



Research Paper

Relationships between urban vegetation and academic achievement vary with social and environmental context

Cody B. Hodson, Heather A. Sander^{*}*Department of Geographical and Sustainability Sciences, University of Iowa, USA*

HIGHLIGHTS

- Graduation rates were lower in high-intensity, low canopy settings overall.
- Rates were higher in low-intensity settings with high canopy or agricultural cover.
- Relationship between graduation and vegetation varied with SES and environment.
- Urban greening cannot be uniformly applied to support academic success.

ARTICLE INFO

Keywords:

Academic performance
Urban environments
Urban vegetation
Student environments
Socioeconomic status

ABSTRACT

Urban greening is increasingly suggested as a means of supporting the health and well-being of urban populations. Given positive relationships between vegetation and academic performance, urban greening could support scholastic achievement, enhancing future success in life. These relationships may vary among social and environmental contexts, however, suggesting greening benefits populations unequally. We investigated relationships between vegetation and academic achievement as indicated by high school graduation rates across social and environmental contexts in the continental US, assessing this variation and the potential for urban vegetation to support academic attainment. We categorized 1,333 public high schools based on the socioeconomic and environmental attributes of their attendance areas using k-means clustering. We identified variation in relationships between vegetation and graduation rate based on socioeconomic status (SES) and estimated conditional mean rates by environmental category using multilevel beta regression. We found significant variation in the relationship between school graduation rate and vegetation among vegetation types and socioeconomic contexts. Graduation rates were lower in high-intensity (i.e., more extensively built), low tree-cover settings and higher in lower-intensity (i.e., less extensively built), high tree-cover or agricultural settings. Agricultural vegetation associated positively with graduation rate broadly, while non-forest vegetation exhibited negative relationships for low-SES, majority Black schools. Tree canopy exhibited positive relationships that were stronger for high-SES schools and low-SES, majority Latino/a schools. These results highlight the importance of social and environmental context in mediating relationships between vegetation and academic achievement and the need to consider these disparities in supporting academic success through urban vegetation management.

1. Introduction

Over half of the global population is urban; nearly 70% will be urban by 2050 (UN Population Division, 2018). Cities ostensibly offer residents high quality-of-life, but may reduce well-being by decreasing exposure to nature, denying urbanites benefits of human-nature interaction (Soga & Gaston, 2016). Exposure may be especially low among poorer, minority populations (Boone et al., 2009; Heynen et al., 2006) that may

experience particularly reduced well-being. Urban vegetation and green space provide benefits to academic performance (Benfield et al., 2015; Donovan et al., 2018; Hodson & Sander, 2017; Kuo et al., 2018; Kweon et al., 2017; Leung et al., 2019; Li et al., 2019; Matsuoka, 2010; Sivarajah et al., 2018; Tallis et al., 2018; Wu et al., 2014). Given links between success in school and success in life (Ceci and Williams, 1997), vegetation and green space represent key resources for supporting urban well-being.

^{*} Corresponding author at: Department of Geographical and Sustainability Sciences, 316 Jessup Hall, University of Iowa, Iowa City, IA 52242, USA
E-mail addresses: cody-hodson@uiowa.edu (C.B. Hodson), heather-a-sander@uiowa.edu (H.A. Sander).

Research supports the role of urban vegetation in supporting the academic achievement of individuals (Benfield et al., 2015; Dadvand et al., 2015; Donovan et al., 2018). Stress Recovery Theory (SRT) and Attention Restoration Theory (ART) identify potential pathways whereby these benefits arise. SRT suggests that exposure to non-threatening natural environments improves stress recovery (Ulrich et al., 1991). Vegetated urban areas may thus reduce stress and enhance academic performance given negative relationships between chronic stress and performance (Schraml et al., 2012). ART suggests that spending time in nature improves recovery from mental fatigue (Kaplan, 1995). Vegetated urban spaces facilitate such recovery, enhancing concentration and task performance (Schutte et al., 2015). These theories support urban greening as a means for enhancing individual performance in school and adult life.

Urban vegetation could benefit well-being more broadly if these effects aggregate to populations. Recent studies indicate positive associations between vegetation and average standardized exam scores (Hodson & Sander, 2017; Kuo et al., 2018; Kweon et al., 2017; Leung et al., 2019; Li et al., 2019; Sivarajah et al., 2018; Tallis et al., 2018; Wu et al., 2014) and graduation rates (Matsuoka, 2010), suggesting benefits accrue to school populations. However, findings include negative (Beere & Kingham, 2017; Browning et al., 2018) and non-significant (Hodson & Sander, 2019; Markevych et al., 2019) associations. Identifying sources for these inconsistencies, which may reflect differences in study methodology or focus or socioeconomic conditions associated with both vegetation and academic achievement (Beere & Kingham, 2017), is critical to creating robust knowledge to support urban environmental design to enhance academic achievement. Identifying whether real differences exist in relationships between vegetation and achievement among environmental and social contexts is particularly important. For example, studies have identified negative relationships between campus tree canopy and sixth-grade mathematics achievement (Sivarajah et al., 2018) and between school attendance area (SAA) tree canopy and graduation rate (Hodson & Sander, 2019) for socioeconomically disadvantaged populations and variation in relationships between canopy and graduation rates with SAA impervious cover (Hodson & Sander, 2019). Better understanding of social and environmental influences on academic benefits of urban vegetation is thus required to support greening efforts that enhance student achievement.

We therefore assessed relationships between SAA vegetation and graduation rate, variation in these relationships among social contexts, and variation in graduation rate among environmental contexts using 1,333 US secondary schools. We focused on graduation rate given relationships between high school graduation and life outcomes (e.g., employment, income) and racial and ethnic gaps in graduation rates (US Department of Education, 2018) that urban greening might help close. We hypothesized positive relationships between graduation rate and vegetation and particularly strong associations with tree cover. We further hypothesized stronger, positive relationships for wealthier schools; weaker or negative relationships for high minority, socioeconomically disadvantaged schools; and higher expected graduation rates for schools in low-intensity, highly vegetated SAA. Low intensity in the context of this study refers to SAA with little built impervious (non-absorbent) cover such as roads, parking lots, and rooftops, moderate intensity refers to those SAA with a moderate degree of built impervious cover, and high intensity refers to SAA covered predominantly by built impervious surfaces. We identified key variation in relationships between vegetation and academic achievement with social context and between environmental contexts, suggesting the need to consider context in urban greening to equitably encourage student success and support the well-being of urban populations.

2. Methods

2.1. Sample selection

We initially selected all public high schools across the continental US with available SAA boundary data operating in the 2009/10 academic year. SAA encompass spaces where students likely spend most of their time, making them useful for representing student environments. We downloaded SAA data from the School Attendance Boundary Information System (SABINS) (College of William and Mary and the Minnesota Population Center., 2011). We investigated only urban schools with SAA centroids within 2010 US Census urban extents (US Census Bureau, 2010). Private, charter, or magnet schools that serve students from an undefined area and schools with missing data were excluded. The final sample selection included 1,333 schools in 36 states.

2.2. Academic achievement measure

We used 2010/11 high school graduation rates from the US Department of Education's EduFacts database (US Department of Education, n.d.) to indicate school-level academic achievement. This was the earliest academic year for which the US Department of Education had such data available. Although we used 2009/10 SAA boundary data, boundaries likely changed little between school years.

2.3. Sociodemographic and school-related variables

We derived several variables to account for socioeconomic status, race, ethnicity, and class influences on graduation rate (SI). We also used these variables to identify sociodemographic contexts in which schools in our sample were embedded, and to examine variation in associations between graduation rate and vegetation with sociodemographic context. Using the 2009/10 Common Core of Data from the National Center for Education Statistics (National Center for Education Statistics, 2010), we calculated school proportion of students eligible for free and reduced lunch, four variables indicating proportion of Black, Latino/a, Asian, or Native American twelfth grade students, and student-teacher ratios. Using data from the 2010 US Decennial Census and 2010 Five-year American Community Survey from American Fact Finder (US Census Bureau, n.d.), we calculated SAA proportions of low, medium, and high educational attainment, household income, and household size. We also used those data to calculate SAA proportion of renters. Precise definitions of low, medium, and high educational attainment, household income, and household size are included in Table 1.

American Fact Finder data were unavailable at the SAA level. We employed dasymetric mapping (Mennis, 2003) with data aggregated to unified school districts (UNSD) to predict SAA counts of people with low/medium/high educational attainment and living in rental housing and counts of small/medium/large and low/medium/high income households. Dasymetric mapping combines ancillary datasets with the original aggregated spatial data using raster overlay operations to model the underlying population density surface of a set of geographic units to which population counts are aggregated. This model is represented by a raster dataset with population density estimates for each grid-cell that are synonymous with expected count per cell, and that may be aggregated to a different set of geographic units within which they are nested.

Dasymetric mapping utilizes data related to the spatial distribution of a population to generate a population density surface model, in this case, degree of impervious coverage. After assigning zero values to locations within our UNSD where population counts were almost certainly zero based on population data at the Census block level, we used a raster dataset from the 2011 National Land Cover Database (NLCD11) (Xian et al., 2011) which we classified into three imperviousness categories, low (L), medium (M), and high (H), using the lower and upper quartiles of the imperviousness distribution after removing all zero values from the dataset. The values in the lower quartile were assigned to class L, the

Table 1
Variable descriptions and summary statistics for each school and SAA.

Variable	Description	Min	Max	Mean	SD
Area	SAA areal extent (km ²)	0.461	769.366	57.263	64.336
Graduation	High school graduation rate	0.050	0.995	0.794	0.137
Reading proficiency	Reading proficiency rate	0.050	0.995	0.708	0.228
Math proficiency	Mathematics proficiency rate	0.025	0.995	0.654	0.235
Lunch	% students eligible for free/reduced price lunch	0.000	99.830	41.411	27.196
Black	12th grade percent Black	0.000	100.000	29.255	30.854
Latinx	12th grade percent Latino/a	0.000	95.620	12.719	16.568
Low income	% households with HHI < \$35 K/yr	4.640	69.894	31.953	11.917
Medium income	% households with \$35 K/yr ≤ HHI < \$100 K/yr	10.663	61.773	44.660	5.659
High income	% households with HHI ≥ \$100 K/yr	1.493	77.890	23.448	12.524
Low education	% individuals 25 or older with no HS degree	0.284	37.265	12.545	5.703
Medium education	% individuals 25 or older with HS or undergraduate degree	51.592	90.597	75.035	5.925
High education	% individuals 25 or older with advanced degree	1.782	45.593	12.419	6.737
Small households	% households with less than 4 people	50.012	88.544	76.538	5.716
Medium households	% households with 4 or 5 people	9.761	38.912	19.333	4.531
Large households	% households with more than 5 people	0.923	21.158	4.148	1.934
Renters	% population living in rental housing	4.399	79.319	34.578	13.080
Student-teacher ratio	Number of students per instructor	6.182	36.625	17.120	3.564
Canopy	Mean tree canopy cover (%)	0.000	86.166	25.983	17.330
Z_canopy	Mean intensity of tree canopy in SAA normalized	−286.534	861.958	12.700	99.998
Non-forest	Non-forest vegetation cover (%)	0.000	0.742	0.049	0.080
Z_non-forest	Normalized non-forest vegetation cover	−135.465	650.413	−6.467	52.112
Agriculture	Agricultural cover (%)	0.000	0.772	0.064	0.106
Z_agriculture	Normalized agricultural	−224.660	743.522	−14.720	83.325
Imperviousness	Mean impervious cover (%)	0.747	84.098	27.474	16.475
Z_imperviousness	Normalized mean impervious cover	−591.986	436.167	−0.731	128.510

values in the upper quartile to class H, and all other values to class M. We used this imperviousness classification scheme to determine weights (total fractions) consisting of six components – three density fractions and three area ratios, one for each imperviousness class. Density fractions were calculated as follows:

$$DF_i = P_i / (P_l + P_m + P_h) \quad [1]$$

DF_i is the density fraction for imperviousness class i , and P_i an estimated population density for imperviousness class i . P_l , P_m , and P_h are estimated population densities for classes L, M, and H, respectively.

Population densities were estimated by identifying a selection of UNSD best represented by the class in question and calculating the aggregate population density of that selection (i.e., total count for the selection divided by total inhabited area of the selection). An ideal selection strategy considers UNSD dominated by a particular class as candidates to estimate the population density of that class. We used this strategy for class M; however, no UNSD had proportional areas of classes L or H over 0.5. For class L, we thus selected UNSD where the proportional area of that class exceeded those of the other two classes (e.g., UNSD with proportional areas of classes L, M, and H equal to 0.4, 0.3, and 0.3 respectively). We selected UNSD in the upper decile of the distribution of class H proportional areas to estimate class H population density from remaining, unsampled UNSD. Area ratios and total fractions were calculated as follows:

$$AR_{ij} = (A_{ij} / A_j) / 0.33 \quad [2]$$

AR_{ij} is the area ratio for imperviousness class i in UNSD j , A_{ij} the area of class i in UNSD j , and A_j the total area of UNSD j . Once the density fractions (one per imperviousness class) and area ratios (one per imperviousness class in each UNSD) were calculated, we combined them to create total fractions:

$$TF_{ij} = \frac{DF_i * AR_{ij}}{DF_l * AR_{lj} + DF_m * AR_{mj} + DF_h * AR_{hj}} \quad [3]$$

TF_{ij} is the total fraction for imperviousness class i in UNSD j , DF_i is the density fraction of class i , AR_{ij} the area ratio for class i in UNSD j , DF_l , DF_m , and DF_h are the density fractions for each respective class, and AR_{lj} , AR_{mj} , and AR_{hj} are the area ratios for each class in UNSD j .

We calculated three per-cell population density estimates per UNSD, one for each imperviousness class, by multiplying by its TF and dividing by the number of inhabited cells in the UNSD. We extracted these values to their corresponding class L, M, and H cells to create the final population surface model. We executed this process for each variable derived from American Fact Finder data, and for UNSD total population, population 25 years of age and over, and number of housing units to transform counts to rates. Spatial non-stationarity in relationships between imperviousness and population density could complicate our approach (Mennis, 2003). While applying the analysis iteratively to sub-regions (i.e., counties) could address this issue, doing so was impractical given the sparse UNSD distribution (approximately 740 UNSD across the US).

2.4. Environmental variables

We calculated four variables to represent vegetated and built characteristics of student environments. We identified SAA mean canopy coverage using the 2011 National Landcover Dataset (NLCD) 30-m cartographic tree canopy product (Coulston et al., 2012, 2013), and SAA mean imperviousness using the 2011 NLCD 30-m percent developed imperviousness product (Xian et al., 2011). We calculated SAA proportional area of non-forest (shrub/scrub, grassland/herbaceous, emergent herbaceous wetland) and agricultural (pasture/hay, cultivated crop) vegetation using the 2011 NLCD 30-m land-cover product (Homer et al., 2015).

We standardized environmental covariates to account for variation in dominant vegetation and development patterns among urban areas. We transformed canopy coverage and imperviousness values to Z-scores to indicate differences between SAA mean canopy coverage and imperviousness and urban area mean canopy coverage and imperviousness, respectively, as:

$$Z_c = (C_{SAA} - C_{UA}) / (SD_{C_{UA}} / \sqrt{g_{SAA}}) \quad [4]$$

Z_c is the transformed canopy or impervious coverage mean for an SAA, C_{SAA} is the original SAA mean, C_{UA} is the urban area mean, $SD_{C_{UA}}$ represents the SD of the urban area canopy or impervious coverage distribution, and g_{SAA} is the number of SAA grid cells. SAA proportional area of non-forest and agricultural vegetation were standardized as:

$$Z_a = \frac{PLC_{SAA} - PLC_{UA}}{\sqrt{\frac{PLC_{UA} * (1 - PLC_{UA})}{g_{SAA}}}} \quad [5]$$

PLC_{SAA} indicates SAA proportion of a given land cover, while PLC_{UA} indicates urban area proportion of that land cover. Transformed in this way, negative values for environmental covariates indicate SAA with lower coverage than their urban area. Positive values indicate SAA with higher coverage than their area.

2.5. Analyses

2.5.1. K-means Clustering

Using K-means clustering, we grouped schools based on socio-demographic and environmental characteristics via two separate analyses in R 3.5.1 (R Team, 2013) using the function “kmeans”. K-means clustering partitions records in a dataset into categories based on the mean values of the dataset attributes. The first analysis involved low/medium/high education and income variables, proportion of renters in SAA, proportions of Black, Latino/a, Asian, or Native American twelfth-grade students, and proportion of students eligible for free or reduced-price lunch. The second analysis involved SAA non-forest vegetated and agricultural land covers, mean canopy coverage, and imperviousness. All variables were mean-centered and scaled by standard deviation, which means that, for each variable, we subtracted the variable mean from each value and then divided the differences by the standard deviation of the variable. This step is necessary for the K-means algorithm to work correctly.

We selected K by examining how the total sum of squared distances (TSS) of data points from the mean-centers of their respective clusters decreased with increases in K , stopping once improvements in TSS were negligible. The K-means algorithm starts by selecting K random records from the input dataset. The attribute values of records in the dataset form coordinates; for example, a dataset with three attributes would have three-dimensional coordinates in the form of (x, y, z). Therefore, a record could be considered a point in Euclidean space. The algorithm then calculates the Euclidean distances between all the other data points and the randomly selected points and assigns each point to the closest randomly selected point, creating groups. K new coordinates are calculated consisting of the means of the attributes for each group. These are the mean-centers. The distances between all data points and the mean-centers are calculated and points assigned to the closest center. The mean-centers are then recalculated, and the process repeats until no points change assignments. To calculate TSS, the distances between data points and their respective mean-centers are squared and summed. As TSS decreases, the categories become more distinct; in other words, there is less variation within dataset groupings and more variation between values in different groupings. TSS will always decrease as K increases, but decreases will become notably smaller after a certain value of K , which is considered the optimal value for K . The quality of the assignments (a smaller TSS is better) varies based on the initial selection of K data points, which is random, thus we ran the algorithm 100 times for each K value to increase the likelihood of optimal solutions. The cluster assignments from both analyses were incorporated as dichotomous variables into subsequent analyses.

2.5.2. Generalized linear regression

We estimated mixed-effects beta regression models with complementary log-log (cloglog) link functions to investigate relationships between environmental covariates and graduation rate using the function “glmmTMB” from the package glmmTMB in R 3.5.1 (Brooks et al., 2017). We chose cloglog based on best fit as determined by Akaike's Information Criterion. Mixed-effects models estimate intercepts and slopes for sub-groups of data based on grouping factors known as random effects, allowing for inferences not otherwise possible, and reducing model error from omitted variables that vary by grouping

factor. In generalized regression, a link function transforms the covariates in the model so the estimated response value falls in a range appropriate for the dependent variable (0 to 1 in the case of graduation rates). The purpose of generalized regression is to provide a set of models that can be reliably fit to data that do not meet certain ordinary least-squares regression assumptions.

Beta regression models non-normal response variables in the range [0–1] and relaxes assumptions regarding heteroscedasticity and over-dispersion (Cribari-Neto & Zeileis, 2009). The typical regression assumption regarding heteroscedasticity is violated when variation is not constant across model residuals; in other words, when variation in subsets of residuals is considerably different from variation among the entire set of residuals. The over-dispersion assumption applies to models designed for count or rate data, Poisson, for example. If the variance of the response variable is greater than its mean, the over-dispersion assumption is violated. We employed beta regression as graduation rate was not normally distributed and was over-dispersed, and inspection of the residuals of a preliminary Poisson model indicated a problem with heteroscedasticity.

We included state random effects to account for differences in graduation requirements among US states and ecoregion random effects to account for differences in vegetation with ecoregional context (e.g., tropical wet ecoregions have more vegetation than semi-arid ecoregions). We used level-II ecoregions from the US Environmental Protection Agency (Omernik & Griffith, 2014). We chose level II as our models would not converge when considering higher levels, likely due to a small sample size given the number of parameters to be estimated. We fit models under the assumption of varying intercepts only for the same reason. Random effects significantly improved model fit based on likelihood ratio tests (ecoregion last).

We fit two models to identify relationships between SAA vegetation variables and graduation rate and variation in mean graduation rate. Model one included canopy coverage, non-forest vegetation, agricultural vegetation, and student-teacher ratio and three dichotomous variables indicating school cluster assignments from our K-means analysis. Model two also accounted for school sociodemographic context and student-teacher ratio and included five dichotomous variables indicating school cluster assignments from our environmental K-means analysis instead of vegetation covariates. We calculated variance inflation factors (VIF) to assess multicollinearity and determined it a non-issue as all VIF were under two. We also tested all two-way interactions between vegetation and sociodemographic context variables in model one.

3. Results

3.1. Sample composition

Our sample included 1,333 high schools from 36 US states (Fig. 1, Table 1). Mean SAA areal extent was 57.26 km² (SD = 64.34). Schools varied considerably in graduation rate (range: 0.050–0.995, mean = 0.794, SD = 0.137), percentage of students eligible for free or reduced-price lunch (mean = 40.860, SD = 27.038), proportion of African-American twelfth-grade students (mean = 29.173, SD = 30.859), and SAA canopy (mean = 25.983, SD = 17.33) and impervious (mean = 27.474, SD = 16.475) cover. Standardized vegetated land-cover measures indicated that SAA often had drastically lower non-forest (mean = -5.792, SD = 51.966) and crop agriculture (mean = -13.577, SD = 82.036) cover than their surroundings.

3.2. School socioeconomic and environmental typology

We identified four school social types using a sociodemographic K-means analysis (Fig. 2a). The first grouping (average SES) included schools near the sample mean for all social and demographic variables. The second group (high SES) included schools with high SAA

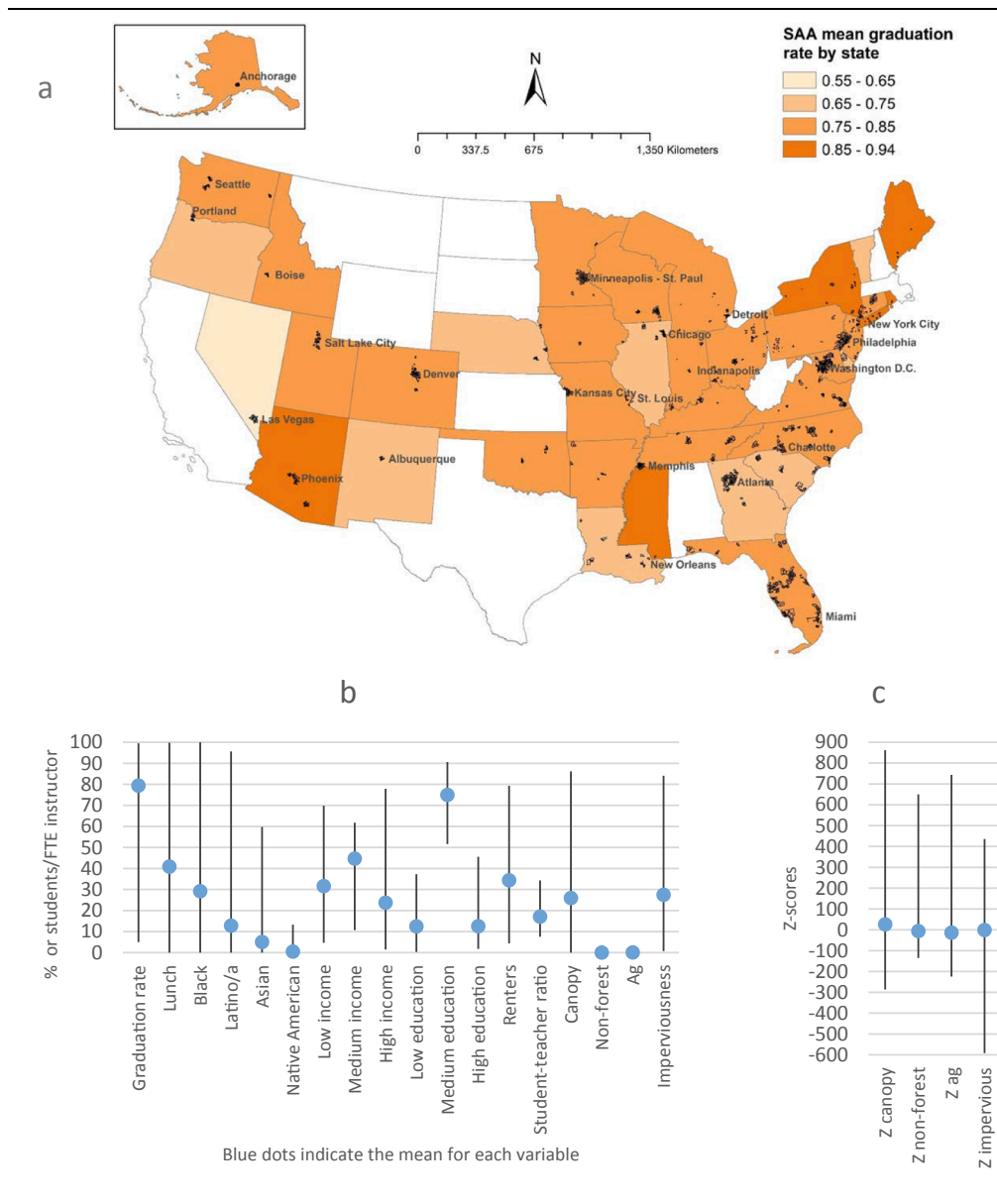


Fig. 1. Sample SAA and mean graduation rates by state (a); min, max, and mean for variables as % or students per full time equivalent (FTE) instructor for student-teacher ratio (b); min, max, and mean for transformed environmental variables (c). “Ag” refers to agricultural land cover.

educational attainment and household income. Groups 3 and 4 included schools characterized by socioeconomic disadvantage. Group 3 schools (low-SES, Black) served SAA with high impoverished and Black twelfth-grade student populations. Group 4 schools (low-SES, Latino/a) served high Latino/a student populations and exhibited low SAA educational attainment.

K-means clustering identified six school environment types (Fig. 2b). The first type (intense urban) included SAA with low canopy coverage (i. e., with few trees and/or few mature trees with broad crowns) and high imperviousness. Type two, three, five, and six schools occupied SAA with below-average imperviousness. Type two (moderate non-forest, low-intensity) exhibited above-average SAA non-forest vegetation coverage. Non-forest vegetation refers to grassland, wetland, or shrubland vegetation. Type three (high-canopy, low-intensity) had very high SAA canopy cover; in other words, type three SAA were those with many trees or large, mature trees with broad crowns and few built impervious surfaces. Type five schools (high non-forest, low-intensity) had exceptionally high SAA non-forest vegetation cover while type six schools

exhibited high SAA agricultural vegetation cover (high-agriculture, low-intensity). Agricultural vegetation cover includes cropland, for example, corn fields, wheat fields, soybean fields, or pasture and hay fields. Type four school (moderate-intensity) SAA were near the average for all variables. Fig. 2b illustrates the categories precisely in terms of standard deviations above or below means on each variable used as inputs for the K-means analysis.

We found a significant, weak correlation between social and environmental types ($V = 0.289$, $p < 0.001$). Inspecting the distribution of social types among environmental types (Fig. 2c) revealed that low-SES, African-American schools commonly occurred in high-intensity urban areas. High-SES schools typically occupied environmental contexts with high levels of vegetation. Low-SES, Latino/a schools, while common in high-intensity settings, typically occurred in highly-vegetated, low urban-intensity environments compared to low-SES African-American schools.

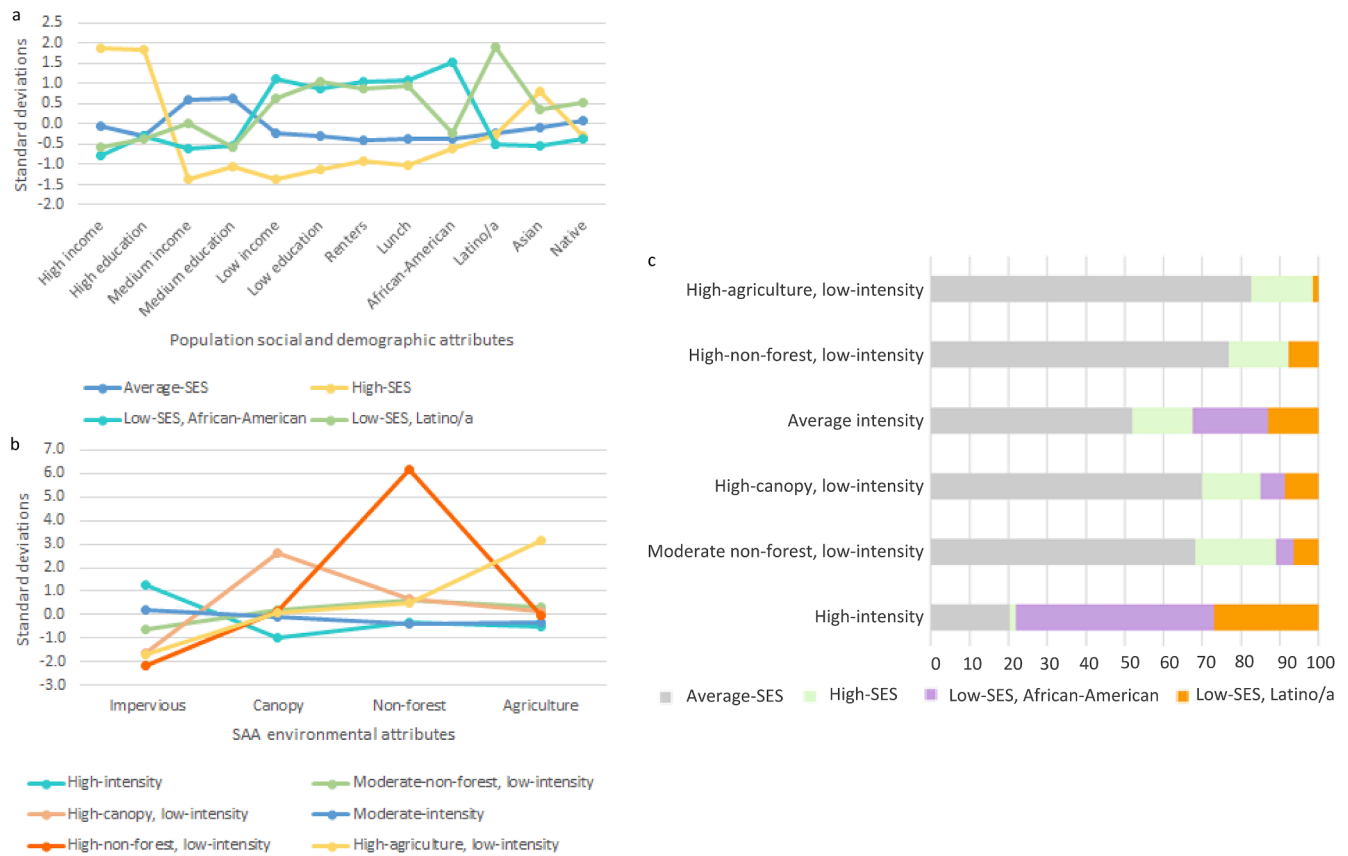


Fig. 2. Parallel coordinate plot of cluster mean-centers identified in (a) sociodemographic and (b) environmental K-means analyses. Vertical axes indicate how many standard deviations above or below an attribute's full sample mean the mean of schools in a cluster is positioned. (c) Stacked bar chart indicating distribution of social contexts across environmental contexts (percentage of observations in each social context type).

3.3. Relationships with graduation rate

All but one variable (non-forest vegetation) in a model that included canopy, non-forest, and agricultural cover; student-teacher ratio; and sociodemographic K-means school cluster assignments exhibited significant associations with graduation rate (Table 2). The sample mean graduation rate estimate (i.e., model intercept) was 0.824. High SES school estimated graduation rates averaged 8.3% above the mean average SES school rate of ($\beta = 0.083, p \leq 0.001$), while low-SES, Black and low-SES, Latino/a schools were estimated at 14.1 and 12.3% below this mean respectively ($\beta_{Soc3} = -0.141, p \leq 0.001$; $\beta_{Soc4} = -0.123, p \leq 0.001$). A one standard deviation (SD) increase in student-teacher ratio corresponded to an increase of 2.1% above the mean expected graduation rate ($\beta = 0.021, p \leq 0.001$). Schools one SD higher in canopy and

agricultural cover had mean estimated graduation rates 0.7 and 0.9% above average ($\beta_{Canopy} = 0.007, p \leq 0.05$; $\beta_{Ag} = 0.009, p \leq 0.001$).

Our second model substituted five dichotomous school environment type variables for vegetation covariates. All model covariates except moderate non-forest and high non-forest, low-intensity environments exhibited significant relationships with graduation rate (Table 2). The sample mean graduation rate estimate was 82.1%. Average estimated high SES school graduation rates were 8.2% above average ($\beta = 0.082, p \leq 0.001$). Expected graduation rates for low-SES, Black and Latino/a schools were 12.9 and 10.9% below the estimated sample mean respectively ($\beta_{Soc3} = -0.129, p \leq 0.001$; $\beta_{Soc4} = -0.109, p \leq 0.001$). A one SD increase in student-teacher ratio corresponded to a predicted graduation rate 2.1% above the estimated sample mean ($\beta = 0.021, p \leq 0.001$). The expected rate for schools in low canopy, high

Table 2
Parameter estimates for Models 1 and 2. Dependent variable is graduation rate.

Parameter	Model 1 β	SE	Z		Model 2 β	SE	Z	
Intercept	0.824	0.061	13.524	***				
Canopy	0.007	0.003	2.544	*				
Non-forest	0.002	0.002	0.905					
Agriculture	0.009	0.002	3.662	***				
High-intensity					-0.042	0.008	-5.168	***
Moderate non-forest, low-intensity					0.009	0.007	1.356	
High non-forest, low-intensity					0.031	0.010	3.003	**
High agriculture, low-intensity					-0.002	0.024	-0.084	
Student-teacher ratio	0.021	0.004	5.265	***	0.042	0.011	3.959	***
High-SES	0.083	0.007	11.945	***	0.082	0.007	11.776	***
Low-SES, African-American	-0.141	0.009	-16.177	***	-0.129	0.009	-14.430	***
Low-SES, Latino/a	-0.123	0.009	-12.970	***	-0.109	0.009	-11.744	***

* ($p \leq 0.05$), ** ($p \leq 0.01$), *** ($p \leq 0.001$).

imperviousness environments was 4.2% below the estimated sample mean ($\beta = -0.042, p \leq 0.001$). Expected high-canopy, low-intensity and high-agriculture, low-intensity school graduation rates were 3.1 and 4.2% above the estimated sample mean respectively ($\beta_{Env3} = 0.031, p \leq 0.01$; $\beta_{Env6} = 0.042, p \leq 0.001$). Average SES and high non-forest, low-intensity environmental classes were reference groups.

We observed significant interactions between canopy and high SES ($\beta = 0.061, p \leq 0.05$); low-SES, Black ($\beta = -0.078, p \leq 0.01$); and low-SES, Latino/a ($\beta = 0.112, p \leq 0.001$) schools and between low-SES, Black schools and non-forest ($\beta = -0.132, p \leq 0.01$) and agricultural ($\beta = -0.102, p \leq 0.05$) vegetation. Significant simple effects occurred in three of five of these interactions: canopy and high SES ($\beta = 0.017, p \leq 0.01$) and low-SES, Latino/a ($\beta = 0.043, p \leq 0.001$), and non-forest vegetation and low-SES, Black ($\beta = -0.044, p \leq 0.01$). Thus, positive associations between canopy coverage and graduation rate were stronger for high SES schools and strongest for low SES, Latino/a schools such that a one SD increase in canopy predicts rates 1.7 and 4.3% above the estimated means for high SES (0.903) and low SES, Latino/a (0.713) schools respectively.

SAA vegetation exhibited a negative association with graduation rate for low-SES, Black schools. Non-forest vegetation was negatively related to graduation rate for these schools such that a one SD increase in non-forest vegetation predicts a rate 4.4% below the estimated group mean (0.673, Fig. 3a). We observed a nearly significant simple effect of canopy for this group ($\beta = -0.016, p = 0.068$), suggesting a negative association between canopy and graduation rate for these schools with a one SD increase in canopy corresponding to a rate 1.6% below the estimated group mean (Fig. 3b).

4. Discussion

4.1. Urban vegetation and population-level academic success

This study addressed critical gaps in our understanding of the relationships between vegetation and academic success, identifying associations between specific vegetation types and academic success, key differences in these associations with social context, and variation in success with environmental context. The observed positive association between canopy coverage and graduation rate adds to the growing literature that supports the idea that trees can benefit population-level academic attainment in urban settings (Hodson & Sander, 2017; Kuo et al., 2018; Kweon et al., 2017; Leung et al., 2019; Li et al., 2019; Matsuoka, 2010; Sivarajah et al., 2018; Tallis et al., 2018; Wu et al., 2014). Furthermore, we document a positive relationship between agricultural vegetation and academic achievement not identified in past

studies, shedding light on additional types of vegetation that could enhance academic performance. This finding, however, may be due to the recency of development in these areas and possible links to particular populations, which warrants exploration. In identifying these associations, we help build robust knowledge of the relationships between specific types of vegetation and academic success that could support the use of urban greening to enhance academic success (e.g., by enhancing urban tree cover or by maintaining some agriculture as urbanization occurs) and thus the future well-being of urban populations.

4.2. Academic achievement and environmental context

We identified clear differences in expected graduation rates among urban environmental contexts. Our models predicted higher graduation rates for schools in SAA with more tree canopy or agricultural cover and low imperviousness, and lower rates for schools with low canopy cover, highly impervious SAA. These findings indicate clear differences in population-level responses to vegetation in highly urban versus suburban settings. Thus, urban greening programs that increase vegetative cover in inner-city environments may have lower returns than those in suburban settings. This relationship may be complex and the mechanisms by which it occurs warrant examination.

4.3. Relationships vary with social context

We also identified great and often striking disparities in the relationships between vegetation and academic success with socioeconomic context. The observed positive relationship between canopy coverage and graduation rate was stronger for high-SES schools, whose populations typically were more educated and higher income, or low-SES schools with high Latino/a student populations. Conversely, we observed a negative relationship with graduation rate for low-SES, majority African-American schools, particularly regarding non-forest vegetation, and possibly canopy coverage. These results support recent studies that identified socioeconomic disadvantage as a moderator of relationships between vegetation and academic achievement (Hodson & Sander, 2019; Kuo et al., 2018; Sivarajah et al., 2018) and suggest that the benefits of urban vegetation to academic performance tend to accrue to students not only at more advantaged schools, but also to Latino/a, low-SES student populations, while the academic success of African-American, low-SES students may be hindered by vegetation. Thus, urban greening may vary in its success at supporting academic achievement with school socioeconomic and demographic composition and greening interventions require tailoring to socioeconomic context.

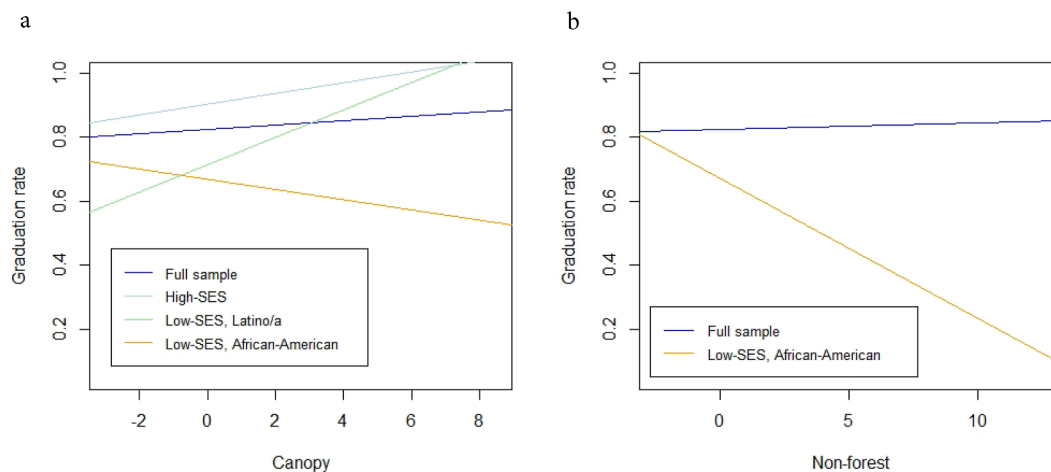


Fig. 3. Relationships between canopy and (a) graduation rate for all schools and for high-SES schools; low-SES, African-American schools; and low-SES, Latino/a schools and (b) non-forest vegetation and graduation rate for all schools and for low-SES, African-American schools.

4.4. Potential drivers of variation in relationships

Variation in these relationships might occur for several reasons. Firstly, vegetation in intensely-urban environments or low-SES, African-American neighborhoods may take a form that does not benefit academic success. For example, vacant or abandoned land may occur more frequently in such neighborhoods. Urban vacancy and abandonment are associated with denser or more extensive vegetation (Endsley et al., 2018; Pearsall & Christman, 2012) that may be less-desirable, unmanaged, and inaccessible, or which may provide locations for activities (e.g., crime, drug use) that reduce academic success. Furthermore, the influence of public and private trees likely varies. Private trees may confer greater benefits because residents experience higher exposure to trees on their property relative to those without trees. Research suggests that high-SES neighborhoods may have better access to private green spaces (e.g., gardens, yards) (Lin et al., 2015). Future research should examine relationships between urban vegetation categorized more specifically based on quality, arrangement, and accessibility, thereby identifying more desirable forms of greening that support, rather than detract from, academic success.

Secondly, student populations in highly impervious settings, which in some regions (e.g., northeastern US) tend to be poorer with a higher proportion of Black residents, may have limited experience with vegetation, suggesting that these populations may have undergone a near complete “extinction of experience” with nature. Such an “extinction of experience” (Pyle, 1993) is posited to lead to a reduced affect for and emotional connection with nature with associated loss of benefits (Soga & Gaston, 2016). Programs that increase vegetation in intensely-urban settings (e.g., urban greening) and that go beyond simple provision of vegetation to encourage use of vegetated environments (Lin et al., 2017), may thus serve to bring about “de-extinction”, reconnecting urbanites to nature and restoring benefits (Soga et al., 2015, 2016; Soga & Gaston, 2016), including benefits to academic achievement. Such increased exposure and experience with vegetated settings may be particularly crucial for children as regular outdoor play during childhood influences nature orientation (Bixler et al., 2002; Thompson et al., 2008).

4.5. A legacy of discrimination and resulting residential segregation may underlie negative relationships for Black populations

Our results suggest that residential segregation may play a role in determining relationships between urban vegetation and academic success. US cities are often highly racially segregated, a reality that particularly impacts Black populations (Rothstein, 2017). The residential choices of Black Americans have been stymied by a legacy of discrimination by federal, state, and local governments, financial institutions, realtors, home-owners associations, employers, and private citizens (Rothstein, 2017). The outcome of this legacy is the concentration of generations of Black residents into intensely-urban, inner-city neighborhoods with little vegetation and high impervious cover. The distribution of school social types across environmental types (Fig. 2c) indicates such a pattern. Many such neighborhoods, particularly in the Great Lakes region, formed during the Great Migration, a decades-long movement of Black Americans from the southeastern US to the north and west between 1915 and 1970 (Tolnay, 2003). Resulting multigenerational lack of exposure may have brought about the extinction of experience with nature posited above. Latino/a populations in many US cities have not experienced this long-term lack of exposure and may thus remain connected to nature and receive its benefits, a notion supported by frequent occurrence of low-SES, Latino/a student populations in more vegetated environments than low-SES, African-American populations (Fig. 2c). The existence and role of experience with nature in determining relationships between urban vegetation and population academic success among inner-city, particularly minority populations, thus warrants careful study to build sufficient understanding to support

urban greening programs that enhance academic achievement.

4.6. Implications and limitations

This research indicates that relationships between vegetation and school graduation rate vary with social and environmental context. While we observed slight, positive associations between vegetation and graduation rate, they were stronger for schools serving wealthier student populations, strongest for schools serving low-SES, primarily Latino/a student populations, and even negative for schools serving low-SES, primarily Black student populations. These findings suggest a need to manage urban vegetation in ways that provide contact where benefits occur and to identify and address factors, including cultural differences and inequities in access to, experience with, and quality of green urban environments, to provide all student populations with urban environments that support their academic achievement. Doing so could weaken barriers to upward social mobility, helping to close gaps between the advantaged and disadvantaged, therefore leading to a more equitable and sustainable society.

Local-level urban greening efforts combined with reformed housing policy could ensure that greening benefits the populations whose academic success it is meant to support. Currently, investment in vegetated public spaces in high social disadvantage neighborhoods often increases property values, displacing low-income residents, a process known as the “Green Space Paradox” (Wolch et al., 2014), or “green gentrification” (Immergluck & Balan, 2018; Rigolon & Németh, 2018). Several policy mechanisms designed to ensure equitable distribution of affordable housing exist, although their effectiveness varies, and implementation can face considerable legal and political opposition (Freeman & Schuetz, 2017). Thus, addressing inequitable access to nature requires change in the policy mechanisms and legal and political landscapes of cities.

Our results should be considered in light of certain limitations. Firstly, the cross-sectional nature of this study precludes causal inference, as observed relationships could result from socioecological conditions peculiar to the study year, and also precludes investigation of whether observed relationships with graduation rate represent the culmination of influences of SAA vegetation on student populations over many years of schooling. The ecological nature of this study prevents extension of conclusions to individuals or subgroups not explicitly considered in these analyses, and proportions of student populations not attending SAA schools (e.g., private or charter schools) were omitted. Although the results may not be generalizable due to the possibility of ecological fallacy, the study sample covered a considerable number of cities in the majority of states across the continental US with a few exceptions (e.g., California, Texas, Montana, North Dakota, South Dakota, and Wyoming). Many of the included cities, however, were larger cities in regions of the country with relatively high amounts of rain fall annually. Results therefore may not apply outside of those social and environmental contexts. Also, the considerable study area extent restricted our choice of environmental data in terms of spatial and thematic resolution, preventing investigation of specific vegetation attributes apparent only at finer resolutions. Additionally, observed relationships may be artifacts of the influences of social processes on vegetation distribution that our models could not assess, although we did our best to eliminate this possibility given data availability.

Future studies that examine these relationships at finer resolutions that facilitate consideration of more detailed vegetation and population attributes or that explore relationships for individuals from different socioeconomic groups and environmental settings could identify how and why relationships between urban vegetation and academic performance vary. Furthermore, while several studies consider vegetation and/or green spaces within the environments of elementary school children (Hodson & Sander, 2017; Kuo et al., 2018; Leung et al., 2019; Sivarajah et al., 2018; Tallis et al., 2018; Wu et al., 2014), to our knowledge, no such studies currently exist at a national extent within

the US, thus limiting our understanding of relationships between vegetation and academic achievement at levels of schooling other than high school to a few cities or states. Such studies would considerably contribute to our ability to manage urban vegetation to support multiple facets of urban sustainability.

CRediT authorship contribution statement

Cody B. Hodson: Conceptualization, Investigation, Methodology, Data curation, Formal analysis, Writing - original draft, Writing - review & editing, Validation, Visualization, Funding acquisition. **Heather A. Sander:** Conceptualization, Methodology, Resources, Supervision, Funding acquisition, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- Beere, P., & Kingham, S. (2017). Assessing the relationship between greenspace and academic achievement in urban New Zealand primary schools. *New Zealand Geographer*, 73, 155–165.
- Benfield, J. A., Rainbolt, G. N., Bell, P. A., & Donovan, G. H. (2015). Classrooms with nature views: Evidence of differing student perceptions and behaviors. *Environment and Behavior*, 47, 140–157.
- Bixler, R. D., Floyd, M. F., & Hammitt, W. E. (2002). Environmental socialization: Quantitative tests of the childhood play hypothesis. *Environment and Behavior*, 34, 795–818.
- Boone, C. G., Buckley, G. L., Grove, J. M., & Sister, C. (2009). Parks and people: An environmental justice inquiry in Baltimore, Maryland. *Annals of the Association of American Geographers*, 99, 767–787.
- Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., ... Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal*, 9(2), 378–400.
- Browning, M. H., Kuo, M., Sachdeva, S., Lee, K., & Westphal, L. (2018). Greenness and school-wide test scores are not always positively associated—A replication of “linking student performance in Massachusetts elementary schools with the ‘greenness’ of school surroundings using remote sensing”. *Landscape and Urban Planning*, 178, 69–72.
- Ceci, S. J., & Williams, W. M. (1997). Schooling, intelligence, and income. *American Psychologist*, 52, 1051–1058.
- College of William and Mary and the Minnesota Population Center. (2011). School attendance boundary information system (SABINS): Version 1.0. University of Minnesota Minneapolis.
- Coulston, J. W., Moisen, G. G., Wilson, B. T., Finco, M. V., Cohen, W. B., Brewer, C. K. (2012). Modeling percent tree canopy cover: A pilot study. *Photogrammetric Engineering & Remote Sensing* 78 (7): 715–727.
- Coulston, J. W., Jacobs, D. M., King, C. R., & Elmore, I. C. (2013). The influence of multi-season imagery on models of canopy cover. *Photogrammetric Engineering & Remote Sensing*, 79, 469–477.
- Cribari-Neto, F., & Zeileis, A. (2009). Beta regression in R. *Journal of Statistical Software*, 34, 1–24.
- Dadvand, P., Nieuwenhuijsen, M. J., Esnaola, M., Forns, J., Basagaña, X., Alvarez-Pedrerol, M., ... Jerrett, M. (2015). Green spaces and cognitive development in primary schoolchildren. *Proceedings of the National Academy of Sciences*, 112, 7937–7942.
- Donovan, G. H., Michael, Y. L., Gatzolis, D., & Hoyer, R. W. (2018). *The relationship between the natural environment and individual-level academic performance in Portland, Oregon: Environment and Behavior* 0013916518796885.
- Endsley, K. A., Brown, D. G., & Bruch, E. (2018). Housing market activity is associated with disparities in urban and metropolitan vegetation. *Ecosystems*, 21, 1593–1607.
- Freeman, L., & Schuetz, J. (2017). Producing affordable housing in rising markets: What works? *Cityscape*, 19, 217–236.
- Heynen, N., Perkins, H. A., & Roy, P. (2006). The political ecology of uneven urban green space - the impact of political economy on race and ethnicity in producing environmental inequality in Milwaukee. *Urban Affairs Review*, 42, 3–25.
- Hodson, C. B., & Sander, H. A. (2017). Green urban landscapes and school-level academic performance. *Landscape and Urban Planning*, 160, 16–27.
- Hodson, C. B., & Sander, H. A. (2019). Relationships between vegetation in student environments and academic achievement across the continental US. *Landscape and Urban Planning*, 189, 212–224.
- Homer, C., Dewitz, J., Yang, L., Jin, S., Danielson, P., Xian, G., ... Megown, K. (2015). Completion of the 2011 National Land Cover Database for the conterminous United States - representing a decade of land cover change information. *Photogrammetric Engineering & Remote Sensing*, 81(5), 345–354.
- Immergluck, D., & Balan, T. (2018). Sustainable for whom? Green urban development, environmental gentrification, and the Atlanta beltline. *Urban Geography*, 39, 546–562.
- Kaplan, S. (1995). The restorative benefits of nature: Toward an integrative framework. *Journal of Environmental Psychology*, 15, 169–182.
- Kuo, M., Browning, M. H. E. M., Sachdeva, S., Lee, K., & Westphal, L. (2018). Might school performance grow on trees? Examining the link between “greenness” and academic achievement in urban, high-poverty schools. *Frontiers in Psychology*, 9.
- Kweon, B.-S., Ellis, C. D., Lee, J., & Jacobs, K. (2017). The link between school environments and student academic performance. *Urban Forestry & Urban Greening*, 23, 35–43.
- Leung, W. T. V., Tam, T. Y. T., Pan, W.-C., Wu, C.-D., Lung, S.-C.-C., & Spengler, J. D. (2019). How is environmental greenness related to students’ academic performance in English and mathematics? *Landscape and Urban Planning*, 181, 118–124.
- Li, D., Chiang, Y.-C., Sang, H., & Sullivan, W. C. (2019). Beyond the school grounds: Links between density of tree cover in school surroundings and high school academic performance. *Urban Forestry & Urban Greening*, 38, 42–53.
- Lin, B., Meyers, J., & Barnett, G. (2015). Understanding the potential loss and inequities of green space distribution with urban densification. *Urban Forestry & Urban Greening*, 14, 952–958.
- Lin, B. B., Gaston, K. J., Fuller, R. A., Wu, D., Bush, R., & Shanahan, D. F. (2017). How green is your garden? Urban form and socio-demographic factors influence yard vegetation, visitation, and ecosystem service benefits. *Landscape and Urban Planning*, 157, 239–246.
- Markevych, I., Feng, X., Astell-Burt, T., Standl, M., Sugiri, D., Schikowski, T., ... Berdel, D. (2019). Residential and school greenspace and academic performance: Evidence from theGINIplus and LISA longitudinal studies of German adolescents. *Environmental Pollution*, 245, 71–76.
- Matsuoka, R. H. (2010). Student performance and high school landscapes: Examining the links. *Landscape and Urban Planning*, 97, 273–282.
- Mennis, J. (2003). Generating surface models of population using dasymetric mapping. *The Professional Geographer*, 55, 31–42.
- National Center for Education Statistics. (2010). Common core of data 2010. Retrieved from <https://nces.ed.gov/ipeds/data/ipedsdata.asp>.
- Omernik, J. M., & Griffith, G. E. (2014). Ecoregions of the conterminous united states: Evolution of a hierarchical spatial framework. *Environmental Management*, 54, 1249–1266.
- Pearsall, H., & Christman, Z. (2012). Tree-lined lanes or vacant lots? Evaluating non-stationarity between urban greenness and socio-economic conditions in Philadelphia, Pennsylvania, USA at multiple scales. *Applied Geography*, 35, 257–264.
- Pyle, R. M. (1993). *The Thunder Tree: Lessons from an Urban Wildland*. Boston, MA: Houghton Mifflin.
- R Team. (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Rigolon, A., & Németh, J. (2018). “We’re not in the business of housing:” environmental gentrification and the nonprofitization of green infrastructure projects. *Cities*, 81, 71–80.
- Rothstein, R. (2017). *The Color of Law: A Forgotten History of How Our Government Segregated America*. New York, NY: Liveright Publishing Corporation.
- Schraml, K., Perski, A., Grossi, G., & Makower, I. (2012). Chronic stress and its consequences on subsequent academic achievement among adolescents. *Journal of Educational and Developmental Psychology*, 2, 69.
- Schutte, A. R., Torquati, J. C., & Beattie, H. L. (2015). Impact of urban nature on executive functioning in early and middle childhood. *Environment and Behavior*, 49, 3–30.
- Sivarajah, S., Smith, S. M., & Thomas, S. C. (2018). Tree cover and species composition effects on academic performance of primary school students. *Plos One*, 13.
- Soga, M., Yamaura, Y., Aikoh, T., Shoji, Y., Kubo, T., & Gaston, K. J. (2015). Reducing the extinction of experience: Association between urban form and recreational use of public greenspace. *Landscape and Urban Planning*, 143, 69–75.
- Soga, M., Gaston, K., & J. (2016). Extinction of experience: The loss of human–nature interactions. *Frontiers in Ecology and the Environment*, 14, 94–101.
- Soga, M., Gaston, K. J., Koyanagi, T. F., Kurisu, K., & Hanaki, K. (2016). Urban residents’ perceptions of neighbourhood nature: Does the extinction of experience matter? *Biological Conservation*, 203, 143–150.
- UN Population Division. (2018). *World Urbanization Prospects: The 2018 Revision*. Department of Economic and Social Affairs, Population Division: United Nations.
- US Census Bureau. (n.d.) American Fact Finder. Retrieved from <https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml#ll>.
- US Census Bureau. (2010). 2010 Census Urban and Rural Classification and Urban Area Criteria. Retrieved from <https://www.census.gov/geo/reference/ua/urban-rural-2010.html>.
- US Department of Education. (2018). *The Condition of Education*. Washington: DC - National Center for Education Statistics.
- US Department of Education. (n.d.) EduFacts. Retrieved from <https://www2.ed.gov/about/units/ed/edfacts/data-files/index.html#acgr>.
- Tallis, H., Bratman, G. N., Samhour, J. F., Fargione, J. (2018). Are California elementary school test scores more strongly associated with urban trees than poverty? *Frontiers in Psychology* 9.
- Thompson, C. W., Aspinall, P., & Montarzino, A. (2008). The childhood factor: Adult visits to green places and the significance of childhood experience. *Environment and Behavior*, 40, 111–143.
- Tolnay, S. E. (2003). The African American “great migration” and beyond. *Annual Review of Sociology*, 29, 209–232.

- Ulrich, R. S., Simons, R. F., Losito, B. D., Fiorito, E., Miles, M. A., & Zelson, M. (1991). Stress recovery during exposure to natural and urban environments. *Journal of Environmental Psychology*, 11, 201–230.
- Wolch, J. R., Byrne, J., & Newell, J. P. (2014). Urban green space, public health, and environmental justice: The challenge of making cities 'just green enough'. *Landscape and Urban Planning*, 125, 234–244.
- Wu, C.-D., McNeely, E., Cedeño-Laurent, J. G., Pan, W.-C., Adamkiewicz, G., Dominici, F., ... Spengler, J. D. (2014). Linking student performance in Massachusetts elementary schools with the "greenness" of school surroundings using remote sensing. *PLoS One*, 9, Article e108548.
- Xian, G. Z., Homer, C. G., Dewitz, J., Fry, J., Hossain, N., & Wickham, J. (2011). Change of impervious surface area between 2001 and 2006 in the conterminous United States. *Photogrammetric Engineering and Remote Sensing*, 77, 758–762.