Landscape and Urban Planning 178 (2018) 69–72

Contents lists available at ScienceDirect

Landscape and Urban Planning

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

Research Note

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"

Matthew H.E.M. Browning[a,b,d,*], Ming Kuo[b,*], Sonya Sachdeva[c], Kangjae Lee[d], Lynne Westphal[c]

[a] Department of Recreation, Sport and Tourism, University of Illinois at Urbana-Champaign, Champaign, IL, USA
[b] Department of Natural Resources and Environmental Sciences, University of Illinois at Urbana-Champaign, Urbana, IL, USA
[c] USDA Forest Service, Northern Research Station, Evanston, IL, USA
[d] Illinois Informatics Institute, University of Illinois at Urbana-Champaign, Champaign, IL, USA

ARTICLE INFO

Keywords:
Academic performance
Remote sensing
Vegetation
Greenness
Replication

ABSTRACT

Recent studies find vegetation around schools correlates positively with student test scores. To test this relationship in schools with less green cover and more disadvantaged students, we replicated a leading study, using six years of NDVI-derived greenness data to predict school-level math and reading achievement in 404 Chicago public schools. A direct replication yielded highly mixed results with some significant positive relationships between greenness and academic achievement, some negative, and some null – but accompanying VIF scores in the thousands indicated untenable levels of multicollinearity. An adjusted replication corrected for multicollinearity and yielded stable results; surprisingly, all models then showed near-zero but statistically significant negative relationships between greenness and performance. In low-green, high-disadvantage schools, negative greenness-academic performance links may reflect the predominance of grass in measures of overall greenness and/or insufficient statistical controls for the moderating effect of disadvantage.

1. Introduction

A growing body of evidence suggests schools with more greenspace around them perform better academically, even after controlling for various confounders. In a 2014 PLOS ONE article, Wu and colleagues reported that the relative greenness in 250 m to 2000 m buffers around Massachusetts public schools correlated with 3rd-grade standardized test scores. Earlier, Matsuoka (2010) discovered that classroom window views of trees and shrubs correlated with graduation rates and academic merit awards in Michigan high schools. Recently, Kweon, Ellis, Lee, and Jacobs (2017) found that more trees in the schoolyards of Washington D.C. public elementary, middle, and high schools also correlated with higher test scores. Hodson and Sander (2017) found that trees correlated with third graders' reading test scores in Minneapolis-St Paul, Minnesota. And Sivarajah, Smith, and Thomas (2018) found that the positive association between tree cover and third grade student performance was most pronounced in the most socio-economically disadvantaged schools of the Toronto District School Board. Based on these findings, it would be tempting to assume the existence of a "greenness-academic performance" (G-AP) link, which describes more green cover generally correlating with greater achievement.

However, we know that greenness seems to impact one academic subject (i.e., reading comprehension or mathematical computation) in some populations but not that same subject in other populations. The study in Toronto found better math scores, but not reading and writing scores, were correlated with tree cover (Sivarajah et al., 2018). The study in Minneapolis found the opposite: reading scores, but not math scores, correlated with tree cover (Hodson & Sander, 2017). And both reading and math scores were associated with tree canopy cover in Washington D.C. (Kweon et al., 2017). Further work on the impact of greenness on both reading and math in diverse contexts is therefore warranted.

Further, G-AP findings to date have been obtained primarily in contexts with relatively high green cover. Massachusetts is the second most tree-covered state in the United States, Michigan is one of the highest tree-covered states in the Midwest, Washington DC and Minneapolis have more canopy cover than the average cover found in at least 20 U.S. cities, and Toronto has higher levels of green views than many cities across the globe (Nowak & Greenfield, 2012; Nowak, Noble, & Sisinni, 2001; O'Neil-Dunne, 2010; Ratti, Seiferling, Li, Ghaeli, & So, 2018). Because past work has been in relatively "green" areas, it's not

* Corresponding authors at: Department of Recreation, Sport and Tourism, University of Illinois at Urbana-Champaign, 1206 S 4th Street, Champaign, IL 61802, USA.
E-mail addresses: brownin@illinois.edu (M.H.E.M. Browning), fekuo@illinois.edu (M. Kuo), sonyasachdeva@fs.fed.us (S. Sachdeva), klee171@illinois.edu (K. Lee), lwestphal@fs.fed.us (L. Westphal).

https://doi.org/10.1016/j.landurbplan.2018.05.007
Received 17 October 2017; Received in revised form 3 May 2018; Accepted 7 May 2018
0169-2046/

Table 1
Measures.

Measure	Original study	Replication
Reading	% 3rd graders progressing toward standard[a]	% 3rd graders meet or exceed standard
Math	% 3rd graders progressing toward standard[a]	% 3rd graders meet or exceed standard
Greenness	Average Normalized Difference Vegetation Index (NDVI) in March, July and October of each year studied (2006–2012) for 250 m, 500 m, 1000 m, and 2000 m radial buffers surrounding each school	Average Normalized Difference Vegetation Index (NDVI) in March, July and October of each year studied (2006–2012) for 250 m, 500 m, 1000 m, and 2000 m radial buffers surrounding each school
Race/ethnicity	% African American; % Asian; % Hispanic; % White; % Native American; % Native Hawaiian	% African American; % Asian; % Hispanic; % White; % Native American; % Native Hawaiian (specific to 3rd grade)
Gender	% female students	% female students (specific to 3rd grade)
Bilingual	% students whose first language is a language other than English ("First Language Not English")	% all students in school (not just in 3rd grade) who did not meet requirements for English screening survey when they first entered Chicago Public School system ("English Language Learners")
Low income	% students who qualified for any one of the following: (a) free or reduced lunch, (b) Transitional Aid to Families benefits, or (c) food stamps	% all students in school (not just in 3rd grade) who were eligible for free lunch
Pupil teacher ratio	Ratio of the number of students in the school divided by the number of all teachers in the school	Ratio of the number of students in the school (not just in 3rd grade) divided by the number of all teachers in the school
Attendance	Ratio of the total number of school days all 3rd graders were enrolled divided by the total number of school days 3rd graders attended[b]	Ratio of the total number of school days all 3rd graders were enrolled divided by the total number of school days 3rd graders attended

[a] Authors described academic performance in two ways: percentage of students progressing toward proficiency and percentage of students exceeding standard (see Table 2 in Wu et al., 2014), but correspondence with authors confirmed the former measurement (John Wu, personal correspondence, August 17, 2016).

[b] Authors described calculating attendance and academic performance with 3rd graders only; whether other variables were specific to 3rd graders or all grades was not described in the original study.

clear whether the G-AP link will hold up in other areas of the country.

Furthermore, most G-AP findings to date have been obtained from school districts serving relatively low levels of disadvantaged students. In regards to economic disadvantage, Minnesota, Michigan, and Massachusetts samples are well below the nationwide free and reduced lunch eligibility average of 48% (U.S. Department of Education, 2017a). Similarly, the G-AP studies in these areas had samples with only 25% to 40% racial or ethnic minorities. One exception is the G-AP study in Washington D.C. Its sample averaged 65% free or reduced lunch eligibility and 94% minority status; its findings provide preliminary evidence that school greening might be a relatively low-cost way to help students in greater need of educational success than more privileged students. However, nearly one-quarter of public schools serve even poorer students than Washington D.C (Rich, Cox, & Bloch, 2016) and academic performance *exponentially* drops with small decreases in parental income (U.S. Department of Education, 2017b). Therefore the extent to which greenspace may improve performance in "highest-poverty" schools serving 75% to 100% of free or reduced lunch eligible students is an important pending question in the G-AP literature.

We examined the G-AP link in low green, high disadvantage schools by replicating a leading G-AP study conducted by Wu et al. (2014) within Chicago Public Schools (CPS). This system includes over 650 schools that serve over 380,000 students, of which 47% Hispanic or Latino, 37% African American, 10% White, and 4% Asian while 17% are English Language Learners (Chicago Public Schools, 2011).

2. General methods

We attempted a replication of Wu and colleague's G-AP study using Generalized Linear Mixed Models (GLMMs), Chicago Public School school-level data, and remote sensing greenness measures. We follow the "close replication" recipe from experimental psychology, in which we aim to recreate a study as closely as possible so that the only differences are the inevitable ones related to different participants (Brandt et al., 2014). Specifically, we tested the relationship between greenness and math and reading outcomes in 3rd grader's math and reading standardized test scores while controlling for confounders used in the Massachusetts study over a six-year time period. Based on the poor model fits of this initial attempt, we adjusted our replication to produce valid estimates of the G-AP relationship in our study context (see Direct replication analysis and results and Adjusted replication analysis and results sections below).

Table 1 shows how our academic achievement, greenness, and confounder variables were analogous to those used by Wu et al. (2014), with one exception. Academic performance was operationalized as the extent to which students progressed *toward* proficiency in the original study rather the extent to which students *met or exceeded* proficiency. Most data were obtained from the CPS Data Portal (http://cps.edu/SchoolData/Pages/SchoolData.aspx) except for the percentage of female students, which was obtained through a CPS data request, and pupil-teacher ratio, which was previously downloaded by the study authors and is no longer available online. In contrast to the original study, which had confounder data for each year of academic achievement data, we applied 2009–2010 school-level characteristics to all years of data (2006–2012), because this was the only year in which complete data were available for all confounds.

Greenness data were identical to the original study and obtained from NASA's Moderate Resolution Imaging Spectroradiometer (MODIS) Vegetation Indices website (https://modis-land.gsfc.nasa.gov/vi.html). Just as in Wu and colleagues' paper, we calculated the normalized difference vegetation index (NDVI) from 250 m MODIS remote sensing images for three months (March, July, and October) in each year of our CPS data (2006–2012). NDVI shows the density of "greenness" and is calculated with the visible and near-infrared light reflected by vegetation. Healthy vegetation absorbs most visible light that hits it and reflects most near-infrared light. Thus, healthy vegetation returns higher NDVI values than unhealthy vegetation.

Data were available from 448 CPS schools. Forty-four were removed because they included non-neighborhood schools (i.e., charter and magnet). The final sample size was 404.

Table 2 shows how Chicago Public Schools provided a high-disadvantage, low-greenness context in which to examined the G-AP relationship. The average proportion of low-income students per school in our sample was nearly twice that of the original study, and the proportion of African American students and Hispanic Students was nearly threefold larger. Also, greenness in our sample was roughly one-half that of the original study.

3. Direct replication analysis and results

Our initial replication of Wu et al. (2014) involved selecting the

Table 2
Descriptive statistics.

Variable	Mean (St. Dev.)		Range	
	Original study	Replication	Original study	Replication
Reading standard (%)[a]	60(19)	32(12)	47–100	0–72
Math standard (%)[a]	61(20)	33(15)	48–100	3–83
Low income (%)	35(30)	83(22)	9–99	7–100
Bilingual (%)	16(19)	17(14)	2–94	0–53
Female (%)	48(3)	49(6)	47–64	22–69
Pupil teacher ratio	14(3)	18(3)	13–73	6–37
Number of students	95(1)	91(2)	95–100	84–98
African American (%)	8(13)	35(39)	1–89	0–100
Asian (%)	6(8)	4(7)	1–71	0–42
Hispanic (%)	16(21)	48(36)	3–99	0–100
White (%)	67(30)	13(19)	48–100	0–76
Greenness[b]	54(18)	29(11)	-30–96	0–71

[a] In the original study, descriptive statistics were reported to be provided for percent exceeding proficiency.
[b] NDVI values were multiplied by 100 in original study and replication.

same variables and applying the same equation to our data. We developed this equation from the original article and confirmed its specifications through communication with the corresponding author (John Wu, personal correspondence, August 17, 2016). We calculated 24 GLMMs to predict achievement in two subjects across six years for three months each year using four buffers and NDVI-derived greenness values. The original equation included a random effect variable for county to control for spatial autocorrelation; in this study, we found that controlling for county was unnecessary because all schools in this sample were from the same county. For our replication, we tested Moran's I on the residuals of reading and math of GLMMs without the random effect variable for county, and there was no spatial autocorrelation for either reading (Global Moran's I = 0.011, Z = 0.3, $p > 0.05$) or math (Global Moran's I = −0.025, Z = −0.5, $p > 0.05$).

However, this replication showed violations of basic model assumptions and its relationships between greenness and test scores were invalid. Variance inflation factor (VIF) levels were considerably above even the most liberal recommended threshold of 10.0 (Field, 2014), extending up to the thousands—approximately 2300 for % white and 8750 for % Hispanic—indicating substantial multicollinearity. Multicollinearity can increase standard errors of coefficients and can even cause coefficients to change sign (direction) from negative to positive and vice versa. Further, we were concerned about temporal autocorrelation, since variations in temperature and precipitation impact leaf health. These variations could result in systematic error in NDVI values (Wang, Rich, & Price, 2003).

4. Adjusted replication analysis and results

Our adjusted replication operationalized race/ethnicity as a single category (white or non-white); this was successful in reducing VIF

Table 3
Results.

Coefficient[a]		Achievement	Greenness
Original study	Adjusted replication	Meet or exceed standard	NDVI buffer month, size
-0.17**	-0.034***	Reading	OCT_2000m
-0.12**	-0.044***	Reading	OCT_1000m
-0.11**	-0.045***	Math	OCT_2000m
-0.07**	-0.049***	Math	OCT_1000m
-0.06**	-0.048***	Reading	OCT_500m
-0.04*	-0.051***	Math	OCT_500m
-0.01	-0.045***	Math	OCT_250m
-0.001	-0.047***	Reading	JULY_500m
0	-0.038***	Reading	OCT_250m
0.02	-0.044***	Reading	JULY_1000m
0.04**	-0.032***	Reading	JULY_2000m
0.05*	-0.041***	Math	JULY_2000m
0.05**	-0.051***	Math	JULY_500m
0.06**	-0.039***	Reading	JULY_250m
0.06**	-0.045***	Math	JULY_1000m
0.09**	-0.047***	Math	JULY_250m
0.18**	-0.027**	Reading	MAR_250m
0.20**	-0.040***	Math	MAR_250m
0.24**	-0.045***	Math	MAR_500m
0.30**	-0.037***	Reading	MAR_500m
0.30**	-0.038***	Math	MAR_1000m
0.32**	-0.038***	Math	MAR_2000m
0.38**	-0.030***	Reading	MAR_1000m
0.42**	-0.027**	Reading	MAR_2000m

[a]Sorted by coefficients from the original study. Blue highlight denotes a significant positive relationship, and yellow highlight denotes a significant negative relationship.
[b]Values from Table 3 in Wu et al. (2014), *$p < 0.05$, **$p < 0.01$, ***$p < 0.01$.

scores to below 3.0 (one of the more conservative recommended thresholds, Field, 2014). To address temporal autocorrelation, we included year as a random effects variable. Again, we calculated 24 models, using GLMMs to account for the random effects of year.

Surprisingly, the adjusted replication showed *only* significant *negative* relationships between greenness and test scores, although the coefficients were near zero —less than a tenth of a percentage point (Table 3).

5. Discussion

We applied a previous greenness-academic performance equation to data from a low-greenness, high-disadvantage context and found no convincing evidence for a positive relationship between greenness and academic performance (G-AP). An initial replication showed violations of basic model assumptions. When the equation was adjusted to address multicollinearity, the findings became valid—but *all* models showed a near zero but significant negative relationship between greenness and academic performance.

Do these findings suggest that the G-AP link is either nonexistent or even slightly negative in low-greenness, high-disadvantage contexts? Perhaps. However, there is reason to think that the positive G-AP relationship found elsewhere also holds in low-greenness, high-disadvantage contexts, and that these contexts simply demand a more nuanced modeling approach than has sufficed elsewhere.

First, in low-greenness contexts, it may be especially important to distinguish between tree cover and other kinds of green cover. NDVI measures do not distinguish between vegetation types, but Urban Tree Canopy assessments do — and show trees to be positively correlated with academic performance, but grass and shrubs negatively correlated with performance (Kweon et al., 2017). Chicago has relatively low levels of tree cover (Nowak & Greenfield, 2012); in low-greenness contexts where NDVI values primarily represent grass cover, NDVI may tend not to be related to academic outcomes, or may tend to be negatively related to academic outcomes despite a positive relationship between *tree* cover and academic outcomes (Kuo, Browning, Sachdeva, Westphal, & Lee, in review).

Second, in high-disadvantage contexts, it may be important to consider potential moderating effects of poverty and race on the G-AP relationship. Socioeconomic status is strongly linked to green space access and exposure, because crime rates (Jansson, Fors, Lindgren, & Wiström, 2013), zoning policies (Wilson, Clay, Martin, Stuckey, & Vedder-Risch, 2003), and vacancy rates (Kremer & Hamstead, 2015) negatively impact the quality of green cover in poor neighborhoods. Socioeconomic status is also one of strongest predictors of academic achievement (Organisation for Economic Co-operation, 2013). And indeed, research in Chicago and Toronto public schools find a significant moderation of the G-AP relationship by disadvantage; further, when this moderator is included in the model, the G-AP relationship is significant and positive (Sivarajah et al., 2018).

6. Conclusion

Despite a handful of studies finding positive correlations between school greenness and academic performance, this replication of a leading study shows that previous models may fail to capture positive G-AP links in low-green, high-disadvantaged contexts. We recommend that future G-AP researchers distinguish between tree and other forms of green cover, and include disadvantage not only as a confounder but also as a moderator of the effects of green cover on academic achievement.

Acknowledgements

This work was supported by the 12-JV-11242309-084 Joint Venture Agreement between the University of Illinois at Urbana-Champaign and the USDA, Forest Service Northern Research Station. The authors thank Sara Dickerson and the entire Chicago Public School Research Review Board for providing data for this study.

References

Brandt, M. J., IJzerman, H., Dijksterhuis, A., Farach, F. J., Geller, J., Giner-Sorolla, R., & Van-Veer, A. (2014). The replication recipe: What makes for a convincing replication? *Journal of Experimental Social Psychology, 50*(C), 217–224. http://dx.doi.org/10.1016/j.jesp.2013.10.005.

Chicago Public Schools. Developing a School Recess Plan (pp. (2011). 1–36) Retrieved from http://www.cps.edu/News/Press_releases/Documents/DevelopingSchoolRecessPlan.pdf.

Field, A. (2014). *Discovering statistics using SPSS* (4 ed.). Thousand Oaks, CA: Sage.

Hodson, C. B., & Sander, H. A. (2017). Green urban landscapes and school-level academic performance. *Landscape and Urban Planning, 160*, 16–27. http://dx.doi.org/10.1016/j.landurbplan.2016.11.011.

Jansson, M., Fors, H., Lindgren, T., & Wiström, B. (2013). Perceived personal safety in relation to urban woodland vegetation: A review. *Urban Forestry & Urban Greening, 12*(2), 127–133. http://dx.doi.org/10.1016/j.ufug.2013.01.005.

Kremer, P., & Hamstead, Z. (2015). Transformation of urban vacant lots for the common good: an introduction to the special issue. *Cities and the Environment, 8*(2).

Kuo, M., Browning, M. H. E. M., Sachdeva, S., Westphal, L., & Lee, K. (in review). Can trees grow school performance? Examining the link between "greenness" and academic achievement in an urban, high-poverty school district. *Frontiers in Psychology*.

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. http://dx.doi.org/10.1016/j.ufug.2017.02.002.

Matsuoka, R. H. (2010). Student performance and high school landscapes: Examining the links. *Landscape and Urban Planning, 97*(4), 273–282. http://dx.doi.org/10.1016/j.landurbplan.2010.06.011.

Nowak, D. J., Noble, M. H., & Sisinni, S. M. (2001). People and trees: Assessing the US urban forest resource. *Journal of Forestry, 3*(1), 37–42.

Nowak, D. J., & Greenfield, E. J. (2012). Tree and impervious cover change in U.S. cities. *Urban Forestry & Urban Greening, 11*(1), 21–30. http://dx.doi.org/10.1016/j.ufug.2011.11.005.

Organisation for economic co-operation and development (2013). *PISA 2012 results: Excellence through equity: Giving every student the chance to succeed (V. II)*. PISA, OECD Publishinghttp://dx.doi.org/10.1787/9789264201132-en.

O'Neil-Dunne, J. P. M. (2010). *A Report on Washington, D.C'.s Urban Tree Canopy* (pp. 1–10). Available: http://www.uvm.edu/~joneildu/Blog/Reports/UTC_Report_DC.pdf.

Ratti, C., Seiferling, I., Li, X., Ghaeli, N., & So, W. (2018). *Exploring the Green Canopy in Cities Around the World*. Available: http://senseable.mit.edu/treepedia/cities/toronto.

Rich, M., Cox, A., & Bloch, M. (2016). *Money, Race and Success: How your School District Compares*. The Upshot: New York Times, April 29, 2016. Available: https://www.nytimes.com/interactive/2016/04/29/upshot/money-race-and-success-how-your-school-district-compares.html?mcubz=0.

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*(2), e0193254–11. http://dx.doi.org/10.1371/journal.pone.0193254.

U.S. Department of Education (2017a). *Number and Percentage of Public School Students Eligible for Free or Reduced-Price Lunch, by State: Selected Years, 2000-01 through 2010-11*. National Center for Educational Statistics, Digest of Education Statistics. Available: https://nces.ed.gov/programs/digest/d12/tables/dt12_046.asp.

U.S. Department of Education (2017b). Promising Results, Continuing Challenges.: *Final Report of the National Assessment of Title I*. Available: https://www2.ed.gov/rschstat/eval/disadv/promisingresults/edlite-exsum.html.

Wang, J., Rich, P. M., & Price, K. P. (2003). Temporal responses of NDVI to precipitation and temperature in the central Great Plains, USA. *International Journal of Remote Sensing, 24*(11), 2345–2364. http://dx.doi.org/10.1080/01431160210154812.

Wilson, J. S., Clay, M., Martin, E., Stuckey, D., & Vedder-Risch, K. (2003). Evaluating environmental influences of zoning in urban ecosystems with remote sensing. *Remote Sensing of Environment, 86*(3), 303–321. http://dx.doi.org/10.1016/S0034-4257(03)00084-1.

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*(10), e108548–9. http://dx.doi.org/10.1371/journal.pone.0108548.