 'a sufficiently dense concentration of people' (1961, p.200). Three indicators were computed: 1) dwelling density (X_{den}^{dwell}), calculated using the people count obtained from the 2016 Census (HKDCS, 2016) divided by area of TPU; (2) work density (X_{den}^{work}) calculated using the same census data and (3) building density (D_{den}^{bldg}) estimated by dividing the building count obtained from the HKLD database (2016) with the TPU area.
- **Tall Building** was blamed for casting 'shadows and misfortune of Rittenhouse Square in Philadelphia' (Jacobs, 1961, p.105), 'handicapped in supervision of children' (ibid, p.397)

and likened to ‘corridors in a bad dream’ (ibid, 1961, p.399). The mean building height ($H^{mean}_{bldg\_h}$) in a TPU was calculated from official GIS database (HKLD, 2016).

Performance indicators

We developed seven indicators to measure the performance of a city district, ranked by the times of mentions by Jacobs or in subsequent literature. The first three (Tweet Density, POI Density, Walking Commute) aimed to capture ‘vitality’, which was mentioned 31 times in Jacobs’ texts (1961) and subject to numerous follow-up studies, whilst indicator 4–7 (Tweet Sentiment, Employment Ratio, Mortality and Premature Mortality) were meant to capture ‘failure or success’.

- **Tweet Density** is an indicator for activities and social interactions. Strictly, it captures the effect of the sub-population group who have an active Twitter account. But since having a Twitter account is itself a measure of social interaction and vitality, it is reasonable to take this as a vitality measure, accepting that fewer Twitter users and lower Tweet rates are both indicators of lower vitality. We calculated Tweet Density by accounting all geo-coded Tweets by a TPU and divided by its footprint area. The Twitter database was acquired via data-mining using the Twitter API streaming, a technical tool allowing exhaustive recording of tweets for a defined geographical (Twitter Inc., 2016). The original dataset has been extensive cleaned; repetitive posts were removed; the top 1% of high frequency user IDs were manually re-viewed to remove fake accounts or commercial advertisement.
- **POI Density** is another accepted measure for urban activities. A POI is a point feature tagged by category, such as a restaurant, grocery store or other amenities which depend on and attract pedestrian activities. We computed POI Density by aggregating geo-coded POIs by TPU, normalized by the unit area. The POI datasets were retrieved from OpenStreetMap (2016), a digital platform enriched by millions of contributors.
- **Walking Commute** was used as a third indicator for vitality, in reference to the ‘sidewalk ballet’ in Jacobs’ text (1961, p.50) and follow up studies. Walking commute was calculated by counting the percentage of people reportedly walking to work, the input data were acquired from the official travel mode survey (HKDCS, 2016).
- **Tweet Sentiment** captures the sentiment tones at the city district level. ‘Happiness’, together with ‘happy’ and ‘sentiment’ appeared 16 times collectively in Jacobs’ text (1961). We counted geo-coded tweets with positive (happy) sentiment tones as a percentage of the total in each TPU. To control the influences of small samples, TPUs with too few tweets ($N < 10$) were excluded from analysis. Sentiment analysis was conducted using the Linguistic Inquiry and Word Count (LIWC) (Kahn et al., 2007) and Emoticon detection (Hogenboom et al., 2013) and results were cross-checked by human reviewers.
- **Employment Ratio** was used to capture ‘delinquency’, a term mentioned 11 times by Jacobs (1961) and it has been measured quantitatively in literature (Chau et al., 2014). Employment ratio measures the ratio of population over 18 participating in the labour force, both as employed or self-employed according to census data (HKDCS, 2016). This measures Jacobs’ interpretation of Alfred Marshall’s (1890) producer agglomeration economies: higher density lowers the cost to workers of finding jobs and lowers the cost to employer in finding workers. It can also relate to the economic vitality in the first category.
- **Mortality and Premature Mortality** measure the death rates of the population and its sub-groups, in reference to Weicher’s work (1973). ‘Death’ was mentioned 4 times Jacobs (1961), although more in a rhetorical sense, meaning ‘deprivation’. The death count was acquired from the municipal vital statistics (HKDH, 2019), aggregated for each TPU by address information, divided by the number of residents in the TPU, and normalized for every 1000

people. Premature deaths, defined locally as deaths before 75 years of age, were computed using the same method.

Statistical analysis and confounders

We deployed a cross-sectional research design, using regression models to determine the associations between the outcomes and inputs. To account for the spatial dependencies of the dependent variables, the Spatial Lag Model (SLM) was used. The X and Y coordinates of each TPU were extrapolated from the longitudinal and latitudinal coordinates of the centre point of its polygon. Confounding influences from access to destinations (e.g., transit, schools, parks or waterfront), demographic and socio-economic factors were modelled to test the independent contributions from Jacobs' specific urban conditions. The dependant variable P, representing either the 'vitality' and 'failure and success' indicators, can be expressed in Formula (3) below

$$\begin{aligned} P = \rho W_P + \beta_0 + \beta_1 M_l + \beta_2 M_h + \beta_3 B_{den}^{int} + \beta_4 A_{bldg\_age}^{mean} + \beta_5 A_{bldg\_age}^{stdv} + \beta_6 D_{den}^{dwell} + \beta_7 D_{den}^{work} \\ + \beta_8 D_{den}^{bldg} + \beta_9 H_{bldg\_h}^{mean} + \beta_{10} C_{topo}^{slope} + \beta_{11} C_{mtr}^{dis} + \beta_{12} C_{sch}^{dis} + \beta_{13} C_{park}^{dis} + \beta_{14} C_{w}^{dis} + \beta_{15} C_{job}^{ser} \\ + \beta_{16} C_{job}^{ind} + \beta_{17} C_{<24}^{pop} + \beta_{18} C_{>65}^{pop} + \beta_{19} C_{male}^{gen} + \mu \end{aligned} \tag{3}$$

where ρ is the spatial regression coefficient; W_P is the spatial lag operator; M_l, M_h, B_{den}^{int}, $A_{bldg\_age}^{mean}$, $A_{bldg\_age}^{stdv}$, D_{den}^{dwell}, D_{den}^{work}, D_{den}^{bldg} are the urban conditions specified previously; C_{topo}^{slope} is the topographical slope; C_{mtr}^{dis}, C_{sch}^{dis}, C_{park}^{dis}, C_{w}^{dis} are the distances to the nearest transit station, school, park and waterfront, calculated using the Euclidean distance to the nearest facilities from the closest point in the TPU (if the feature was contained in the TPU, then the distance is zero); C_{job}^{ser} is the percentage of service jobs, such as management, finance and professionals; C_{job}^{ind} is the percentage of industrial jobs, including manufacturing, transportation and construction; $C_{<24}^{pop}$, $C_{>65}^{pop}$ and C_{male}^{gen} are the percentages of population that are under 24, above 65 or male; $\beta_0 - \beta_{19}$ are the regression coefficients; μ is the unmodelled error term.

We checked the data distribution of each variable used in the regression against a normal distribution curve. Multicollinearity was checked and the Pearson's correlation values among all variables were below 0.8, a threshold considered as high degrees of collinearity. Regression models were fitted using STATA MP/16.1 software package.

Results and discussion

Data characteristics

We obtained 236621 tweets in HK between May 1 and October 31, 2016, each consisted of the text, time stamp and basic account profiles. 86,706 contained GPS coordinates and can be geo-coded with one of the 209 TPUs (Supplemental Figure S3). Meanwhile, we acquired 8972 POIs from OSM, divided into commercial services such as retails and restaurants (39.5%), public amenities such as playground, swimming pool (46.3%) and public services (7.8%). Supplemental Table S2, S3 and S4 summarized the mean, range and distribution of performance indicators, urban form conditions and confounders, respectively.

We found significant spatial dependencies in OLS regression models. The Moran's I, an indicator for spatial dependencies, was 3.457 (p = .001) for the Tweet Density Model, 3.250 (p = .001) for the Walking Commute Model, and 7.326 (p = .007) for Premature mortality model, all above the threshold value of 2.0, suggesting that OLS is no longer suitable. A further test of spatial correlations among the residuals using Lagrange multiplier sores suggested that the Spatial Lag Model (SLM) scored higher compared with the Spatial Error Model in most cases. The above results justified the choice of the former over the latter. The spatial diagnostics were performed using STATA MP/16.1 software package, details are included in Supplemental Table S8.

Testing of Jane Jacobs' theory

Overall, we found strong associations between urban conditions and 'vitality', support of Jacobs' theory. The same associations cannot be observed between urban conditions and 'failure and success' indicators. Table 1 presents the results for eight SLM regression models, three for 'vitality' and four for 'success and failure'. The three vitality models have high R-squared values between 0.74 and 0.82, suggesting Jacobs' urban conditions plus the confounders can explain the variance of Tweet Density, POI Density and Walking Commute reasonably well. In comparison, we found no association between urban conditions and Tweet Sentiment nor Mortality. The Employment Ratio was an exception, which associated negatively with Aged Building and Housing Mix with controls, contrary to what Jacobs had theorized. The associations between Jacobs' urban conditions and performance indicators were summarized below, ranked by the consistency of signs and statistical significance level of the parameter estimates.

- Our finding is a strong support for Jacobs' assertion that short block is conducive to vitality. B_{den}^{int} associated positively with Tweet Density ($p < .001$), POI Density ($p<0.001$), and Employment Ratio ($p < .01$). Supplemental Figure S4 shows a scatter plot of B_{den}^{int} and Tweet Density in logarithmic scale. Our finding is consistent with evidence from the six-Italian-city study (De Nadai et al., 2016a). Our short block variable (road intersection density) is capturing a population density effect, so that the dwelling density variable is insignificant with intersection density in the model. This is important for our result since it implies that 'short block' offers more explanation of tweet vitality than dwelling density alone. It is capturing population density and more. We hypothesise that 'the more' may be unmeasurable street-level vitality encouraged by short blocks. We also note that the relationship between short blocks and Tweet Density controls for two significant confounders: building height and building density. This is more evidence that street layout is independent of population density, which is captured by the two controls. Short blocks are also predictive of POI density, and this is also holding building density constant (with highly statistical significance). We regard this as strong support for Jacobs, in that for a given building density, POI-vitality increases with road intersection density.
- Density associated positively with vitality, in line with Jacobs' observations. We found that building density, instead of the dwelling density that Jacobs pointed to, associated positively with Tweet Density ($p < .05$), POI Density ($p < .001$), and Walking Commute ($p < .001$) with controls. No performance indicators, other than the Employment Ratio, associated with Work Density. Replacing the dwelling and work density with a combined indicator, the Ambient Population Density, yielded consistent results. Our findings differ from De Nadai et al., 2016a; Sung et al., 2015, which suggest that work density contributes positively to vitality. It is likely that our building density variable captured the effect from population or work density, so that the latter two becomes insignificant with vitality in the model. Also, work density may not feature in HK's vitality because a sizeable proportion of HK residents live in new towns and planned housing estates and non-service jobs tend to be highly concentrated in CBDs and industrial centres such as logistics and port hubs.
- Aged building associated positively with POI and walking vitality, not tweet vitality. $A_{bldg\_age}^{mean}$ associated positively with POI Density ($p < 0.01$) and Walking Commute ($p < 0.05$). The first of these is probably due to the location of shops and restaurants on the ground floor of older

Table 1. SLM analysis results of regressing performance indicators and Jacobs' urban form conditions.

Dependent Variables		'Vitality' Indicators						'Failure and Success' Indicators							
		Tweet Density (log10 N) (#/km^2)		POI Density (#/km2)		Walking Commute (%)		Tweet Sentiment (% positive)		Employment Ratio (%)		Premature Mortality Rate(<75)		Population Mortality Rate	
Independent Variables		Coef. (95% CI)	*p*-value	Coef. (95% CI)	p-value	Coef. (95% CI)	*p*-value	Coef. (95% CI)	*p*-value	Coef. (95% CI)	*p*-value	Coef. (95% CI)	*p*-value	Coef. (95% CI)	*p*-value
Mixed Use	Land Use Mix	2.717 (1.117,4.318)	0.001**	0.879 (−0.351, 2.110)	0.161	10.007 (2.290, 17.723)	0.011*	10.947 (−5.123, 27.016)	0.182	−1.810 (−3.566, −0.053)	0.043*	1.416 (−0.536, 3.367)	0.155	1.497 (−4.858, 7.851)	0.644
	Housing Mix	−0.135 (−1.024,0.755)	0.767	0.026 (−0.650, 0.702)	0.941	−0.838 (−5.173, 3.497)	0.705	0.723 (−8.349, 9.795)	0.876	−2.532 (−3.534, −1.530)	0.000***	0.553 (−0.547, 1.653)	0.324	−2.031 (−5.642, 1.580)	0.270
Short Block	Road Intersection Density (1,000 /km^2)	7.751 (5.262, 10.240)	0.000***	4.523 (2.560, 6.486)	0.000***	10.588(−1.491, 22.667)	0.086	9.764 (−14.774, 34.302)	0.435	3.507 (0.843, 6.171)	0.010**	−1.791 (−4.787, 1.205)	0.241	−0.365 (−10.041, 9.311)	0.941
Building Age	Mean (yr.)	0.030 (−0.001, 0.061)	0.056	0.033 (0.010, 0.057)	0.006**	0.181 (0.032, 0.330)	0.017*	−0.043 (−0.357, 0.270)	0.787	−0.042 (−0.077, −0.008)	0.016*	−0.016 (−0.053, 0.022)	0.421	−0.108 (−0.231, 0.015)	0.086
	Standard Deviation (yr.)	−0.025 (−0.081, 0.032)	0.395	−0.037 (−0.080, 0.007)	0.096	−0.068 (−0.344, 0.208)	0.627	−0.374 (−0.951, 0.204)	0.205	0.056 (−0.007, 0.120)	0.082	−0.007 (−0.077, 0.063)	0.850	−0.006 (−0.235, 0.223)	0.957
Density	Dwelling (1,000/km^2)	−0.035 (−0.112, 0.041)	0.367	−0.058 (−0.117, 0.001)	0.052	−0.326 (−0.702, 0.051)	0.09	−0.251 (−1.035, 0.533)	0.530	−0.191 (−0.280, −0.102)	0.000***	0.002 (−0.095, 0.099)	0.965	−0.002 (−0.316, 0.313)	0.991
	Work (1,000/km^2)	0.065 (−0.086, 0.217)	0.397	0.120(0.004, 0.235)	0.042*	0.550 (−0.190, 1.291)	0.145	0.429 (−1.115, 1.973)	0.586	0.352 (0.177, 0.527)	0.000***	0.001 (−0.191, 0.192)	0.992	0.024 (−0.596, 0.644)	0.940
	Bldg. (1,000/km^2)	0.433 (0.077, 0.788)	0.017*	0.597 (0.326, 0.868)	0.000***	3.9718 (2.226, 5.715)	0.000***	−0.716 (−4.352, 2.920)	0.699	0.157 (−0.245, 0.559)	0.444	−0.014 (−0.455, 0.427)	0.952	−0.410 (−1.848, 1.028)	0.576
Tall Bldg.	Mean Building Height (m)	0.024 (0.004, 0.044)	0.016*	0.009 (−0.006, 0.025)	0.226	0.134 (0.036, 0.232)	0.007**	0.170 (−0.037, 0.377)	0.107	0.000 (−0.023, 0.022)	0.979	−0.014 (−0.039, 0.011)	0.288	−0.070 (−0.151, 0.012)	0.094
Confounder	Topographical Slope (degree)	0.003 (−0.036, 0.042)	0.875	−0.015 (−0.045, 0.016)	0.342	−0.173 (−0.364, 0.017)	0.075	−0.005 (−0.405, 0.394)	0.979	0.063 (0.019, 0.108)	0.006***	0.015 (−0.034, 0.063)	0.553	0.068 (−0.090, 0.225)	0.399
	Distance to MTR Station (km)	−0.001 (−0.310, 0.309)	0.996	−0.111 (−0.347, 0.124)	0.354	−1.675 (−3.187, −0.164)	0.03*	0.002 (−0.001, 0.005)	0.198	0.458 (0.106, 0.810)	0.011*	0.052 (−0.331, 0.435)	0.790	0.895 (−0.351, 2.142)	0.159
	Distance to School	0.945 (−2.622, 4.512)	0.604	−1.290 (−4.009, 1.430)	0.353	0.243 (−17.187, 17.673)	0.978	0.021 (−0.015, 0.058)	0.248	2.177 (−1.775, 6.130)	0.280	0.070 (−4.371, 4.511)	0.975	−2.882 (−17.285, 11.521)	0.695
	Distance to Park	−4.212 (−11.710, 3.286)	0.271	−2.997 (−8.763, 2.768)	0.308	−6.295 (−42.802, 30.212)	0.735	0.062 (−0.013, 0.138)	0.107	−8.042 (−16.639, 0.556)	0.067	−8.412 (−17.631, 0.807)	0.074	−32.991 (−62.991, −2.992)	0.031
	Distance to Waterfront	−0.117 (−0.267, 0.033)	0.127	−0.120 (−0.235, −0.006)	0.040*	−0.431 (−1.156, 0.294)	0.244	0.808 (−0.001, 0.002)	0.301	−0.045 (−0.212, 0.122)	0.598	0.119 (−0.069, 0.307)	0.213	0.577 (−0.030, 1.185)	0.062
	Job Type 1 (%)	−0.048 (−0.111, 14.194)	0.130	−0.046 (−0.093, 0.002)	0.060	0.092 (−0.214, 0.398)	0.554	−0.094 (−0.733, 0.545)	0.773	0.397 (0.327, 0.467)	0.000***	−0.010 (−0.088, 0.068)	0.794	0.031 (−0.224, 0.286)	0.814
	Job Type 2 (%)	−0.113 (−0.187, −0.039)	0.003**	−0.086 (−0.143, −0.029)	0.003**	0.138 (−0.223, 0.498)	0.454	−0.005 (−0.752, 0.743)	0.990	0.323 (0.241, 0.405)	0.000***	0.068 (−0.023, 0.160)	0.144	0.197 (−0.100, 0.494)	0.194
	% 24	−0.062 (−0.178, 0.055)	0.298	−0.037 (−0.126, 0.051)	0.411	0.005 (−0.564, 0.575)	0.985	−0.377 (−1.568, 0.814)	0.535	0.154 (0.024, 0.284)	0.020*	−0.006 (−0.152, 0.139)	0.932	0.173 (−0.304, 0.651)	0.477
	% ≥ 65	0.007 (−0.093, 0.107)	0.892	0.014 (−0.061, 0.090)	0.709	0.487 (0.008, 0.967)	0.046*	0.102 (−0.900, 1.103)	0.842	−0.251 (−0.361, −0.140)	0.000***	0.068 (−0.053, 0.190)	0.269	0.668 (0.265, 1.070)	0.001
	% Male	0.083 (−0.036, 0.202)	0.172	0.123 (0.032, 0.214)	0.008**	0.939 (0.362, 1.515)	0.001**	−0.135 (−1.339, 1.070)	0.826	−0.550 (−0.684, −0.416)	0.000***	0.047 (−0.099, 0.194)	0.526	−0.065 (−0.542, 0.411)	0.788
Cons.		1.528 (−6.296, 9.352)	0.702	1.528 (-5.455, 7.374)	0.769	−72.517(−108.435, −36.599)	0.000***	118.300(−10.016, 246.615)	0.071	60.909 (42.485, 79.334)	0.000***	2.662 (−6.284, 11.608)	0.560	16.251 (−13.123, 45.625)	0.278
Spatial Regression Coefficient (ρ)		0.736 (0.377, 1.096)	0.000***	0.231 (−0.341, 0.803)	0.428	0.711(0.253, 1.170)	0.002**	−0.332 (−1.641, 0.978)	0.619	−0.187	0.221	−3.037 (−4.400, −1.674)	0.000***	−4.084 (−5.503, −2.666)	0.000***
R2		0.823		0.804		0.737		0.165		0.962		0.455		0.508	

*$p < .05$, ** $p < .01$, *** $p < .001$

residential buildings in HK, where access to customers and rents are favourable. This indeed is capturing a Jacobs effect. The association we found between aged buildings and, respectively, median household income, employment ratio and walking commute aligned with Jacobs' observation that aged buildings contribute to housing affordability. HK's older neighbourhoods tend to have POI-vitality because of liberal laws for co-locating people and local commercial uses. The high densities of people and buildings creating a Jacobs-style vibrant neighbourhood. The lack of significance of building age in the Italian study cited (De Nadai et al., 2016a), was said to be due to the prevalence of old buildings in Italian cities. HK offers a better test in this sense, and we corroborate the role played by older neighbourhoods in providing lower rent premises for the kind of people-facing businesses that make for urban vibrancy.

- We found mixed evidence on mixed use in association with 'vitality'. M_l associated positively with Tweet Density ($p < .01$) and Walking Commute ($p < .05$), but not with POI Density. The tweet and walking vitality results are interesting since they control for building density, tall buildings and short block morphology. This suggests that M_l might be an independent influence on twitter vitality and walking. Perhaps tweeting density and pedestrian activities are partially a function of stimulus: people tweet about something they experience, and mixed use provides a denser and more heterogeneous set of potential experiences. A previous study conducted by Sung et al. (2015) also found no statistically significant associations between M_l and walking activities; neither was the association significant between M_l and mobile phone activities, according to De Nadai et al., (2016a). A plausible reason is perhaps M_l is no longer a reliable estimate of the diversity of city functions, given the vertical mixture of functions is common in dense cities. We found no statistically significant association between Housing Mix (M_h) and vitality indicators, except that M_h is negatively associated with Employment Ratio ($p < .001$), suggesting that lower employment can explain the presence of public housing.
- Tall building was found to have played a positive role in city vitality, contrary to Jacobs' observation. The mean building height associated positively with Tweet Density ($p < .05$), and Walking Commute ($p < .01$) with controls. It also associated positively with POI Density and Tweet Sentiment, although not at a statistically significant level. The finding echoes with previous housing research literature from HK (Yeung and Wong, 2003) and Singapore (Yuen et al., 2006), which credit high-rise public housing programmes as a success in both cities, with the majority of residents living in high-rise housing apartments being satisfied. There are certainly few if any 'corridors in a bad dream' nor living spaces 'handicapped in supervision of children' of Jane Jacobs' time. Rather, our evidence in HK asserts that high-rise housing, if properly designed and managed, can offer a viable solution to house the population.

Discussion

We place findings from this study in the context of the existing literature in Table 2, which reveal that Jacobs' urban conditions are linked to 'vitality' indicators, but not to "failure and success" ones. We found consistent support for Jacobs' theory in explaining city vitality measured in Twitter Activities, POI density (this study), mobile phone activities (De Nadai et al., 2016a), walking behaviour (Sung et al., 2015 and this study). However, there are no evidence of extending Jacobs' observations to explain "Failure and Success' of city districts measured in Mortality, Crime or Employment. Neither this study nor Weicher (1973) found statistically significant associations between Jacobs' urban conditions and death rate; neither Fowler (1987) nor Schmidt (1977) found statistically significant associations between crime incidents and urban conditions in statistically significant ways, although aspects of the people dynamics of Jacobs' theory, i.e. 'eyes on street' measured by population diversity and the ratio of visitors, have been confirmed to have associated with the probably of crime hotspots occurring in London (Traunmueller et al., 2014).

The availability of new urban data provides opportunities to evaluate and further develop classic planning theories. As Table 2 suggests, studies using new data source, i.e., this study and those of De Nadai et al., 2016a; Sung et al., 2015, are statistically more powerful and more likely to yield observations in consistency with each other. They are consistent in the sign and statistical significance of the regression coefficients for Mixed Use, Short Block, Aged Building and Density, despite the three studies are from three countries spanning two continents. In comparison, earlier studies using survey (Fowler, 1987) or reported crime (Schmidt, 1977) arrived at contradictory conclusions with each other. This inconsistency may have more to do with the choice of performance indicators than geographical differences. The new urban data can dynamically measure a substantial proportion of the population, equipping contemporary studies with models of higher statistical power compared with earlier studies. An example is the use of the density of geo-coded tweets, which is accessible from the public domain. It is a good supplement to data from Point-of-Interest, travel behaviour or mobile phone activities.

The strength of this study is multi-dimensional measurement of urban performances using independent data sources. Social media data, for example can be 'a' measure of urban vitality not necessarily the 'best measure', since several studies pointed to the underrepresentation of the elderlies, minorities or the low-income groups by platforms such as Twitter and Foursquare (Anselin and Williams, 2016). Similarly, the POI data from voluntary contributors can be more developed in wealthier neighbourhoods, i.e., HK's urban core, and less so in rural areas (Haklay et al., 2010). Whilst a single indicator may be flawed, a combination of them is less likely to distort the picture. A multi-dimensional measurement is a key step towards refined theory and good science.

Inevitably, our study is reductionist in nature, and the mathematical expression cannot capture the richness of Jacobs' work in its entirety. The performance indicators developed in this study cannot exhaustively capture the complex human-city interactions. Rather, they are the closest approximation to her main constructs in our comprehension. There are others we were unable to include, given the practical limits of an empirical paper. An example is 'crime', mentioned 22 times by Jacobs (1961). It was not included in our study since crime data at district level were unavailable; its impact on the results is expected to be small, given the low crime rates in HK.

Findings from this study can help identify factors that are positively or negatively related to performance outcomes of HK's city districts. This is timely for the city's planners, who placed vitality, diversity and walkability on their priority list, and they should encourage urban form conditions such as short block, building density, tall and aged buildings. Such findings may not be automatically extended elsewhere, since Jane Jacobs' intention was not to provide a one-size-fit-all formulation, but to call attention to observed issues within specific geographical and temporal domains.

We postface our study with a cautionary note. The quest behind our paper and the others we cite is important but contrived. Important because Jacobs' theories have turned into long-lived doctrine that still influence the way cities are planned. Contrived, because they are not really theories. They were observations at a particular time and place.

To make progress in the theory of urban design, it may well be important to do what we have done in this paper and present simple statistical demonstrations that Jacobs' four conditions are indeed positively associated with vitality measured with urban big data. But the vitality measures we have used are only intermediate measures, and can only be so, because vitality is an unmeasurable concept. By contrast, the 'life and death' measures we used, measure impact in a more direct way. In that sense, the two types of tests we have applied are not equal. To thoroughly test the impact of urban design on economic success or failure, or happiness or health requires a different approach – one involving large number panel studies or natural experiments, with well-specified models and

Table 2. A Comparison of regression results from six evaluation studies including this one. We reported only the sign of beta coefficients due to the difference in adopted unit system among studies.

		Vitality & activities					Failure & success			
Performance indicators		Twitter activities	Mobile phone activity	Walking commute	Walking activities	POI density	Mortality rate	Standardized death rate	Perceived crime	Reported crime
Description	Study (location)	This study (Hong Kong)	De Nadai et al., 2016a (Italy)	This study (Hong Kong)	Sung et al., 2015 (Seoul)	This study (Hong Kong)	This study (Hong Kong)	Weicher, 1973 (chicago)	Fowler, 1987 (toronto)	Schmidt, 1977 (denver)
Mixed use	Land use mix	+**	*n.a*	+*	*n.a*	+	+	–	–	+
	Housing mix	–	*n.a*	–	+***	+	*n.a*	*n.a*	*n.a*	*n.a*
Short block	Road intersection density (count/km^2)	+***	+***	+*	+***	**+*** **	–	–	–	+
Aged bldg	Mean bldg. Age (yr.)	+	*n.a*	+***	+***	**+** **	+	+	+	+
Density	Dwelling (1000/km^2)	–	*n.a*	–	*n.a*	–	+	–**	–	+
	Work (1000/km^2)	+	+***	+	+***	+*	–	+	n.a	n.a
	Building (1000/km^2)	+*	n.a	+***	n.a	**+*** **	n.a	n.a	n.a	n.a
Others	Tall bldg. (m)	**+* **	n.a	**+** **	n.a	**+**	–	n.a	n.a	n.a
	Border vacuum	n.a	–***	n.a	–***	n.a	n.a	n.a	n.a	n.a
R-squared		0.82	0.77	0.74	n.a	0.80	0.51	0.53	0.62	0.51

*p < .05, ** p < .01, *** p < .001

accurate data. Sarkar et al. (2017), for example showed that obesity (adiposity) in a sample of over 400,000 British members of the UK Biobank panel, is significantly correlated with dwelling density, controlling for socio-demographics, prior health conditions and multiple built environment features around their homes. The pattern followed an inverted hockey-stick curve, with odds of obesity falling steadily with density from the average British suburb as you move towards the city centre. This also tests Jacobs' idea, but with a finer lens. Living at higher density produces more vibrancy as people walk more in response to a higher density of walkable destinations. From an urban designer's perspective, denser places are more vital; from an individual and population health perspective, the by-product of street 'vitality' is better health. Jacobs may not have guessed the link with obesity, or with the other wellness outcomes now associated with higher density living. But now we can measure such fine-grained direct impacts of urban morphology. Just as we can measure conceptual ideas such as vitality more reliably (same results with repeated measures) using social media and other urban big data. And so, as expressed by authors cited (De Nadai et al., 2016a; Fowler, 1987; Schmidt, 1977; Sung et al., 2015; Weicher, 1973), we should move on from Jacobs and take our performance-related design hypotheses from what has now become a much more mature scientific field. The teaching of Jane Jacobs' doctrine should include its evaluation by our study and the five others we cite, either as an addendum or a footnote to the original text. Such endeavours can strengthen the scientific rigor of urban design discipline and fulfil Jacobs' own wish to 'test and further develop the theories' (Jacobs, 1961, p. 22). She should remain an inspiration, shedding light on the search for new theories based on a mixture of good science and refined professional judgement.

Conclusion

The study examined Jacobs' doctrine of whether urban form conditions contribute to 'vitality' and 'failure and success' of city districts. We measured 'vitality' using Twitter activity, POI density, walking activities, and 'Failure and success' using tweet sentiment, mortality and employment ratio. Findings lend support for the most aspects of Jacobs' theory in HK, six decades after the publication of the original text. Her four primary conditions were positively related to vitality measured in social media activities, walking and POI density, except that one needs to substitute dwelling density for building density. No consistent associations were found between Jacobs' urban conditions and measures of 'failure and success' of city districts, such as mortality, employment ratio. Contrary to Jacobs' observation that tall buildings contribute negatively to urban life, we found them to be positively related to vitality in HK. Whilst Jacobs should remain as an inspiration, future researchers and practitioner should move forward in search of new theories that are sensitive to geographical, political, environmental and social contexts. New urban data, as exemplified in this paper, can serve as a powerful tool for this purpose.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Natural Science Foundation of China (51978594).

ORCID iDs

Jianxiang Huang https://orcid.org/0000-0002-0027-3944
Yi Lu https://orcid.org/0000-0001-7614-6661

Supplemental Material

Supplemental material for this article is available online.

References

Anselin L and Williams S (2016) Digital neighborhoods. *Journal of Urbanism* 9(4): 305–328. DOI: 10.1080/17549175.2015.1080752.

Bertrand K and Bialik M, Vidree K, et al. (2013) *Sentiment in New York City: A High Resolution Spatial and Temporal View*, NECSI report. New York: Arxiv. Available at: http://arxiv.org/abs/1308.5010

Booth C (1895) Political science quarterly. In *Life and Labour of the People in London*. London: McMillan and Co, Vol.5, p. 727.

Bouyer J, Inard C and Musy M (2011) Microclimatic coupling as a solution to improve building energy simulation in an urban context. *Energy and Buildings* 43(7): 1549–1559. DOI: 10.1016/j.enbuild.2011.02.010.

Chau PH, Gusmano MK, Cheng JOY, et al. (2014) Social vulnerability index for the older people—hong kong and New York city as examples. *Journal of Urban Health* 91: 1048–1064. DOI: 10.1007/s11524-014-9901-8

De Nadai M, Staiano J, Larcher R, et al. (2016a) The death and life of great italian cities: a mobile phone data perspective. In: Proceedings of 26th International ACM Conference on World Wide Web. Montreal, Canada, 2016

Ewing R and Cervero R (2010) Travel and the built environment: a meta-analysis. *Journal of the American Planning Association* 76(3): 265–294.

Flyvbjerg B (2006) Five misunderstandings about case-study research. *Qualitative Inquiry* 12(2): 390–404. DOI: 10.1177/1077800405284363

Fowler EP (1987) Street Management and City Design. *Social Forces* 66(2): 365–389. DOI: 10.1093/sf/66.2.365

Glaeser E (2010) *Preservation Follies: Excessive Land Marking Threatens to Make Manhattan a Refuge for the Rich*. New York: City Journal. Available at: https://www.city-journal.org/html/preservation-follies-13279.html

Glaeser E (2011) *How Skyscrapers Can Save the City*. Washington, DC: The Atlantic. Available at: http://www.theatlantic.com/magazine/archive/2011/03/how-skyscrapers-can-save-the-city/308387/(accessed 20 June 2016).

Go-Globe (2015) *Social Media Use in Hong Kong - Statics and Trends*, Hong Kong: Go-Globe. Available at: http://www.go-globe.hk/blog/social-media-hong-kong/(accessed 10 September 2016).

Haklay M, Antoniou S, Basiouka S, et al. (2010) How many volunteers does it take to map an area well? the validity of Linus' law to volunteered geographic information. *Cartographic Journal* 47(4): 315–322. DOI: 10.1179/000870410X12911304958827

Harris R (2011) The Magpie and the Bee: Jane Jacobs's magnificent obsession. In: Page M and Mennel T (eds), *Reconsidering Jane Jacobs*. Chicago, IL: Planners' Press.

Hill DR (1988) Jane jacobs' ideas on big, diverse cities: a review and commentary. *Journal of the American Planning Association* 54(3): 302–314. DOI: 10.1080/01944368808976491.

HKDCS (2016) *Department of Census and Stastistics: District Profile of 2011 Population Census*. Hong Kong: HKDCS.Available at: http://www.census2011.gov.hk/en/major-housing-estates.html (accessed 26 June 2016).

HKDH (2019) *Department of Health Vital Statistics*. Sacramento, CA: CHP. Available at: https://www.chp.gov.hk/en/static/24025.html (accessed 1 April 2020).

HKHS (2016) *Hong Kong Housing Society: Subsidised Housing*. Hong Kong: HKHS. Available at: http://www.hkhs.com/eng/business/renhousing.asp (accessed 14 September 2016).

HKLD (2016) *Hong Kong Lands Department Digital Maps*. Hong Kong: HKLD. Available at: http://www.landsd.gov.hk/mapping/en/digital_map/intro.htm (accessed 15 September 2016).

Hogenboom A, Bal D, Frasincar F, et al. (2013) Exploiting emoticons in sentiment analysis. In: Proceedings of the 28th Annual ACM Symposium on Applied Computing - SAC '13. Amsterdam, Netherlands, 18 March 2013

Huang J, Kamrowska-Zaluska H, Kamrowska-Zaluska D, et al. (2021) The image of the city on social media: a comparative study using "big data" and "small data" methods in the tri-city region in Poland. *Landscape and Urban Planning* 206: 103977.

Huang J, Chang Y and Chang H, (2022) A city is not a tree: a multi-city study on street network and urban life. *Landscape and Urban Planning* 226, p. 104469.

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

Kahn JH, Tobin RM, Massey AE, et al. (2007) Measuring emotional expression with the linguistic inquiry and word count. *The American Journal of Psychology* 120(2): 263–286.

Kang C, Fan D and Jiao H (2020) Validating activity, time, and space diversity as essential components of urban vitality. *Environment and Planning B: Urban Analytics and City Science* 48: 1180–1197. DOI: 10.1177/2399808320919771

Katz P (1993) *The New Urbanism: Towards an Architecture of Community*. New York: McGraw-Hill Education.

Kotkin J (2000) *The New Geography: How the Digital Evolution is Reshaping the American Landscape*. New York: Random House.

Long Y and Huang C (2017) Does block size matter? impact of urban design on economic vitality for chinese cities', *Environment and Planning B: Urban Analytics and City Science*, 46, pp. 406–422. DOI: 10.1177/2399808317715640

Lu Y (2019) Using Google Street View to investigate the association between street greenery and physical activity. *Landscape and Urban Planning* 191: 103435. DOI: 10.1016/j.landurbplan.2018.08.029

Marshall A (1890) *Principles of Economics*. London: Macmillan.

Marshall S (2012) Science, pseudo-science and urban design', In: *Urban Design International*. Berlin: Nature Publishing Group 17(4): 257–271. DOI: 10.1057/udi.2012.22

Montgomery C (2013) *Happy City: Transforming Our Lives Through Urban Design*. New York: Farrar, Straus and Giroux.

Newman O (1975). Reactions to the "Defensible Space" study and some further findings. *International Journal of Mental Health* 4(3): 48–70. Available at: http://www.jstor.org/stable/41344045

Ng M (2000) Urban regeneration in Hong Kong: problems and prospects. In: Yim C (ed), *Asian Planning Schools Association Papers: Prospect and Retrospect of Planning in Asia at the Turn of the Century.* Seoul, Korea: Seoul National University, pp. 169–180.

Oda M, Wen ML, Higurashi M, et al. (1989) Effects of high-rise living on physical and mental development of children. Journal of Human Ergology 18(2): 231–235.

OSM (2016) *Data Retieved from Open Street Map*. Norway: OSM

Quercia D, Hare NO and Cramer H (2014) Aesthetic capital : what makes london look beautiful, quiet, and happy? Proceedings of the 17th ACM Conference on Computer Supported Cooperative work & Social Computing, 15 February 2014.

Sarkar C, Webster C and Gallacher J (2017) Association between adiposity outcomes and residential density: a full-data, cross-sectional analysis of 419 562 UK Biobank adult participants. *The Lancet Planetary Health* 1(7): e277–e288. DOI: 10.1016/S2542-5196(17)30119-5

Schmidt CG (1977) Influence of land use diversity upon neighborhood success: An analysis of Jacobs' theory. *The Annals of Regional Science* 11(1): 53–65. DOI: 10.1007/BF01287248

Sung H, Lee S and Cheon S (2015) Operationalizing jane jacobs's urban design theory: empirical verification from the great city of seoul, korea. *Journal of Planning Education and Research* 35(2): 117–130. DOI: 10.1177/0739456X14568021

Traunmueller M, Quattrone G and Capra L (2014) Mining mobile phone data to investigate urban crime theories at scale. In: *Lecture Notes in Computer Science (including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* UK: Publons

Trenkner T (2011) *Is It Time to Retire Jane Jacobs' Vision of the City? Triumph of the City author Edward Glaeser challenges long-accepted views on urban planning, promoting a more market-friendly set of policies*, Folsom, CA: Governing:The Future of States and Localities. Available at: https://www.governing.com/is-it-time-to-retire-jane-jacobs-vision-city.html (accessed 31 December 2019).

Twitter Inc (2016) *Twitter API*. San Francisco, CA: Twitter Inc. Available at: https://dev.twitter.com/streaming/overview (accessed 15 May 2016).

Webster CJ (2015) Refutation and the scientific knowledge base of urban planning practice, In: Silva E, Healey P, Harris P, et al. (eds), *The Routledge Handbook of Planning Research Methods*. England, UK: Routledge.

Weicher JC (1973) A test of jane jacobs' theory of successful neighborhoods. *Journal of Regional Science* 13(1): 29–40. DOI: 10.1111/j.1467-9787.1973.tb00375.x

Xia Chang, Yeh Anthony Gar-On and Zhang Anqi (2020) Analyzing spatial relationships between urban land use intensity and urban vitality at street block level: A case study of five Chinese megacities. *Landscape and Urban Planning* 193: 103669. DOI: 10.1016/j.landurbplan.2019.103669

Ye Y, Li D and Liu X (2018) How block density and typology affect urban vitality: an exploratory analysis in Shenzhen, China. *Urban Geography* 39(4): 631–652. DOI: 10.1080/02723638.2017.1381536

Yeung YM and Wong TKY (2003) *Fifty Years of Public Housing in Hong Kong: A Golden Jublee Review and Appraisal*. Hong Kong: The Chinese University Press.

Yue Y, Yeh Y, Zhang Y, et al. (2017). Measurements of POI-based mixed use and their relationships with neighbourhood vibrancy. *International Journal of Geographical Information Science* 31(4): 658–675. DOI: 10.1080/13658816.2016.1220561

Yuen B, Yeh A, Appold A, et al. (2006) High-rise living in singapore public housing. *Urban Studies* 43(3): 583–600. DOI: 10.1080/00420980500533133

Author Biographies

Dr. Jianxiang Huang is an assistant professor in the Department of Urban Planning and Design at the University of Hong Kong. He studies the human-built environment interactions through the lens of "urban new data", i.e., social media, POI, behaviors, and wellbeing. He also studies sustainable and intelligent design, including urban environment and energy modelling and design optimization based on artificial intelligence. He has developed a series of design and planning software tools for urban heat island mitigation, building energy efficiency, traffic noise mitigation, and enhance of occupant thermal comfort, which are applied in real-world projects.

Ms. Cui Yuming is a Ph.D. student at the University of Hong Kong. Her research direction is urban data analysis, especially the spatial network analysis and geographic information system. She has previously obtained a Master of Philosophy from the University of Hong Kong, a Master of Arts in Urban Design from Cardiff University, and a Bachelor of Architecture from Shenyang Architectural University.

Dr. Mengdi Guo is a young associate professor in the Faculty of Architecture at Tianjin University. Her research focuses on noise assessment in high density cities via measurement, simulation, and big data analytics, in support of the optimization of early-stage urban design and healthy city studies. She teaches bachelor architecture design studio. She received a PhD from the University of Hong Kong, a European Post-graduate Master in Urbanism from Delft University, and a Bachelor of Architecture from Tsinghua University.

Dr. Hung Chak Ho (Derrick Ho) is a Research Assistant Professor in the Department of Urban Planning and Design at the University of Hong Kong. His research focuses on urban resilience and environmental hazards, with particular interests in spatiotemporal modelling and machine learning applications for environmental exposure and health assessments. He co-developed the government-based Heat Exposure Integrated Deprivation Index for disaster risk planning and public health surveillance. Derrick received his PhD from Simon Fraser University and was trained as a research student in the British Columbia Centre for Disease Control.

Dr. Lishuai Li is an assistant professor at the Faculty of Aerospace Engineering at the Delft University of Technology (TU Delft). She is a member of the Air Transport and Operations (ATO) section of the Department of Control and Operations. She received a M.Sc. and Ph.D. degrees in Air Transportation Systems from the Department of Aeronautics and Astronautics at Massachusetts Institute of Technology (MIT) in 2009 and 2013, respectively. Prior to MIT, she obtained the B.Eng. degree in Aircraft Design and Engineering from Fudan University in 2007.

Dr. Yi Lu is an associate professor at City University of Hong Kong. He is interested in the behavioral impact of built environment, such as how site design of an urban smart development project influence walking, how nursing unit design influence staff behavior. He has published in leading journals such as Environment and Behavior, Environment and Planning B. He has also developed rigorous spatial description of physical environment and related assessment tools. He received his bachelor's degree of architecture from Southeast University in Nanjing China, his master's degree from National University of Singapore, and his PhD from Georgia Institute of Technology, USA. He practiced in architectural firms in Singapore and China for several years.

Prof. Chris Webster is a leading urban theorist and spatial economic modeller. He has published over 150 scholarly papers on the idea of spontaneous urban order and received over US$20M grants for research and teaching and learning projects. He was co-editor of Environment and Planning B for ten years. He has five prize-winning academic papers on urban theory. His present professional mission is to change the way cities are planned in China and his current research agenda is to establish systematic evidence for the relationship between urban configuration (planned and spontaneous) and individual health. He has degrees in urban planning, computer science, economics and economic geography.