Residential green space and behavioral and mental health in early childhood

Marnie F. Hazlehurst

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2021

Reading Committee:

Catherine Karr, Chair

Anjum Hajat

Pooja Tandon

Program authorized to offer degree:

Epidemiology

University of Washington

Abstract

Residential green space and behavioral and mental health in early childhood

Marnie F. Hazlehurst

Chair of the Supervisory Committee:

Catherine Karr

Departments of Environmental and Occupational Health Sciences, Pediatrics, and Epidemiology

Background

Natural environments, including urban green spaces, have been associated with a range of health outcomes across the life course. Green space may promote healthy development, even early in childhood. The aim of this study was to investigate the relationships between residential green space exposures and child behavioral and mental health in a socio-demographically diverse cohort in Memphis, TN. We also explored the relationship of green space measures to a broad set of neighborhood conditions, including socioeconomic and education resources.

Methods

We assessed three green space exposures—residential surrounding greenness, tree cover, and park proximity—within the Conditions Affecting Neurocognitive Development and Learning in Early Childhood (CANDLE) cohort. Behavioral and mental health outcomes at age 4, including both externalizing and internalizing behaviors, were assessed via parent-report on the Child Behavior Checklist (CBCL). Linear regression models adjusted for individual, household, and neighborhood-level confounders across multiple domains, were fit to assess the relationship between green space and child behavior. Effect modification by neighborhood socioeconomic opportunity and child sex were explored using multiplicative interaction terms.

Results

Higher residential surrounding greenness was associated with lower internalizing behavior scores. Observed associations were generally robust across a suite of sensitivity analyses, including adjustment for potential mediators. In secondary analyses, these associations were most consistent for the anxious/depressed and somatic complaints syndrome scales. We did not find any associations with externalizing behaviors or attention problems. In this study, residential surrounding greenness and tree cover were higher in neighborhoods with a higher homeownership rate, more early childhood education resources, and a lower percentage of Black residents.

Conclusions

Findings from this dissertation add to the accumulating evidence of a protective effect of residential green space for mental health in early childhood. Strengths of this study include the spatial resolution of green space measures and the assessment of behavior outcomes across a continuum. Future work may improve our understanding of these relationships by incorporating child care or school-based exposures and assessing time spent in green spaces. This research can inform the design of new green spaces or the conservation and management of existing urban green spaces, including improving access to nature features such as trees within historically underserved communities, as well as interventions targeting families with young children.

Table of Contents

List of Figures

List of Tables

Table A2. Distribution of NDVI exposure in multiple buffer sizes, derived from annual average 2011 satellite imagery and weighted by address history over the year prior to the Age 4 study visit.
Table A4. Distribution of percent tree canopy at the census block group level across multiple exposure windows in the CANDLE cohort.
Table A5. Distribution of park proximity exposure metrics derived from ParkServe data.
Table A6. Spearman correlations between primary green space measures at the current (age 4-6 visit) address and averaged across the residential history over the prior year.

Acknowledgements

This study was supported by the ECHO-PATHWAYS consortium, funded by National Institutes of Health (NIH): 1UG3OD023271and 4UH3OD023271. CANDLE is funded by the Urban Child Institute and NIH 1R01HL109977. We would like to thank the CANDLE participants, staff, and investigators for their contributions.

I would like to say a special thank you to my dissertation committee chair and mentor, Dr. Catherine Karr. Her continued support of my development as an epidemiologist, critical thinker, and writer, particularly with the transition to remote work during the COVID-19 pandemic, has made this work possible. I would also like to thank my reading committee, Dr. Anjum Hajat and Dr. Pooja Tandon, who have provided such valuable input and who have supported my scholarly development, both in the work contained in this dissertation and on a number of other projects over the years. I also want to thank the other members of my dissertation committee, Dr. Joel Kaufman, Dr. Adam Szpiro, and Dr. Liliana Lengua. I am grateful for the time and expertise that each committee member has graciously shared with me.

Last but not least, I would like to recognize my colleagues, friends, family, and partner, who have supported me through both the challenges and triumphs of completing this work. You all continue to inspire me and I am truly grateful.

Introduction

Background

Recent scientific literature has investigated how the structure of neighborhoods and urban landscapes, including urban green spaces and green infrastructure, may influence human health and well-being (1). Contact with nature and green space exposure has been conceptualized in terms of quantity of vegetation or green spaces, access to or proximity of green spaces, time spent in parks and green spaces, or views of natural environments and green spaces (2,3). Green space may include wooded areas, street trees, gardens, grassy areas, and parks (4,5). Understanding the potential role of these natural features in promoting health may improve urban planning and help guide intervention strategies.

A range of health outcomes have been considered in the existing research literature on green space, including both physical and mental health clinical outcomes and intermediate measures. Multiple mechanisms for the effects of green space on these health outcomes have been theorized. Much of the literature has focused on how green space may act through reducing harmful exposures such as air pollution, restoring capacity in psychological and cognitive processes, or promoting health behaviors such as physical activity (6–8). For children, nature contact and natural environments may also promote creative, imaginative, and collaborative play behaviors (9). The relevance of pathways in each of these domains may vary depending on the health endpoint under consideration. In pediatric populations, studies have largely focused on physical activity, though an accumulating literature has also investigated behavioral and mental health outcomes (10,11).

For example, the evidence to date suggests an inverse relationship between green space and externalizing behavior problems in childhood. Outcomes in this domain, characterized by behavior directed outwards at an individual's environment, may include attention problems, rule-breaking, and aggressive behaviors (12). The onset of externalizing behavior problems in early childhood may interfere with functioning in school and social settings and are associated with further externalizing behaviors in adolescence, as well as academic problems, substance use disorders, accidental injury, and obesity across the life course (13,14). Prior research on green space has focused on hyperactivity and inattention outcomes, including Attention Deficit/Hyperactivity Disorder (ADHD) diagnosis and symptoms (10). The prevalence of ADHD is estimated to be 10.5% among children ages 3-5 years old (15). Approximately half of children with ADHD in the preschool years are diagnosed with the predominantly hyperactive-impulsive subtype of ADHD and 23% with the predominantly inattentive subtype. Symptoms of hyperactivity and impulsivity are often observed by the time a child reaches four years old. Among school-age children ages 6-12 years old, the prevalence of ADHD is estimated to be 11.4%, with the inattentive subtype comprising the largest percentage of cases (15). Symptoms of ADHD have been linked to deficits in executive functions including working memory and inhibition, even during preschool years (16,17). Working memory, including both verbal and visuospatial working memory, involves holding and manipulating information in the short term (18). Inhibitory control of attention, or attentional control, is the ability to direct attention towards a target while simultaneously ignoring distractions.

The research examining relationships of natural environments with attention and executive functioning includes clinical diagnoses as well as intermediate outcome measures. In a national cohort of Danish children, residential green space in early childhood was associated with a lower risk of ADHD diagnosis (19). Studies of residential green space have also suggested associations with executive functioning, including working memory and attentional and inhibitory control (20,21). Psychologists have theorized that natural environments directly facilitate recovery of cognitive processing resources, as described by Attention Restoration Theory (ART) (22–24). Experimental evidence suggests that cognitive processing tasks requiring control of attention towards both internal and external stimuli or suppression of both internal and external distractors may be sensitive to features of natural environments (24). Alternatively, urban green space may provide spaces for physical activity, which may in turn impact executive functions in children with and without an ADHD diagnosis (25). Associations between green space and externalizing behavior problems may also occur through reduction of environmental hazards such as air pollution, which has been identified as a potential contributor to the development of externalizing behavior problems (26).

Existing literature, primarily in adolescent and adult populations, also suggests associations between green space and anxiety and depression. Behavior problems in the internalizing domain—behaviors directed towards the self, including depression, anxiety, and social withdrawal—can emerge early in childhood (27). Anxiety disorders are the most frequent mental health condition among children, with some prevalence estimates ranging up to 20% (28). While some children may exhibit transient symptoms as part of typical development, others experience wide-ranging and lasting effects on health and well-being (27). Risk factors for early childhood internalizing behaviors have been conceptualized as a cumulative risk model of multiple factors within multiple domains; environmental factors, parental factors, and child factors may all influence the risk of behavioral problems (29). The residential neighborhood context has been identified as an important contributor to behavioral development (30–32).

Green space may be a significant feature of the residential neighborhood contributing to healthy child behavioral development and mental health. A small number of studies of residential green space have observed associations of higher green space with reduced symptoms of depression or anxiety in school-age children and adolescents (33–35). Observational studies have also identified associations with emotional or peer problem behaviors earlier in childhood (36,37). Physical activity, social contacts, stress reduction, and attention restoration may play a role in the relationships between natural environments and mental health. Green space may promote child physical activity and free play, or facilitate social contacts, which in turn may reduce internalizing behaviors (38,39). Others have hypothesized that Stress Recovery Theory (SRT) may explain relationships between natural environments and health (40). Specifically, contact with nature may directly impact perceived stress and physiological measures of the stress response, including regulation of autonomic nervous system activity and the hypothalamic–pituitary–adrenal (HPA) axis. In adults, studies suggest both acute and long-term effects on these systems with exposure to natural environments (41).

Green space may be particularly important for children and families with access to fewer resources. Low SES populations are more likely to experience higher levels of adverse stressors and environmental exposures, or may be more reliant on resources within the residential neighborhood, each of which may suggest a greater potential benefit of residential green space exposures (42). Several studies have identified relationships between green space and child behavioral and mental health that are more pronounced among those with lower parent education levels or lower household income (43–45). However, others have identified stronger associations between green space and health in higher income areas and suggest that this may be due to higher perceived quality of green space translating to increased use of those spaces and improved health outcomes (46,47).

Despite recent interest and expansion of the scientific literature on nature contact and human health, our understanding of relationships between natural environments and pediatric health is still limited. Epidemiologic studies have varied widely in design, including in selection of study population, geographic location, and exposure measurement. Many prior studies have recruited pediatric samples from cities in Europe and Australia, but differences in ecology, meteorology, and urban structure may limit also generalizability of those studies to urban settings in the US. Studies in the US have used small sample sizes or high-income samples. Exposure assessment is commonly limited in both spatial and temporal resolution and analyses have frequently included only partial adjustment for potential confounders (3).

This study addresses several current gaps in our knowledge of the relationship of these spaces to children's neurodevelopment within urban settings in the US. Understanding the health consequences of green space among vulnerable groups, particularly children, is increasingly important as urbanization reduces and restricts access to these places. Furthermore, access to nearby nature may be more limited for some children, including children of color and families with lower incomes. Research in this area may inform urban design and planning in the building of new green spaces or in the conservation and management of existing urban green spaces.

Objectives

The goal of this work was to investigate the relationships between residential green space exposures and child behavioral and mental health within the Conditions Affecting

Neurocognitive Development and Learning in Early Childhood (CANDLE) study. We leveraged a wealth of existing health and covariate data within the CANDLE cohort, in combination with several green space exposure metrics developed for this project, calculated for the residential address history of each participant. The work is presented in three chapters as follows.

Chapter 1 details the distribution of green space exposures in the CANDLE cohort by a suite of neighborhood characteristics. Multiple statistical approaches were utilized in order to examine two distinct scientific questions. First, we estimate the magnitude of the relationship between a broad set of neighborhood conditions and green space, and identify the neighborhood features driving those relationships. We hypothesized that more neighborhood resources across multiple domains is associated with higher levels of green space. Second, we explore models to predict green space exposures based on either neighborhood or individual-level factors compared to models with multi-level predictors. We hypothesized that adding individual-level covariates to the model would improve prediction of green space measures.

Chapter 2 describes an analysis of green space and attention and externalizing behavior problems in 4 to 6-year-olds. Externalizing behaviors were assessed in the CANDLE cohort using the Child Behavior Checklist (CBCL). We examined both the broadband externalizing behavior score and the attention problems syndrome scale in primary analyses and the aggressive behavior syndrome scale in a secondary analysis. We explored adjustment of the primary model for factors potentially on the causal pathway or proxies for those causal pathways, including physical activity, screen time, sleep, and near-road residence, as well as associations with exposure during multiple windows. Effect modification was examined by neighborhood socioeconomic opportunity and child sex.

Chapter 3 examines the relationship between green space and internalizing behavior problems in early childhood. Our primary analysis utilized the broadband internalizing score from the CBCL. We investigated the related syndrome subscales—emotionally reactive, anxious and depressed, somatic complaints, and withdrawn behavior—in secondary analyses. As was done in chapter 2, we explored whether these associations varied by neighborhood SES and child sex, and conducted a number of sensitivity analyses to assess the robustness of modeled relationships.

A descriptive analysis of the green space measures used in this work—residential surrounding greenness, tree canopy, and park proximity—is included in Appendix A. This includes the primary metrics included for the current participant addresses in chapter 1 as well as the exposures averaged over multiple windows in chapters 2 and 3.

The CANDLE study in Memphis, Tennessee

The Conditions Affecting Neurocognitive Development and Learning in Early Childhood (CANDLE) study is a pregnancy cohort located in Shelby County, TN (48,49). This longitudinal, observational cohort is one of three cohort studies within the ECHO-PATHWAYS consortium (50). Pregnant women (N=1503) were enrolled in CANDLE between 2006 and 2011 (births in 2007-2011). These women were recruited from community obstetric practices and an urban hospital obstetric clinic in Memphis, TN. Enrollment included healthy women between the ages of 16 and 40 during weeks 16 through 27 of their pregnancy, who were able to speak and understand English and were a resident of Shelby County, TN. Women were excluded from participation in the CANDLE study if they were considered to have a high medical risk pregnancy, including those with an existing chronic disease requiring medication or a known pregnancy complication at the time of recruitment. Eligibility for CANDLE was further restricted to singleton pregnancies and women planning to deliver at one of the four participating study hospitals.

Among women who remained in the study through delivery, there were 1,436 live births. Mother-child dyads in CANDLE have completed follow-up study visits at multiple intervals, with current study visits ongoing in 2021. Data collection has included biospecimens, questionnaires, clinic visits, home visits, and frequent phone contacts for both data collection and study retention purposes. The CANDLE study has collected a complete residential history for each participant through more than fifteen points of contact with study staff, on the phone or in person, over the first 4-6 years of each child’s life. CANDLE staff have successfully maintained high rates of mother and child participation in the study over time; 1,157 mother-child dyads were seen at the age 4-6 study visit.

The CANDLE study was established specifically to investigate determinants of child neurodevelopment in a population representative of Shelby County, TN. Approximately 66% of the women enrolled in CANDLE are African-American/Black and 31% White, 63% are single mothers, about 30% have a bachelor’s degree or higher, and about 30% report a household income less than $20,000 a year. These distributions of participant characteristics are generally similar to the city of Memphis and Shelby County, TN as a whole (51). Based on 2010 American Community Survey estimates, Memphis is a city of 650,000 people with 7.6% of the population under 5 years of age (52). Sixty-four percent of residents are Black/African-American and 29% are White; 25% of adults older than 25 years have a bachelor’s degree or higher. The estimated median household income in Memphis is $38,230, with more than 25% of the population under 100% of the federal poverty level.

The CANDLE cohort is located in Memphis, within Shelby County in the southwestern corner of Tennessee. The region is classified as a warm temperate, fully humid, hot summer climate zone. While most of the annual precipitation occurs during winter and early spring, there is also typically significant precipitation during the summer, with average temperatures generally ranging from 34ºF to 91ºF across the year. This climate and the location of Memphis lends itself to a long growing season. Where natural vegetation has not been converted to agriculture, the region is primarily deciduous, mesophytic forests (53).

Chapter 1. The distribution of green space by individual and neighborhood-level characteristics in Shelby County, TN

Introduction

Access to green space may be distributed differentially across the population; ecologic studies in a variety of urban settings suggest that neighborhoods with lower median income, lower adult educational attainment, or a higher proportion of African-American residents may have less green space (54–59). However, these patterns are not consistent across all urban settings or for all types of green space. Much of the prior literature has also focused specifically on park spaces. In some US cities, low socioeconomic position has been associated with access to lower quality parks, but this finding is not consistent across cities or for park quantity versus quality (60–63). Most of the literature to date has focused on correlations between neighborhood characteristics and green space; however, individual-level characteristics may also be correlated with green space exposures (57). Additionally, no prior studies have included spatial measures of racial residential segregation. Given the array of health benefits hypothesized for access and exposure to green space, as well as the potential for confounding by other individual and neighborhood characteristics in epidemiologic studies, further investigation of how these factors predict green space in specific contexts is warranted.

In this study, we examined the relationships between individual and neighborhood characteristics and green space within the CANDLE cohort in Memphis, TN. Specifically, we investigated two distinct research questions using two different analytic approaches. First, we estimated the overall association between an index of neighborhood conditions and green space, and identified which neighborhood measures were driving those relationships. Second, we evaluated whether neighborhood-level, individual-level, or the combination of neighborhood and individual-level factors best predicted participant residential surrounding greenness, tree cover, or park access.

Methods

Study population

This analysis used participant-level data from the CANDLE study, one of three cohort studies within the ECHO-PATHWAYS consortium. CANDLE is a pregnancy cohort in Shelby County, TN (48,49). The CANDLE study was established specifically to investigate determinants of child neurodevelopment in a representative sample from Shelby County and includes a large proportion of participants who live in under-resourced communities. Pregnant women living in Shelby County, TN were enrolled in CANDLE between 2006 and 2011 (N=1503) and followed-up at a study visit when children were between 4 and 6 years old. The current address at the time of the age 4-6 follow-up visit was used to assess green space exposure and neighborhood characteristics. The sample was limited to those with a valid geocoded current address within Shelby County.

Green space

We used three measures of exposure to green space in this study: residential surrounding greenness, tree cover, and park proximity. Details on the data sources and data processing are included in Appendix A and described briefly here. Residential surrounding greenness was defined as the Normalized Difference Vegetation Index (NDVI). NDVI was calculated using NASA Global Web-Enabled Landsat Data (GWELD) to assess overall vegetation (64). These satellite images are available at a 30m resolution. Primary exposure measures were derived using the GWELD annual dataset for 2011, which was calculated as the highest NDVI in each pixel from December 1st, 2010 through November 30th, 2011. For this analysis pixels with a value less than zero, which indicates water, were set to missing. The weighted average of non-missing values was calculated within a 300m radial buffer of the current address of participants at the time of the age 4-6 study visit. The 300m buffer size was chosen as the primary spatial scale for this exposure because it has been hypothesized to be a relevant scale for psychological mechanisms of the effect of green space on health; prior studies of mental health outcomes have commonly used smaller buffers than physical health outcomes (3). Others have also cited distances of approximately 300m as a reasonable walking distance for families to enjoy nature and green spaces, and because this distance is commonly utilized in policy and programs as a measure of the residential neighborhood (65). Buffer sizes of 100m, 500m, and 1000m were examined in sensitivity analyses.

Tree cover data were obtained from the US Environmental Protection Agency EnviroAtlas Memphis Community Dataset (66). This metric was derived from 1m resolution land cover data and reported as percent of each census block group covered by tree canopy. Exposures were linked to participant addresses by the census block group identifier for the current participant address at the time of the age 4-6 study visit.

The third green space measure used was the distance in meters from the participant address at the age 4-6 study visit to the edge of the nearest public park. Park boundaries were derived from ParkServe data compiled by the Trust for Public Land (67). This dataset includes publicly owned local, state, and national parks; school parks with a joint-use agreement with local government; and privately owned parks that are managed for full public use.

Neighborhood characteristics

Neighborhood conditions were conceptualized as both socioeconomic resources and educational resources in the neighborhood, using indicators from the Childhood Opportunity Index (COI). The COI was developed for the purposes of examining neighborhood resources and conditions that influence children's health and development. and has been found to be associated with multiple health outcomes (68,69). The COI was compiled from multiple data sources in order to capture a variety of potential resources including the availability and quality of neighborhood institutions and neighborhood social structure and economic resources (70,71). The specific variables within each domain and subdomain are defined in Table 1-1. The socioeconomic domain includes variables related to economic opportunities and to economic and social resources. The educational opportunity domain includes variables related to early childhood education, elementary education, secondary and postsecondary education, and educational and social resources. Data at the census tract level in 2010 were used to calculate z-scores for each variable.

The environmental justice literature has highlighted the role of racial residential segregation in determining exposure to environmental factors in the US. However, most studies have relied solely on racial composition measures which do not account for levels of segregation in surrounding neighborhoods. In this study, we used measures of both racial composition and a spatial measure of racial residential segregation. Racial composition measures (% Black and % White) were obtained at the census tract level from the 2006-2011 American Community Survey (ACS). To assess spatial patterns of segregation in the city, we used the Getis-Ord G_i* statistic (72–74). Though multiple approaches to assessing the different dimensions of racial residential segregation have been used in studies of segregation and health, the clustering dimension assessed by the G-statistic has been theorized to be a relevant measure for understanding historical patterns of land-use decision-making and community investment in physical infrastructure (75). This statistic describes whether the racial composition within the census tract and in the surrounding tracts, deviates from the overall county mean. Surrounding tracts were defined as those that share a boundary with the index tract. The G-statistic was calculated using the racial composition data from the 2006-2011 ACS. In this study, a higher, positive value of the G-statistic z-score indicates an overrepresentation of Black residents within the census tract and surrounding tracts compared to Shelby County as a whole.

Population density at the census tract level, derived from the 2006-2011 ACS, was used as a measure of urbanicity. All neighborhood-level predictors were assessed for the participant address at the time of the age 4-6 year study visit.

Individual and household characteristics

Individual and household-level characteristics were assessed at the age 4-6 study visit and include maternal age, race, education, and marital status; and household income and household size. Maternal race was categorized as African-American/Black, White, or other. Maternal education was categorized as less than a high school degree, high school diploma or equivalent, technical school, college degree, and graduate or professional degree. Marital status was categorized in two groups: married or living with a partner and single, which included never married, widowed, divorced, and separated. Household income was reported in 11 categories: $0-4,999; $5,000-9,999; $10,000-14,999; $15,000-19,999; $20,000-24,999; $25,000-34,999; $35,000-44,999; $45,000-55,999; $55,000-64,999; $65,000-74,999; or $75,000 and over. The midpoint of each category was selected and treated as a continuous variable. Those in the highest income category were assigned to $80,000. Continuous variables were standardized to a mean of zero and standard deviation of one.

Statistical analyses

Descriptive statistics and Spearman correlations were used to examine the distribution of and correlations between the green space measures and each of the predictors.

<u>WQS Regression</u>

Weighted Quantile Sum (WQS) regression (76,77) was used to estimate the association between a suite of neighborhood conditions and green space, and to examine which predictors were driving those associations. Each predictor of interest was divided into quintiles. The first step in WQS regression is a bootstrap sampling procedure to estimate weights for the WQS index from the predictors of interest. The second step is to estimate the coefficient for the resulting WQS index. The WQS index is defined as $WQS = \sum_{j=1}^{c} \overline{w}_j q_i$ where w_j is the weight ($0 \leq w_j \leq 1$) for the *j* variable with quintile values q_j, and the weights sum to 1. The final weights are calculated through bootstrap sampling (B=1000 samples) with replacement as: $\overline{w} = \frac{1}{B}\sum_{b=1}^{B} w_{j(b)} f(\hat{\beta}_{1(b)})$, where $f(\hat{\beta}_{1(b)})$ is a pre-specified signal function, which uses the t-statistic for the β_1 estimate from each bootstrap sample to calculate a weighted mean across bootstrap samples such that samples with a higher signal are given a higher relative weight in calculating the index. Two separate models were calculated by constraining the association between the WQS index and the green space measure to be either positive or negative. We hypothesized that a higher WQS index, representing more neighborhood resources, would be associated with higher levels of green space. Our primary analysis examined associations in a model constrained to assess an association in this hypothesized direction. As a secondary analysis, we also examined a WQS index in the opposite direction of that hypothesized, where the model was constrained to non-positive beta estimates.

The WQS index was then included in an ordinary least squares regression model. For each measure of green space as the dependent variable, the coefficient β_1 for the WQS index was estimated, adjusting for urbanicity. The WQS regression yields a single coefficient to describe the relationship between neighborhood characteristics and residential green space. We adjusted for population density as a measure of urbanicity due to the geographic extent of Shelby County, which includes census tracts across an urban to rural gradient. In the primary analysis where the model was constrained to be non-negative, the WQS index was calculated for the associations between a higher index and either more surrounding greenness and tree canopy or closer proximity to green space. The weights in the WQS index were examined to identify the factors driving observed associations. In a secondary analysis, we also examined model results when the WQS index was calculated for associations that were constrained to be in the opposite of the hypothesized direction.

WQS regression was developed in the context of studies of chemical mixtures, but has been extended to model socioeconomic indices (78,79). The neighborhood characteristics contributing to the WQS index in this analysis are poverty rate, public assistance rate, homeownership rate, high-skill employment, median household income, employment rate, commute duration, percent single-headed households, school poverty, teacher experience, adult educational attainment, early childhood education (ECE) centers, high-quality ECE centers, ECE enrollment, third grade math proficiency, third grade reading proficiency, high school (HS) graduation rate, Advanced Placement course enrollment, college enrollment in nearby institutions, percent Black/African-American, and the G-statistic. Each of the variables described here, across multiple domains, were included separately within the WQS index. Variables were reverse-scored as necessary, so that all variables were coded such that higher quintiles were hypothesized to be associated with more green space or represented more neighborhood resources.

LASSO Regression

LASSO regression was used to compare predictive models of green space using neighborhood-only, individual-only, or multi-level sets of predictors. Model A for each green space measure contained only neighborhood-level predictors. The neighborhood variables in these LASSO models included 21 neighborhood characteristics, including measures of socioeconomic and education resources, and racial composition and residential segregation. Model B contained only individual-level predictors: maternal race, maternal education, marital status, maternal age, household income, and household size at the age 4-6 study visit. Model C contained all predictors, including both individual and neighborhood-level variables. All continuous variables were standardized to have a mean of 0 and standard deviation of 1. For each LASSO model, 10-fold cross-validation was used to identify the minimum point in the estimated test mean squared error (MSE) curve and select the corresponding tuning parameter (λ). Coefficients were then estimated in the full sample using the selected tuning parameter. The minimum MSE from 10-fold cross-validation was compared across models A-C to assess the predictive ability of the three sets of variables.

All data manipulations, visualizations, and statistical analyses were conducted in R 3.6 (The R Foundation for Statistical Computing; Vienna, Austria). Processing of spatial data was implemented using the sf and raster packages, WQS analyses were conducted using the gWQS package, and LASSO analyses were conducted using the glmnet package.

Results

CANDLE participants were included in this analysis if they had a valid geocoded address in Shelby County, TN at the time of the age 4-6 study visit (n=1012). LASSO models were further restricted to those with complete data for individual-level predictors at the age 4-6 visit (n=917).

The distribution of residential surrounding greenness, tree canopy, and park access in this cohort are shown in Table 1-2. NDVI levels were generally high, on average 0.596 (SD 0.084) in 300m buffers. Tree canopy (mean 37.8% [SD 12.5]) at the census block group level was correlated with NDVI in 300m buffers at 0.60. Distance to the nearest park was on average 795m (SD 862) and 28% of the cohort lived within 300m of the nearest park; park proximity was not highly correlated with the other green space measures (Spearman correlations of -0.03 with NDVI and 0.02 with tree canopy). A more detailed description of the various green space metrics explored, as well as results from descriptive analyses and maps of these various measures, are included in Appendix A.

Additional characteristics of the neighborhood environment, including social and economic resources, educational resources, racial composition, and residential segregation, are shown in Table 1-3. Relative to other US cities, neighborhoods in Memphis on average tend to have fewer neighborhood-level resources as assessed by the various components of the Childhood Opportunity Index (70). The mean percent of students in elementary schools eligible for free or reduced-price lunches is 71% for the neighborhoods in which CANDLE participants live, compared to the national average of 53%. In this cohort, the average neighborhood poverty rate (percent of individuals living in households below 100% of the federal poverty threshold) was 23%, compared to the national average of 15%. However, the distribution of other characteristics was similar to the national average. For example, the mean high school graduation rate is 76% in CANDLE participant neighborhoods, compared to the national average of 75%. Prior work using the COI has also identified larger gaps in Memphis between neighborhoods with many resources and those with few resources across multiple domains, compared with other US cities (70). We observed some variability in z-scores of all of these measures within the CANDLE cohort as well (Figure 1-1). Spearman correlations between the neighborhood variables ranged from 0.01 to 0.9, with higher correlations between variables within the same conceptual domain (Figure 1-2). Bivariate correlations of neighborhood characteristics with green space measures were generally low.

Measures of racial residential segregation and racial composition are also included in Table 1-2. Census tracts in which CANDLE participants lived were on average 57% African-American/Black. CANDLE participants on average also tended to live in neighborhoods with a higher dissimilarity measure of residential segregation. Both percent black and the residential segregation measure were generally negatively correlated with socioeconomic and education neighborhood opportunity measures (Figure 1-2).

The distributions of individual and household-level characteristics for CANDLE participants in Shelby County at the time of the age 4-6 study visit (N=1012) are shown in Table 1-4. In this sample of mother-child dyads, 63% of women were African-American/Black and 30% White. The mean annual household income was $37,691 and 53% of participants had at least a technical school or college degree.

WQS regression

Results from the WQS regression for NDVI and tree canopy, when the association was constrained to estimate the association between higher opportunity (higher WQS index) and more green space are shown in Figure 1-3. A 1 quintile increase in the WQS index was associated with a 0.021 unit higher NDVI (95% CI: 0.014, 0.028) and this association was largely driven by higher homeownership rate, closer proximity of early childhood education (ECE) centers and higher enrollment, and lower percent African-American/Black. A 1 quintile higher WQS index was associated with a 4.9 percent higher tree canopy coverage (95% CI: 3.8, 6.0) and this index appears similar to that for NDVI, with the same four variables having the highest weights. A 1 quintile higher WQS index was associated with living 358 m closer to a park (95% CI: -427, -288) and the early childhood education variables were weighted heavily, along with teacher experience.

Results from the WQS regression when the association was constrained to estimate the association between a higher WQS index and a lower level of green space or further distance from a park are shown in Figure 1-4. A 1 quintile higher WQS index was associated with 0.028 lower NDVI (95% CI: -0.036, -0.021) when HS graduation rate and percent new teachers (reverse coded) were highly weighted. A 1 quintile higher WQS index was associated with 1.9 lower percent tree canopy (95% CI: -3.0, -0.9), when HS graduation rate was heavily weighted. A 1 quintile higher WQS index was associated with a 184m further distance to the nearest park (95% CI: 148, 221) when 3rd grade reading proficiency and percent single-headed households (reverse coded) were highly weighted.

In sensitivity analyses using NDVI calculated in varying buffer sizes around the residence as the dependent variable, a larger magnitude of association was observed for NDVI in the 100m buffer size relative to the larger buffer sizes (Figure 1-5). Homeownership rate was consistently identified as the variable driving this association, with weights ranging from 0.32 to 0.49, along with ECE-related variables.

LASSO regression

For models predicting NDVI, including both individual and neighborhood characteristics in the model (model C) only slightly reduced the MSE compared to models A and B, suggesting that adding the individual-level predictors does not meaningfully improve prediction accuracy. Adding individual-level predictors to a model that already includes the suite of neighborhood predictors did not improve prediction of green space for tree canopy or park proximity. In prediction models for each of the three types of green space, the MSE was highest when only individual-level covariates were included in the model (Table 1-5).

Coefficients from these models using the tuning parameters identified from cross-validation are shown in Tables 1-6, 1-7, and 1-8, for models A, B, and C, respectively. For models of NDVI and tree canopy, positive coefficients indicate predictors of more green space. Lower percent Black in the census tract and higher homeownership rate predicted higher NDVI (Table 1-6). In this model, lower poverty rate and lower college enrollment nearby predicted lower NDVI. In the model with individual predictors (Table 1-7), marital status (married or living with a partner) and maternal race (Black/African American) predicted NDVI. Similar patterns were observed for neighborhood predictors when both neighborhood and individual predictors included in the same model (Table 1-8). More coefficients for individual-level variables were estimated to be non-zero in this combined model, though the magnitude of these coefficients was small relative to those for neighborhood variables.

In the model of neighborhood characteristics and tree canopy, the largest coefficient was for the residential segregation index, with a higher value of the index indicating more residential segregation of Blacks and predicting lower tree cover (Table 1-6). Similar to NDVI, a lower poverty rate and higher college enrollment in nearby institutions predicted lower tree canopy coverage in this model. In a model with individual-level predictors only, the largest coefficient was for the maternal education of a graduate or professional degree (Table 1-7).

Park access was operationalized as distance to the nearest park; in contrast to the NDVI and tree canopy measures, smaller values of the park proximity variable are interpreted as greater access to green space. Third grade reading proficiency predicted closer park proximity, but third grade math proficiency predicted a further distance from the nearest park, in neighborhood models (Table 1-6). ECE centers predicted closer proximity to a park, but high-quality ECE centers predicted a larger distance to the nearest park. Higher median household income predicted living closer to a park. In model B with only individual-level predictors, only two factors remained in the model (Table 1-7). The maternal education category of high school degree and household income both predicted further distance to the nearest park. In contrast to the individual-level model, almost all coefficients including those at the individual level remained in the combined model (Table 1-8). Trends in the combined model for early childhood and elementary education variables were similar to the neighborhood-only model.

Discussion

We used two distinct modeling approaches to explore relationships between green space and a range of socioeconomic, education, and racial composition and residential segregation measures of the neighborhood. In WQS regression, associations between the neighborhood social, economic, and educational environment and higher measures of greenness were driven by factors within both the education and socioeconomic domains. Homeownership rate was consistently the most highly weighted neighborhood variable in the WQS index. The early education variables had smaller weights for each variable, but collectively suggested the importance of the early educational environment. Homeownership rate and ECE variables were similarly important predictors of green space in LASSO regression models using neighborhood-level predictors. Adding individual-level predictors to the LASSO regression models improved prediction of NDVI only slightly and did not improve prediction of tree canopy or park proximity. However, individual-level predictors were still frequently included in these combined models.

While some variables were influential in both modeling approaches, we also observed some differences across the WQS and LASSO approaches. One explanation for this may be that these two modeling approaches handle correlated predictors differently. The WQS model tends to split weights across the group of variables that are highly correlated, while the LASSO model selects a single predictor out of a group of correlated predictors (77).

Much of the prior literature examining the distribution of green space has focused on measures of income and poverty in relation to urban green space quantity or proximity. Single-city studies in Europe and a multi-city study in Australia identified better access to public green spaces in neighborhoods characterized by higher incomes (54,55,57). In the US, a national study of census tracts from the year 2000 identified lower levels of greenness in areas with a higher concentration of poverty (80). We did not observe consistent results across measures of poverty and income. Median household income was less influential than housing tenure in WQS models. Others have also observed associations with housing tenure and suggested that renters may have less ability or incentive to influence green infrastructure (81,82). Homeownership is also an indicator of overall wealth, particularly among those with lower incomes in the US (83).

Fewer studies have considered education in addition to income in models of green space. A study of census tracts across ten US cities identified more consistent relationships between education and greenness than income (56). The ten cities included spanned multiple eco-regions in the US, with variable urban development structures. Adult educational attainment tended to have the strongest association with greenness and woody vegetation. However, income had a larger role in cities with relatively lower per capita incomes. In the smallest city included in this study—Jacksonville, FL—the authors observed only weak relationships between measures of income, education, and urban vegetation. In our study, adult educational attainment was one of the top predictors of tree canopy when using LASSO regression, but was not weighted heavily in WQS analyses.

This hypothesis-generating analysis explored a wider selection of neighborhood resources than prior work. In addition to multiple indicators of socioeconomic conditions of the neighborhood, we explored a range of neighborhood indicators related to educational opportunities. The weights for the proximity of early childhood education centers and elementary school test scores in WQS analyses suggest a relationship with the early education environment. Given the high cost of early childhood education in the US, disparities in access to these education opportunities indicated by variables such as the percentage of 3 and 4-year-olds enrolled in preschool, likely reflect disparities in access to a broader set of resources for families with young children (84). Some of these measures may reflect the population density of young children, which may vary relative to the total population density that we controlled for in this analysis, and has been correlated with measures of urban trees in other studies (85). School grounds may also contribute to the quantity of green space in the neighborhood, particularly in high-income neighborhoods (86).

Several studies have observed less green space access in neighborhoods with a higher percentage of African-American residents, including in Atlanta, GA and Baltimore, MD (58,60). We observed a similar pattern of total vegetation by neighborhood racial composition in Memphis. A history of discriminatory policies shaping investment in neighborhoods have influenced spatial patterns; studies of historical redlining practices found that worse Home Owners Loan Corporation (HOLC) grades in the 1930s were associated with decreased present-day greenness and tree canopy (87,88). In our analysis, higher residential segregation modeled using the Getis-Ord G statistic was the largest neighborhood coefficient predicting less tree canopy in LASSO models. Others have hypothesized that tree canopy in particular is more reflective of long-term investments in green infrastructure, due to the time it takes for large trees to grow, relative to overall measures of greenness or measures of herbaceous vegetation (81).

Patterns observed across multiple domains of neighborhood conditions in relation to urban green space may also vary depending on the type of green space considered. Some studies have observed consistent patterns of tree canopy by income (59); others observed the largest differences in street trees by race and education (85). Prior studies have also observed variability in these patterns of tree canopy across cities in the US (81,89). A study of 40 US cities found that the distribution of urban tree canopy was less dependent on socioeconomic factors in cities with higher levels of tree canopy relative to cities with lower tree canopy overall (89). The authors suggest that in cities with lower overall tree canopy, trees tend to be more concentrated within specific parcels of land such as within parks. Participants in the CANDLE cohort generally had higher levels of residential surrounding greenness than in many prior studies (mean of 38% tree canopy). The climate in Shelby County is conducive to widespread vegetation growth and deciduous tree cover. Average levels of tree canopy were similar in prior studies of Charlotte, NC (47%) and Pittsburgh, PA (38%) (81,89). Differences across cities in multi-city studies suggest that the generalizability of our study may be limited and that extrapolation to other urban settings with lower levels of urban tree canopy should be done cautiously.

Simple measures of park access tend to not have the same associations with socioeconomic conditions that have more consistently been observed for measures of total vegetation or urban tree canopy (90). A study in Denver, CO found that low-income neighborhoods had better park access, but these patterns are not observed for acres of parks per youth, high quality parks, or safer parks (91). Similarly, a study of five other US cities found that access to safe parks, was more limited in low-income communities (92). In our analysis, patterns of park access by income and education were inconsistent and appeared sensitive to modeling choices. However, an important limitation of this analysis was that we were unable to refine our assessment to account for a measure of park facilities such as playgrounds or other qualities of the parks.

The distributions of green space measures may also vary based on the size or spatial scale of the urban green space (93). In our study, NDVI was assessed on a smaller scale than tree canopy; NDVI was also the measure for which adding individual-level covariates to the LASSO models slightly improved prediction over the model only including neighborhood-level covariates. In the WQS analysis, we examined NDVI at multiple distances from the residential location, but observed similar trends across all buffer sizes.

There are several limitations that should be considered in interpretation of these analyses. The variability in overall greenness is restricted, as Memphis is a generally green city. The location and climate of Memphis is a conducive growing environment and patterns observed here may not be generalizable to other regions. We used census tracts to define neighborhoods when assessing neighborhood resources and characteristics, as is commonly done in the literature, but these administrative boundaries may not accurately reflect the way that residents would define their neighborhood or interact with the spaces surrounding their residence. The spatial mismatch between the neighborhood census tract variables and NDVI measures, which were measured within buffers around the residential locations irrespective of census geographical borders, is an additional limitation in this analysis. Given these limitations, the hypothesis-generating analyses presented here should be interpreted cautiously.

This analysis is distinct from prior studies in that we used measures of green space based on participant residential locations from a cohort study of mother-child dyads. Prior studies have primarily examined greenness on a larger spatial scale than the 300m buffers used in our primary analysis (94). The results in this aim suggest several factors that may be considered as confounders when investigating relationships between green space and health. In particular, confounding by homeownership and early childhood education opportunities may be of concern for health outcomes known to be affected by these neighborhood resources. Additionally, the inclusion of a host of characteristics with variable distributions across the population, and by exploring relationships with multiple measures of green space exposure and access, this exploratory analysis suggests several neighborhood resources for further investigation in future work.

Tables & Figures

Table 1-1. Variable definitions for components within the socioeconomic and education domains, obtained from the Childhood Opportunity Index (70).

Subdomain	Variable Name	Variable Definition and Units
Social and Economic Domain		
Economic and social resources	Poverty rate	Percentage of individuals living in households with incomes below 100% of federal poverty threshold (reversed coded in z-score)
	Public assistance rate	Percentage of households receiving cash public assistance or Food Stamps/Supplemental Nutrition Assistance Program (reverse coded in z-score)
	Homeownership rate	Percentage owner-occupied housing units
	High-skill employment	Percentage of individuals ages 16 and over employed in management, business, financial, computer, engineering, science, education, legal, community service, health care practitioner, health technology, arts and media occupations
	Median household income	Median income of all households
	Single-headed households	Percent family households that are single-parent headed (reverse coded in z-score)
Economic opportunities	Employment rate	Percent adults ages 25-54 who are employed
	Commute duration	Percent workers commuting more than one hour one way (reverse coded in z-score)
Education Domain		
Educational and social resources	School poverty	Percentage students in elementary schools eligible for free or reduced-price lunches (reverse coded in z-score)
	Teacher experience	Percentage of teachers in their first and second year (reverse coded in z-score)
	Adult educational attainment	Percentage adults ages 25 and over with a college degree or higher
Early childhood education (ECE)	ECE centers	Number of ECE centers within a 5-mile radius, converted to natural log units
	High-quality ECE centers	Number of NAEYC accredited centers within a 5-mile radius, converted to natural log units
	ECE enrollment	Percentage of 3 and 4 year olds enrolled in nursery school, preschool, or kindergarten
Elementary education	Third grade math proficiency	Percentage of third graders scoring proficient on standardized math tests, converted to NAEP scale score points
	Third grade reading proficiency	Percentage of third graders scoring proficient on standardized reading tests, converted to NAEP scale score points
Secondary and postsecondary education	High school graduation rate	Percentage of ninth graders graduating from high school on time
	Advanced Placement course enrollment	ratio of students enrolled in at least one AP course to the number of 11th and 12th graders
	College enrollment in nearby institutions	Percentage of 18-24 year-olds enrolled in college within 25 mile radius

Table 1-2. Distribution of green space measures in the CANDLE cohort for the residential address reported at the time of the age 4-6 year study visit (n=1012).

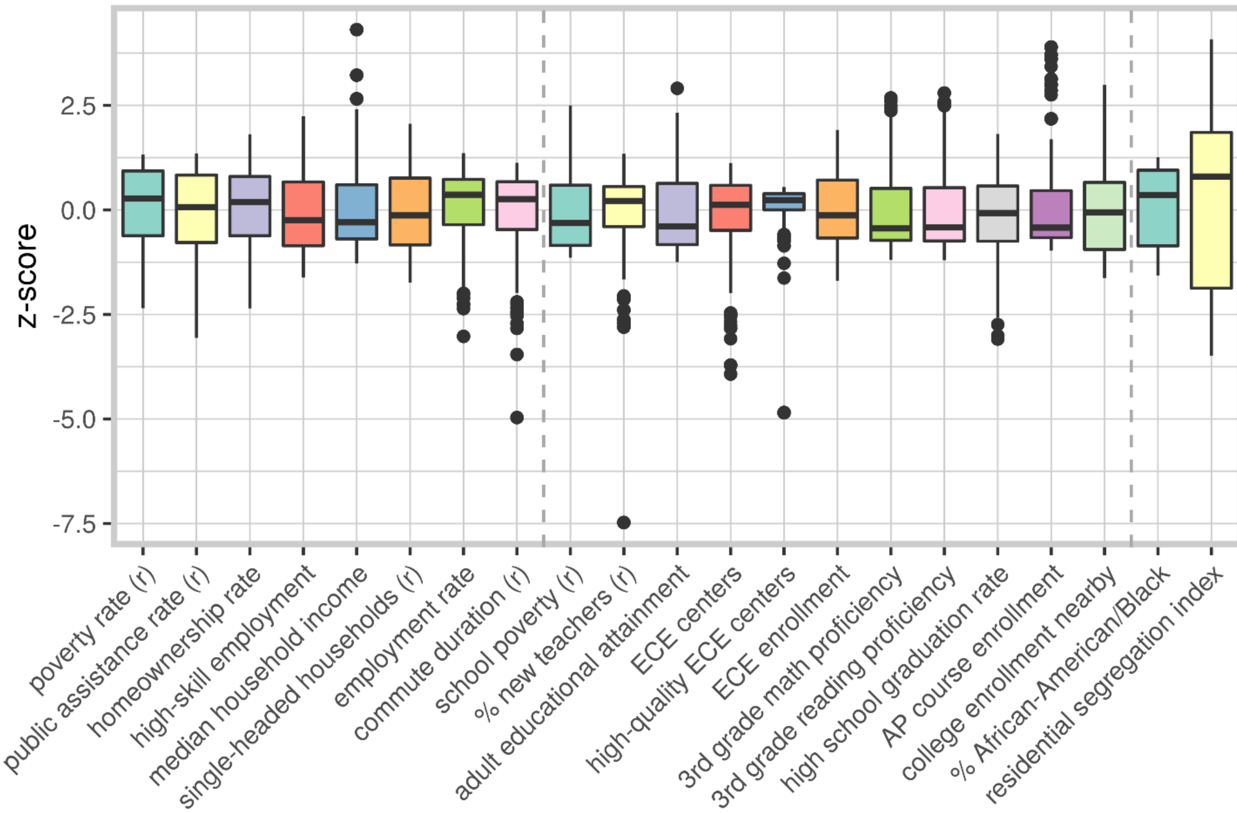
Green space	Mean	SD	Min.	Q1	Median	Q3	Max.
NDVI in 300m buffer	0.596	0.084	0.251	0.544	0.599	0.657	0.810
Tree canopy (%)	37.8	12.5	4.4	28.3	37.8	46.8	81.2
Park proximity (m)	795	862	4	282	549	987	8417

Table 1-3. Characteristics of CANDLE participant neighborhoods, including socioeconomic and education environments and racial composition and residential segregation measures.

Neighborhood characteristics at age 4-6 year study visit [a]	Mean (SD)
COI Socioeconomic Domain [b]	
Poverty rate (% households)	22.9 (16.3)
Public assistance rate (% households)	23.5 (16.4)
Homeownership rate (%)	58.1 (21.6)
High-skill employment (%)	29.5 (16.1)
Median household income ($)	50853 (29184)
Single-headed households (%)	55.0 (25.1)
Employment rate (%)	71.8 (12.6)
Commute duration (% commuting >1hr one way)	3.1 (2.5)
COI Education Domain [b]	
School poverty (% of students)	71.1 (25.0)
Teacher experience (% in 1st or 2nd year)	7.5 (5.6)
Adult educational attainment (% w/ college degree)	24.2 (18.6)
Early childhood education centers (n in 5-miles)	4.9 (0.7)
High-quality early childhood education centers (n)	1.1 (2.9)
Early childhood education enrollment (%)	45.9 (24.8)
Third grade math proficiency (NAEP scale score points)	157.4 (72.8)
Third grade reading proficiency (NAEP scale score points)	145.5 (77.9)
High school graduation rate (%)	76.1 (9.9)
Advanced Placement course enrollment (ratio)	0.12 (0.09)
College enrollment in nearby institutions (% in 25 miles)	36.0 (2.6)
Racial composition and residential segregation [c]	
% Black/African-American	57.0 (33.9)
% White	36.9 (32.1)
Residential segregation G-statistic	0.13 (1.87)

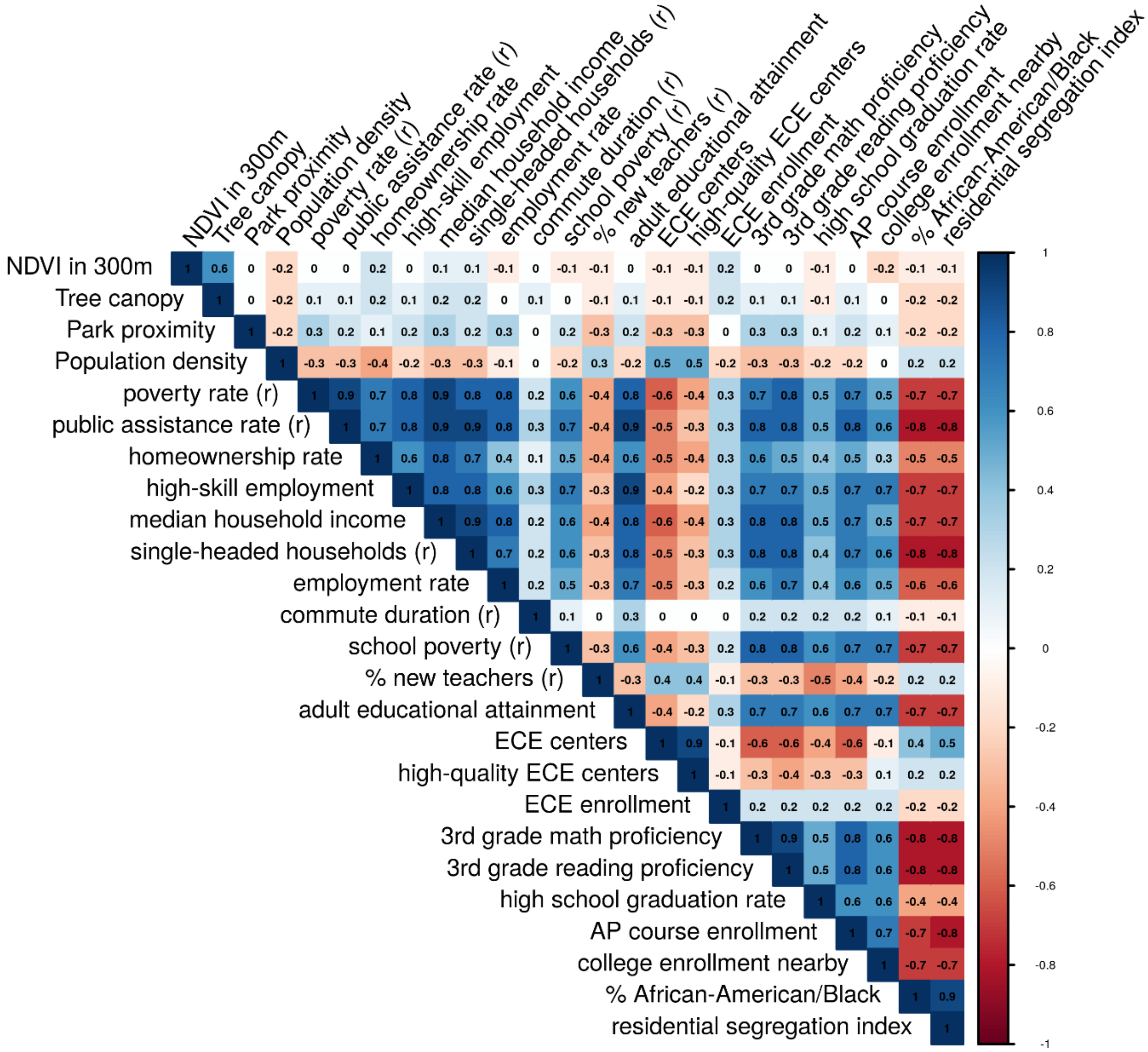
[a]Neighborhood is defined as the census tract of the address reported by participants at the time of the age 4-6 study visit.
[b]Variables in the education and socioeconomic domains were obtained from the Childhood Opportunity Index (COI).
[c]Data on race at the census tract level was obtained from the American Community Survey 2006-2011. Residential segregation calculated as the Getis-Ord G-statistic.

Figure 1-1. Distribution of neighborhood characteristics in the CANDLE cohort.



Boxplots show z-scores of neighborhood-level measures normed to Shelby County, TN for CANDLE participants in Shelby County at the age 4-6 visit (N=1012). Medians are shown by the horizontal black bar, interquartile range (IQR) is indicated by colored boxes, whiskers indicate either the minimum or maximum of the data or 1.5 times the IQR whichever is closer to the median, and circles indicate data outside of 1.5 times the IQR, for z-scores of neighborhood predictors for CANDLE participants (N=1012). Variables that were reverse coded are indicated by (r) in the label.

Figure 1-2. Spearman correlations among different domains of the neighborhood environment.



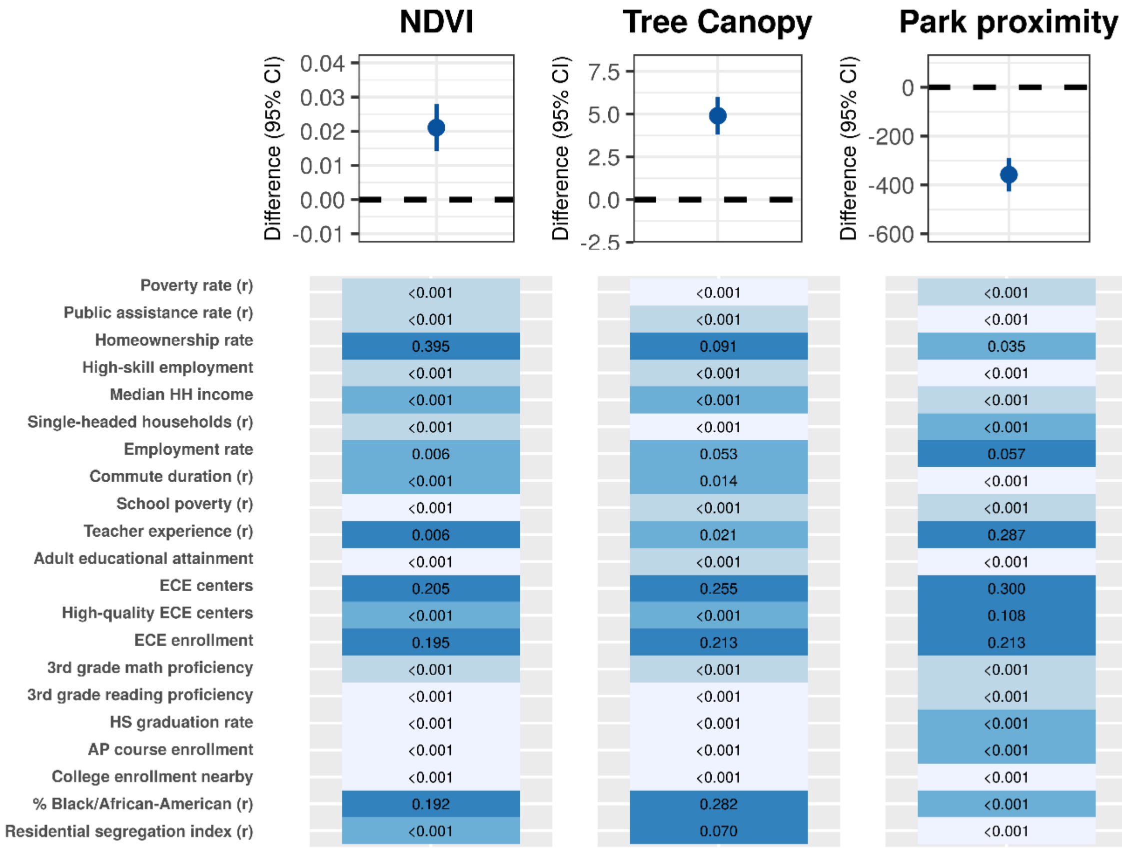
Variables derived from the Childhood Opportunity Index are coded such that higher values represent more opportunity; variables that were reverse coded are indicated as (r). Green space measures are included for the current address at the age 4-6 study visit.

Table 1-4. Individual and household characteristics of CANDLE participants (n=1012) at the age 4-6 year study visit.

Participant characteristics at age 4-6 year study visit	**Mean (SD) or N (%)** [a]
Maternal age (years), mean (SD)	31.4 (5.4)
Maternal race, n (%)	
Black/African-American	621 (63)
White	298 (30)
Other	63 (6)
Maternal education, n (%)	
Less than high school	48 (5)
GED or high school diploma	401 (41)
Technical school	121 (12)
College degree	246 (25)
Graduate or professional degree	159 (16)
Marital status, n (%)	
Married/living as married	562 (58)
Single/living as single	415 (42)
Annual household income ($), mean (SD)	37691 (27858)
Household size, mean (SD)	4.53 (1.42)

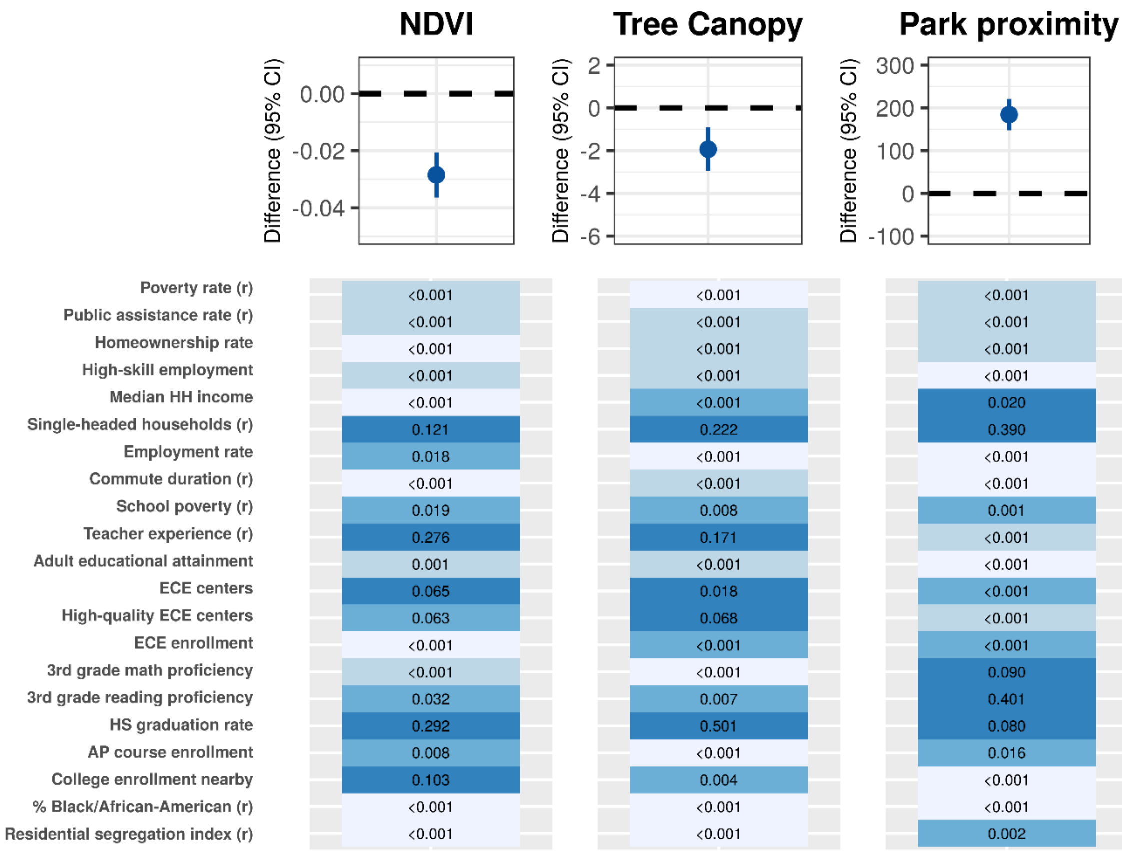
[a]Missing maternal age (11), maternal education (7), marital status (5), household income (42) and household size (65).

Figure 1-3. Difference (95% confidence interval) in surrounding greenness, tree canopy, or park proximity per 1 quintile higher WQS index and the corresponding weights for each variable in the WQS index.



The darker shades of blue indicate higher weights in the WQS index, with the darkest shade indicating the highest quartile of weights. The first column shows weights in the model of NDVI in a 300m buffer, the second column shows weights in the model of tree canopy, and the third column shows weights in the model of park proximity. Some variables, indicated by (r), were reverse coded before being considered in the WQS index so that variables contributing to the WQS index were all coded such that a higher value of the variable represents more opportunity (educational opportunity, socioeconomic opportunity) or was hypothesized to be associated with more green space. A 1 quintile increase in the WQS index was associated with a 0.021 unit higher NDVI (95% CI: 0.014, 0.028) and this association was largely driven by homeownership rate, ECE centers and enrollment, and racial composition. A 1 quintile higher WQS index was associated with a 4.9 percent higher of tree canopy coverage (95% CI: 3.8, 6.0) and this index appears similar to that for NDVI, with % black (reverse coded), ECE centers and enrollment, and homeownership rate weighted most heavily. A 1 quintile higher WQS index was associated with living 358 m closer to a park (95% CI: -427, -288) and the early childhood education variables were weighted heavily, along with teacher experience.

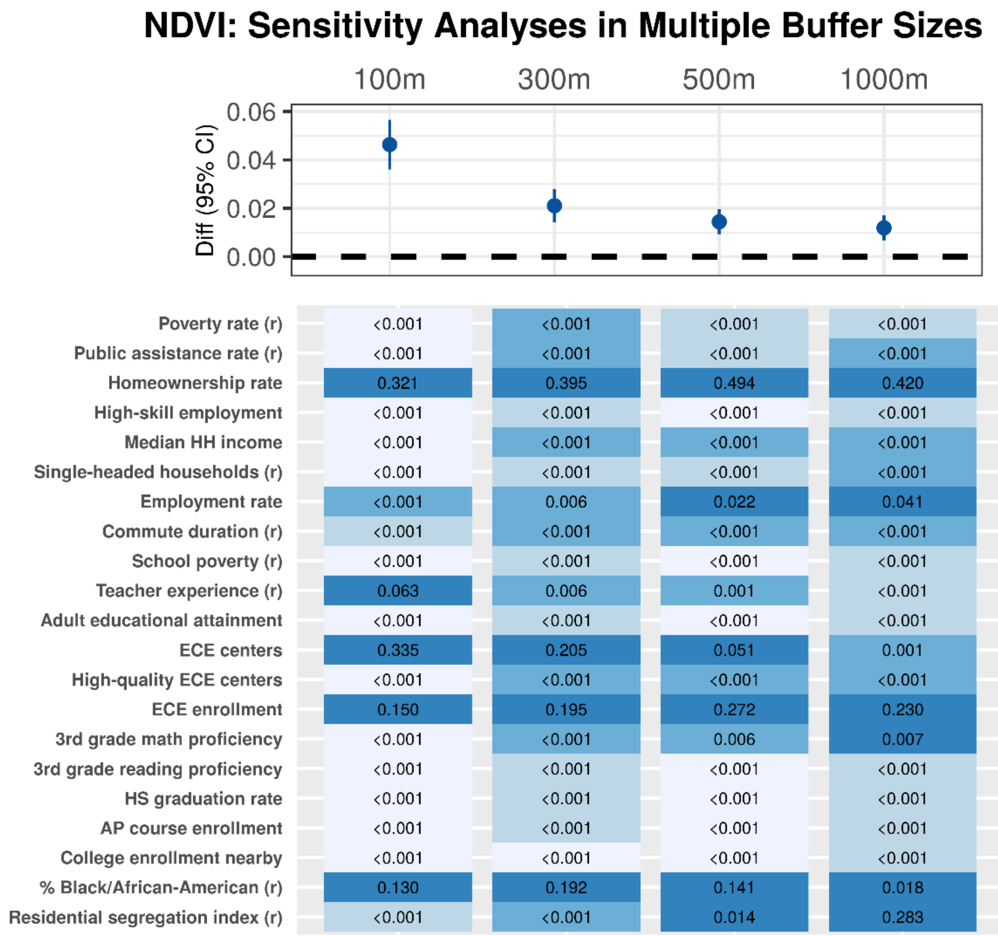
Figure 1-4. Difference (95% confidence interval) in surrounding greenspace per 1 quintile higher WQS index in a model where the association is constrained such that higher opportunity is associated with lower surrounding greenness or further distance to a park.



The darker shades of blue indicate higher weights in the WQS index, with the first column showing weights in the model of NDVI in a 300m buffer, the second column showing weights in the model of tree canopy, and the third column showing weights in a model of distance to the nearest park. Some variables in the WQS index were reverse coded, indicated by (r), so that variables contributing to the WQS index were coded such that a higher value of the variable represents more opportunity (educational opportunity, socioeconomic opportunity) or was hypothesized to be associated with more green space. A 1 quintile higher WQS index was associated with 0.028 lower NDVI (95% CI: -0.036, -0.021) when high school (HS) graduation rate (reverse coded) and teacher experience (reverse coded) were highly weighted. A 1 quintile higher WQS index was associated with 1.9 lower percent tree canopy (95% CI: -3.0, -0.9), when

HS graduation rate (reverse coded) was heavily weighted. A 1 quintile higher WQS index was associated with a 184m further distance to the nearest park (95% CI: 148, 221) when 3rd grade reading proficiency (reverse coded) and single-headed households (reverse coded) were highly weighted.

Figure 1-5. Difference (95% confidence interval) in surrounding NDVI in various buffer sizes per 1 quintile higher WQS index and corresponding weights.



The darker shades of blue indicate higher weights in the WQS index. The buffer size of the NDVI measure for each column of results is indicated at the top of each column of results. Some variables were reverse coded, indicated by (r), so that variables contributing to the WQS index were coded such that a higher value of the variable represents more opportunity (educational opportunity, socioeconomic opportunity) or was hypothesized to be associated with more green space. Associations were largely driven by homeownership rate across all buffer sizes.

Table 1-5. Minimum MSE from cross-validation of LASSO models of NDVI, tree canopy, and park proximity.

	NDVI	Tree canopy	Park proximity
(A) Only neighborhood-level predictors [a]	0.005995	123.2869	511426.9
(B) Only individual-level predictors [b]	0.007085	153.6273	682977.9
(C) Both individual and neighborhood predictors [c]	0.005973	125.4239	534456.0

[a] Model includes the following measures at the census tract level: poverty rate, public assistance rate, homeownership rate, high-skill employment, median household income, employment rate, commute duration, single-headed households, school poverty, teacher experience, adult educational attainment, ECE centers, high-quality ECE centers, ECE enrollment, 3rd grade math proficiency, 3rd grade reading proficiency, high school graduation rate, Advanced Placement course enrollment, college enrollment in nearby institutions, racial composition (% Black/African-American) and racial residential segregation (Getis-Ord Gi* statistic).

[b] Model includes the following measures, reported at the CANDLE age 4-6 study visit: maternal education (categorized as <high school, high school degree, technical school, college degree, graduate or professional degree), maternal race (categorized as Black/African-American or other), annual household income, household size, marital status (married/living with partner or not), and maternal age.

[c] Model includes all predictors specified for models A and B.

Table 1-6. Standardized coefficients from the LASSO models of NDVI, tree canopy, and distance to the nearest park, including only neighborhood-level predictors. [a]

	NDVI [b]	**Tree Canopy** [b]	**Park Proximity** [c]
Neighborhood-level predictor			
Poverty rate (r)	-0.024	-4.91	179.6
Public assistance rate (r)	-0.007	-0.54	29.3
Homeownership rate	0.028	3.07	-83.7
High-skill employment	-0.002	0.24	28.4
Median household income	0.004	0.64	-263.3
Employment rate			-95.7
Commute duration (r)	0.004	0.74	-26.3
Single-headed households (r)		1.65	116.0
School poverty (r)	-0.004	-1.25	-71.8
% new teachers (r)	-0.009		-168.6
Adult educational attainment		2.25	90.5
ECE centers	-0.004	0.49	-637.2
High-quality ECE centers	-0.002	0.36	259.9
ECE enrollment	0.010	1.20	-50.2
Third grade math proficiency			724.2
Third grade reading proficiency	-0.006	-0.50	-911.4
High school graduation rate			-69.3
Advanced Placement course enrollment	-0.013	-0.76	60.0
College enrollment in nearby institutions	-0.024	-4.89	63.7
% Black/African-American	-0.039	-0.69	237.9
Residential segregation (G-statistic)	-0.005	-6.09	-247.3

[a] Reported coefficients are from models using the tuning parameter identified by the minimum MSE from 10-fold cross-validation. Blank cells indicate variables where coefficients were estimated to be zero in the LASSO model.
[b] For NDVI and tree canopy models, a higher value represents more green space.
[c] In contrast, the parks measure is operationalized as distance to the nearest park (in meters) with smaller values being interpreted as greater access to green space.

Table 1-7. Standardized coefficients from LASSO models of NDVI, tree canopy, and distance to the nearest park, including only individual-level predictors. [a]

	NDVI [b]	Tree Canopy [b]	Park Proximity [c]
Individual-level predictor			
Maternal race: Black/African-American	-0.013	-1.46	
Maternal education: HS degree			62.3
Maternal education: technical school		-1.32	
Maternal education: college degree			
Maternal education: grad/prof degree		2.46	
Income	0.0004	0.73	95.0
Household size	0.003	0.67	
Married/living with partner	0.012	1.59	
Maternal age		0.05	

[a] Blank cells indicate variables where coefficients were estimated to be zero in the LASSO model. Reported coefficients are from models using the specified tuning parameter identified by the minimum MSE from 10-fold cross-validation.
[b] For NDVI and tree canopy models, a higher value represents more green space.
[c] In contrast, the parks measure is operationalized as distance to the nearest park (in meters) with smaller values being interpreted as greater access to green space.

Table 1-8. Coefficients from LASSO models of NDVI, tree canopy, and distance to the nearest park, including predictors at both the neighborhood and individual levels.

	NDVI	Tree Canopy	Park Proximity
Neighborhood-level predictors			
Poverty rate (r)	-0.023	-4.93	158
Public assistance rate (r)	-0.007	-0.53	8
Homeownership rate	0.026	2.87	-61
High-skill employment	-0.003	0.52	34
Median household income	0.002	0.56	-244
Employment rate			-70
Commute duration (r)	0.004	0.82	-21
Single-headed households (r)	0.000	1.76	99
School poverty (r)	-0.003	-1.22	-72
% new teachers (r)	-0.009		-163
Adult educational attainment		1.81	69
ECE centers	-0.005	0.29	-611
High-quality ECE centers	-0.002	0.40	248
ECE enrollment	0.010	1.23	-51
Third grade math proficiency			556
Third grade reading proficiency	-0.009	-1.04	-726
High school graduation rate		-0.05	-47
Advanced Placement course enrollment	-0.013	-0.75	51
College enrollment in nearby institutions	-0.025	-5.03	60
% Black/African-American	-0.036	-0.66	192
Residential segregation (G-statistic)	-0.006	-6.21	-233
Individual-level predictors			
Maternal race: Black/African-American	-0.010	0.35	32
Maternal education: HS degree	-0.004	0.02	69
Maternal education: technical school degree	-0.008	-1.70	96
Maternal education: college degree	-0.004	1.32	
Maternal education: graduate/professional degree			
Household income	0.005	0.76	-2
Household size	0.005	0.92	3
Marital status: married/living with partner	0.005	0.82	26
Maternal age	-0.001	-0.25	

[a] Reported coefficients are from models using the specified tuning parameter identified by the minimum MSE from 10-fold cross-validation. Blank cells indicate variables where coefficients were estimated to be zero in the LASSO model.
[b] For NDVI and tree canopy models, a higher value represents more green space.
[c] In contrast, the parks measure is operationalized as distance to the nearest park (in meters) with smaller values being interpreted as greater access to green space.

Chapter 2. Associations of green space with externalizing behaviors and attention problems in the CANDLE cohort

Introduction

In addition to individual-level risk factors, neighborhood exposures may impede or promote healthy neurodevelopment (95). Early childhood is an influential window for these contextual exposures. In particular, natural environments are hypothesized to positively affect child behavioral health. Specifically, Attention Restoration Theory (ART) posits that natural environments facilitate recovery from attention fatigue, improving executive functioning (23). Green space may also provide opportunities for physical activity, in turn leading to improved attention and reduced behavioral problems (96). Alternatively, green space may reduce exposures to environmental hazards such as air pollution, which have been linked to adverse externalizing behavioral outcomes (6).

Prior studies of green space and child behavior identified more consistent associations with attention problems, including Attention-Deficit/Hyperactivity Disorder (ADHD) symptoms, than for other behavior domains (11). Additionally, several studies have identified associations between green space and executive functions such as working memory, as well as with lower risk of ADHD diagnosis (19,20). However, some studies have not accounted for potentially important confounders or utilized study samples across quite distinct urban settings and ecological regions. Others have only identified associations within subgroups such as with residential surrounding greenness only among girls or park access only among boys or for various green space measures in groups with lower education or income levels (43–45,45,97).

In this study, we investigated associations of surrounding residential greenness and tree canopy, and residential proximity to parks, with externalizing behaviors among children 4 to 6 years old. Specifically, we examined relationships with broadband externalizing behavior scores, attention problems, and aggressive behavior in a well-characterized cohort in Memphis, TN. We also explored effect modification by neighborhood SES and child sex.

Methods

Study population

We used data from the CANDLE cohort, which is described above in detail. Briefly, pregnant women (N=1503) were enrolled in CANDLE between 2006 and 2011. Enrollment included women in the second trimester of their pregnancy, who were considered to have a low medical risk pregnancy. Between ages 4 and 6, CANDLE participants completed a suite of online surveys, as well as an in-person study visit. This analysis was limited to those who completed the age 4-6 visit. The analytic sample was further restricted to those who reported a current address within Shelby County, TN or who reported a residential history within Shelby County for at least 75% of the year prior to the study visit.

Externalizing behaviors and attention problems

The preschool (1.5-5 years) version of the Achenbach System of Empirically Based Assessment Child Behavior Checklist (CBCL) was administered at the CANDLE age 4-6 study visit. Mothers reported “not true” (coded as 0), “somewhat or sometimes true” (coded as 1), or “very true or often true” (coded as 2), for 99 child behaviors. Each question was asked in reference to the two months prior to the study visit and missing responses were treated as zeroes. Responses to these individual items were summed to calculate the broadband externalizing score as well as related syndrome subscales.

The primary outcomes in this analysis were the broadband externalizing score and the attention problems syndrome scale. At the preschool age, the attention problems syndrome scale contributes to the broadband externalizing score and has shown diagnostic value in screening for ADHD (98). We considered the aggressive behavior syndrome scale as a secondary outcome. The raw scores for each syndrome scale and for the broadband externalizing score were modeled continuously in primary analyses.

Each scale was also considered as a binary outcome using clinical and borderline-clinical thresholds. These thresholds are defined based on t-scores, which are normed based on a referent sample and truncate raw scores at the low end of the scale; we used raw scores rather than t-scores as the primary outcome due to this truncation. In secondary analyses, a t-score greater than or equal to 60 and less than or equal to 63 was considered in the borderline range and t-scores above 63 were considered above the clinical cutoff for the broadband externalizing score. For all subscales, t-scores in the 65-69 range were defined as borderline and scores above 69 were considered above the clinical threshold.

Green space

In this study, we examined three distinct measures of exposure or access to green space. Details on the derivation of these measures is included in Appendix A. First, the Normalized Difference Vegetation Index (NDVI) was used to assess the overall greenness of the area surrounding the residential location. We calculated NDVI at each participant residence using 2011 annual data at 30m resolution from the NASA Global Web-Enabled Landsat Data (GWELD) (64). Water was excluded, resulting in a scale ranging from 0 to 1 in each pixel of the dataset. We considered NDVI within a 300m buffer as the primary exposure, and examined NDVI in smaller (100m) and larger (500m and 1000m) buffers in further sensitivity analyses.

The second green space measure used in this study was the percent of land area covered by tree canopy. Tree cover data were obtained from the US Environmental Protection Agency (EPA) EnviroAtlas and reflect the percent of the census block group covered by trees (66). This measure includes street trees, parks, urban forests, and single trees on various properties, which were derived from 1m resolution landcover data and aggregated to the census block group level. Percent tree canopy was weighted by the residential history across the year prior to the age 4-6 study visit.

In the primary analysis of NDVI and tree cover, green space exposures were weighted by each of the addresses where the child lived in the year prior to the date of outcome ascertainment. In sensitivity analyses of NDVI and tree cover, we used the residential history from birth to outcome assessment, the residential history from age 1 to age 4, the address at the time of the age 4-6 study visit, or the address at which the child lived the longest.

In addition to these exposure measures assessing residential surrounding greenness and tree canopy, we also examined a measure of access to green spaces. Park proximity to the residence was calculated as Euclidean distance from the home location to the nearest boundary of a park in meters, using data on park location and boundaries compiled by the Trust for Public Land (67). Distance to the nearest park was calculated for the address at the time of outcome assessment as the primary measure and for the address at which the child lived the longest in a sensitivity analysis. In further sensitivity analyses, we calculated the distance to the nearest small park (< 2 acres), neighborhood park (2-20 acres), and community park (>20 acres), separately. Unlike the other two green space measures for which a higher value represents greater exposure to natural environments, park proximity was coded such that a lower value represents closer proximity, conceptualized as better access.

Covariates

An extensive suite of variables has been collected in CANDLE, including both maternal and child information. Maternal education was reported in five categories (<HS, HS degree, technical school, college degree, or graduate/professional degree). Maternal race was included as African-American/Black or White/Other. Household income was reported in 8 categories ($0-$15,000 was the lowest category, each of the next 6 categories were in increments of $10,000, and the highest category was $75,000 or more) at the age 4-6 visit. We converted income to a continuous variable by selecting the midpoint of each category; in the highest category the Pareto distribution was used to assign the income level. This continuous income variable was then adjusted for the number of adults and children in the household at the age 4-6 visit using the OECD equivalence scale (99). Maternal IQ was assessed using the WASI short form and maternal depression was assessed using the Center for Epidemiological Studies-Depression scale (CES-D) at the age 4-6 study visit; both were included as continuous covariates. Maternal tobacco smoking during pregnancy was defined as either self-report of tobacco use or a urinary cotinine level greater than 200µL in a 2^{nd} or 3^{rd} trimester sample. Childhood secondhand smoke exposure in the household (yes/no) was assessed via questionnaire at the age 4-6 visit. Gestational age at birth and birthweight were obtained from medical records and dichotomized; preterm was defined as a gestational age <37 weeks and low birthweight was defined as <2500 grams. Residential instability was calculated as number of changes of address between the child's birth to the age 4-6 study visit from reported residential history.

Several child behaviors were also reported at age 4-6 in CANDLE and were included in extended models. Physical activity was reported as how many times in a normal week the child engaged in vigorous physical activity in three categories (never or occasionally, once or twice per week, and three or more times per week). Screen time, specified as including watching television and using a computer, was reported as the number of hours per day. Sleep habits were reported via the Children's Sleep Habits Questionnaire and a continuous total sleep score was calculated.

Several neighborhood-level measures and features of the residential location were obtained. Urbanicity was assessed at the census tract level using census designations based on population density. Neighborhood resources were operationalized using the socioeconomic and education opportunity subscales of the Childhood Opportunity Index (COI) (70). The socioeconomic scale is comprised of several variables at the census tract level, including poverty rate, homeownership rate, median household income, and employment rate. The education scale includes factors related to early childhood education, elementary education, secondary and post-secondary education, and educational resources. In extended models, we accounted for living near a major roadway. Distance to the nearest major roadway (class A1, A2, or A3 road) was dichotomized at 150m to indicate a near-road residence.

Statistical Analyses

Descriptive statistics were calculated for exposures, outcomes, and covariates. We used linear regression with robust standard errors to assess the association between green space and the continuous raw CBCL scores. Logistic regression was used to estimate associations with CBCL scores dichotomized at clinical and borderline clinical thresholds.

We used a staged model approach to covariate adjustment. Model 1 was considered minimally-adjusted, with only child sex and child age at outcome assessment included as covariates in the model. Model 2 was additionally adjusted for socioeconomic status, neighborhood resources, and proxies for socioeconomic resources. Given the history of residential segregation and discrimination in the US, maternal race was considered a proxy measure for access to such resources and corresponding exposure to stressors. Covariates in model 2 included maternal education, household income adjusted for household size, maternal race, residential stability, COI socioeconomic scale, COI education scale, and urbanicity. Model 3 includes all of the covariates in model 2 and further adjusts for some additional factors related to child neurodevelopment, including maternal IQ, maternal depression, maternal smoking during pregnancy, childhood secondhand smoke, preterm birth, and low birthweight. Model 3 was considered our primary model. Sensitivity analyses included a set of extended models, each with further adjustment for another individual or environmental factor, including child sleep score, child physical activity, child screen time, and a residential location near a major roadway. Effect modification by neighborhood SES and child sex was assessed by inclusion of a multiplicative interaction term in the model.

All data manipulations, visualizations, and analysis were conducted in R 3.6 (The R Foundation for Statistical Computing; Vienna, Austria).

Results

The analytic sample included 943 CANDLE participants (Figure 2-1). In CANDLE, 1,030 participants completed the CBCL at the age 4-6 study visit. Participants were further excluded if they did not have a valid geocoded address history such as an address resolution only at the zip code level or an address outside of Shelby County, TN, for more than 25% of the year prior to CBCL outcome ascertainment.

Characteristics of the participants in the sample overall and by quartile of NDVI exposure are included in Table 2-1. Approximately half of the sample were boys and the mean age of children at the time of the CBCL assessment was 4.3 years (SD 0.4). In this sample, 65% of mothers were African-American/Black and 41% had a college or graduate/professional degree. Fewer participants in the highest quartile of NDVI exposure were African-American/Black compared to those in the lowest quartile of NDVI. In the highest quartile of NDVI, 47% had a college or graduate/professional degree compared to 43% in the lowest quartile of NDVI.

The distribution of CBCL scores in the externalizing domain, as well as the percent of children scoring above clinical or borderline-clinical thresholds are shown in Table 2-2. In this sample, 7.7% and 6.8% of children scored above borderline-clinical thresholds on the broadband externalizing score and attention problems syndrome scale, respectively.

The distributions of the three primary measures of green space are shown in Table 2-3. NDVI in this cohort is generally high (mean 0.59 [SD 0.08]). Mean tree cover was 38% (SD 12). 28% of the cohort lived within 300m of the nearest park; 11% lived within 300m of a community park (>20 acres). Further descriptives of green space measures are included in Appendix A.

Effect estimates for NDVI exposures and each externalizing outcome scale were in the hypothesized direction but all confidence intervals included the null (Figure 2-2). For example, in the fully adjusted model a 0.1 unit higher NDVI in 300m buffer was associated with a 0.29 lower externalizing score (95% CI: -0.91, 0.34). Attention problems were not associated with NDVI (-0.06 [95% CI: -0.22, 0.11] difference in attention problems score per 0.1 unit higher NDVI). Estimated differences in externalizing broadband scores tended to be larger when NDVI was assessed in larger buffer sizes (Table 2-4). Regression coefficients for NDVI tended to be similar across model staging adjustment.

Estimates of associations between tree cover and each of the child behavior scales had confidence intervals that included the null (Figure 2-2). In this sample, a 10% higher tree canopy was associated with a 0.27 points lower externalizing domain score (95% CI: -0.65, 0.11) with confidence intervals including the null. In sensitivity analyses, we observed similarly null associations across varying exposure windows. Additional adjustment for several covariates did not change effect estimates.

In primary analyses, no associations were observed with distance from the nearest park and externalizing behavior (Figure 2-2). In a sensitivity analysis looking at distance to the nearest parks when classified by park size, there was some suggestion that living 500m farther from the nearest community park (>20 acres) was associated with a higher score on the attention problems scale (0.05, 95% CI: 0, 0.11, p=0.06).

We assessed effect modification by neighborhood socioeconomic conditions and child sex (Table 2-5). Regression coefficients for NDVI tended to be larger for the group with lower socioeconomic opportunities, though p-values for interaction were well above 0.05. For the attention problems scale, we observed a stronger association with NDVI in the hypothesized direction among boys, but p-values for interaction were greater than 0.05. No effect modification was observed in analyses of tree canopy or park proximity and externalizing scores.

When the CBCL scales were dichotomized, we observed lower odds of an aggressive behavior score above the clinical threshold (OR 0.56, 95% CI: 0.32, 0.97) per 10% higher tree cover (Table 2-6).

Discussion

In this study, we did not observe an association between greenness and attention problems or externalizing scores in 4 to 6 year old children. For residential surrounding greenness and tree cover exposures, we generally observed effect estimates in the hypothesized direction (more greenness associated with fewer behavioral problems) though confidence intervals included the null. In a sensitivity analysis, higher tree canopy was associated with lower odds of aggressive behavior scores above clinical thresholds. When parks were limited to community parks (<20 acres), there was some suggestion that living further from a park was associated with more attention problems. However, many tests were conducted with primarily null results and therefore results from sensitivity analyses should be interpreted cautiously.

Several prior studies have identified associations between green space and externalizing behavior among children. Bijnens *et al*. estimated a lower CBCL externalizing score (2.0 points lower externalizing t-score [95% CI: -3.5, -0.4]) per IQR higher green space in a 3km residential buffer (100). Others have also identified associations between higher neighborhood greenness and lower scores on broadband externalizing measures among nationally representative samples of children in Australia and South Korea (37,101). In studies that further probed associations with subsets of behaviors within the externalizing domain, several have identified relationships with attention problems and hyperactivity scores (36,101–103). This prior research includes studies of NDVI in buffer sizes down to 100m, as well as exposures weighted by both residential and preschool locations (36,102). Protective associations have also been observed when access to green space was operationalized as proximity to the nearest urban green space or city park (43,104). There is some variability in the published literature on this topic; several studies have reported effects only in some subgroups or null findings across all externalizing scales (37,44,45). While prior research has generally suggested more consistent results for externalizing behaviors related to inattention, a small number of studies have also identified associations between more green space and lower conduct problem or aggressive behavior scores in older children and adolescents (35,101,102,105). In our study, though some sensitivity analyses hinted at associations with attention problems or aggressive behavior, we conducted many tests without adjusting for multiple comparisons. In our primary analysis specified a priori, we did not observe an association between green space and externalizing behavior. Our study included a younger sample than most prior studies, which were generally restricted to samples ranging in age from 7 to 17 years. Many behaviors within the externalizing domain, including those related to ADHD, may be more likely to be identified as problems once children reach school-age.

Large population-based cohorts have utilized ADHD diagnosis as the primary study outcome. Strong evidence comes from a cohort study of a cohort of children ages 2-17 years (n=814,689, mean age at diagnosis of 11.5 years) in Denmark (19). Importantly, the authors accounted for multiple measures of individual and neighborhood level socioeconomic status in their primary model. Thygesen *et al*. estimated a 3% higher risk of ADHD across childhood and adolescence (95% CI: 1.02, 1.03) per 0.1 unit decrease in NDVI during early childhood (0-5 years). The authors found that this association was attenuated when adjusted for exposure to nitrogen dioxide (NO_2). Other large cohorts, including a study of 10 to 14-year-olds in Germany (n=66,823), have similarly observed associations between higher greenness and lower risk of ADHD (106). Some of these studies suggest that associations were sensitive to the choice of exposure metric. A study from seven cities in China observed lower odds of ADHD (OR 0.87, 95% CI: 0.83, 0.91) per 0.1 unit increase in NDVI within 500m of the school location, but did not observe an association with greenness around the residential location (107). A large New Zealand-based study estimated an association between minimum NDVI between age 2-18 and ADHD (108). Taken together, these large cohort studies from urban settings in multiple countries suggest a protective effect of higher NDVI exposure for ADHD that may be explained in part by reduced exposure to NO_2.

To explore potential effects of green space on specific executive functions, several studies have utilized computerized tests of attention and working memory. Performance on tests of attention was associated with residential surrounding greenness in a cohort of children 4-7 years old in Spain (21). Surrounding greenness, particularly at the school location, was also associated with 12 month changes in attention and working memory among children 7-10 years old (20). Specifically, this study found an increase in the progress of working memory and a reduction in inattentiveness with more greenness, and traffic-related air pollution appeared to mediate this association. A study of spatial working memory—one of four components of working memory—found that more green space was associated with better spatial working memory in children age 11 (109). However, associations with executive functioning were not observed in a sample of adolescents (110). An important contribution of these studies is the use of executive functioning assessments to both address limitations of using parent-reports of child behavior and provide insight into the specific functions that may be most affected by exposure to natural environments. While our study adjusted for multiple factors known to influence reporting of child behavior such as maternal depression, CBCL outcomes may still be limited by maternal reporting on behavioral problems.

Some evidence for an effect of natural environments on attention and executive functions among children comes from small short-term experimental studies in both clinical and nonclinical study populations. Most of these experimental studies have compared a nature walk with an urban walk, and have observed effects such as a reduction in symptoms among children diagnosed with ADHD and improved performance on an attention task among a sample of preschool-age children (111,112). In contrast, improved performance on measures of attention was not observed in a sample ages 10-14, though the study did observe faster and more stable response times on the assessment after the nature walk (113). These studies benefit from cross-over experimental designs to address the between-person confounding that threatens validity in the observational literature and suggest there is a short-term benefit of time actually spent in nature settings. In our study, we were unable to assess time spent in green spaces and this may in part explain our null results. Alternatively, the short-term effect observed in some experimental studies that may not translate to a sustained impact on children's behavior over longer periods of time.

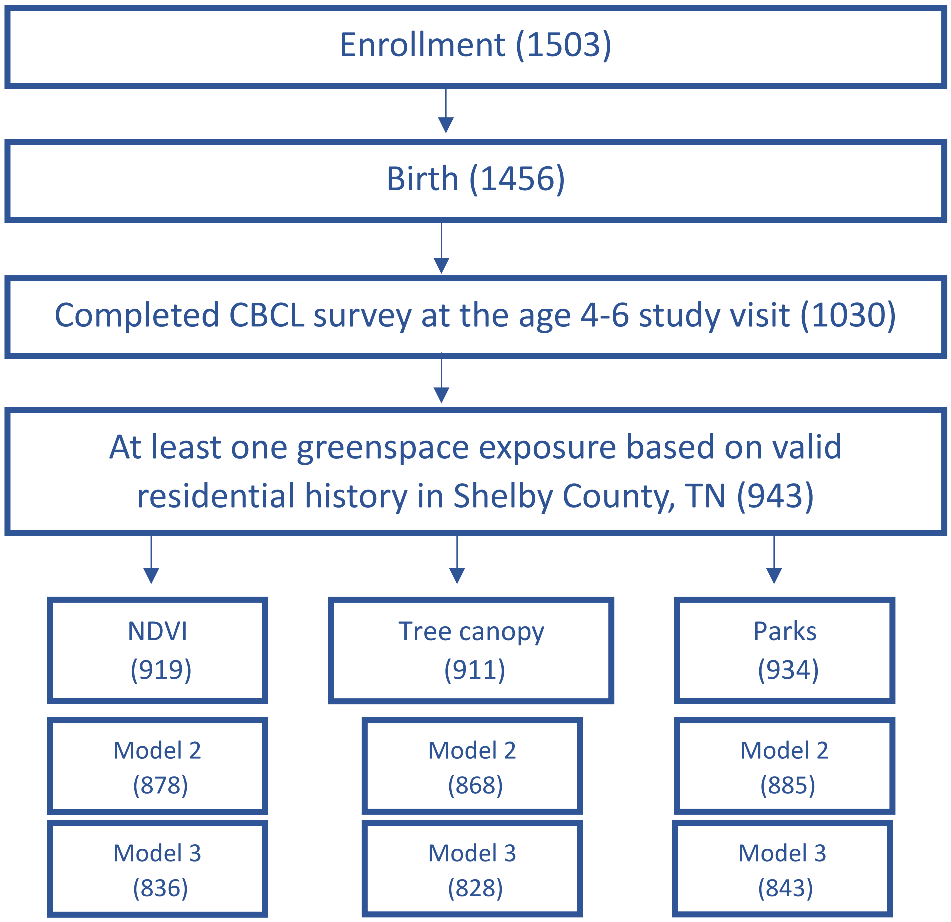
Our study had several limitations. First, we did not have information on whether or not CANDLE participants were attending preschool or child care at a location other than the residence; our exposure assessment was restricted to the home location. Several studies have observed more consistent associations with green space at school, than for green space at home (36,107). Use of 2011 NDVI data assumed a constant spatial distribution such that there were not changes in exposure over time from sources other than changes in the residential history. Furthermore, the spatial resolution of the tree canopy measure was limited to the census block group geography, which may not be the most relevant assessment of a neighborhood tree effect. The park proximity measure might be improved by restricting to parks that include a playground, which may be more likely to reflect actual use of those parks, particularly by families with young children. While views of green space from inside a house, car, or classroom may impact attention, we did not have any measures of actual interaction with nature in this study. We were also unable to assess the time children spent outside in green spaces. Due to the overall high levels of green space in Memphis, there was limited exposure variability for this cohort and this may explain the null results observed here. While the CBCL is a well-established, validated, widely-used tool in research settings, it relies on parent report of child behaviors and may be subject to outcome misclassification. Some mothers may be more or less likely to report problems to clinicians in a research setting or responses may be relative to other children the parent observes in the child's peer group, which may have on average either more or fewer problems that national norms. Underreporting of behavior may have contributed to attenuation of any true effects in this analysis.

Our study also had several strengths. The CANDLE cohort is a well-characterized cohort, with frequent contact throughout childhood, which provided a more refined residential history in calculating exposures as well as robust adjustment for many potential confounders and precision variables at multiple levels. In particular, prior studies have cited potential residual confounding due to insufficient adjustment for parental mental health or neighborhood resources. We were able to adjust for maternal depression, as well as neighborhood opportunity in two domains. Furthermore, the address history collected across childhood was highly resolved, allowing us to explore multiple exposure windows. Spatial resolution of greenness exposure was also improved in this study, using source data for NDVI was more highly resolved than the 250m resolution data used in many prior studies.

To enhance our understanding of this research area, future work to consider more specific measures of attentional control and working memory will be useful, as will examining developmental trajectories of externalizing behaviors. Additionally, improving exposure assessment by incorporating measures of green space at locations where children may be more likely to spend time outdoors such as schoolyards and playgrounds or measures of time spent in green spaces is an important component to strengthen this literature.

Tables & Figures

Figure 2-1. Flowchart of inclusion in the analytic sample for each analysis.



Between enrollment and birth, 47 mothers were lost to follow-up. At the age 4-6 follow-up visit, the outcomes were ascertained for 71% of the children in the cohort at birth. The primary exposures in this analysis required a valid address history over the year prior to the age 4-6 visit (for NDVI and tree canopy) or a valid address at the time of the visit (for park proximity). Addresses were considered invalid for the purposes of this analysis if they could not be geocoded to a more granular level than the zip code or fell outside of the boundary of Shelby County, TN. When averaging exposures across the year prior to the study visit, exposures were excluded if more than 25% of the relevant residential history was missing. Reductions in sample size in Model 2 were primarily due to missing household income data and in Model 3 were primarily due to missing maternal depression.

Table 2-1. Characteristics of the analytic sample overall and by quartile of NDVI.

	Overall (n=943)		NDVI							
			Q1 (n=230)		Q2 (n=230)		Q3 (n=229)		Q4 (n=230)	
Boys, n (%)	465	(49)	115	(50)	113	(49)	108	(47)	120	(52)
Girls, n (%)	478	(51)	115	(50)	117	(51)	121	(53)	110	(48)
Child age in years, mean (SD)	4.3	(0.4)	4.3	(0.4)	4.3	(0.4)	4.3	(0.4)	4.3	(0.4)
Maternal race, n (%)										
African-American/Black	611	(65)	166	(72)	156	(68)	151	(66)	123	(53)
Not African-American/Black	332	(35)	64	(28)	74	(32)	78	(34)	107	(47)
Maternal education, n (%)										
<HS	52	(6)	8	(3)	15	(7)	7	(3)	20	(9)
HS degree	376	(40)	93	(40)	100	(44)	94	(41)	81	(36)
Technical school	124	(13)	30	(13)	29	(13)	38	(17)	21	(9)
College degree	237	(25)	65	(28)	53	(24)	52	(23)	61	(27)
Graduate/professional degree	147	(16)	34	(15)	28	(12)	38	(17)	45	(20)
Adjusted household income, mean (SD)	17903	(13514)	17129	(13224)	17302	(13170)	17932	(13699)	19697	(13973)
Maternal IQ percentile, mean (SD)	40.3	(30.7)	40.2	(29.4)	36.7	(29.8)	40.7	(31.4)	44.1	(32.4)
Maternal depression score, mean (SD)	8.6	(7.2)	9	(8.4)	9	(7.2)	7.8	(6.3)	8.2	(6.8)
Maternal smoking during pregnancy, n (%)	81	(9)	8	(3)	25	(11)	21	(9)	22	(10)
Child secondhand smoke exposure, n (%)	287	(31)	71	(31)	74	(32)	67	(30)	65	(28)
Low birthweight, n (%)	66	(7)	18	(8)	16	(7)	17	(7)	13	(6)
Preterm birth, n (%)	83	(9)	20	(9)	17	(7)	26	(11)	19	(8)
Sleep score, mean (SD)	46.6	(7.3)	46.9	(7.4)	46.8	(7.1)	46.4	(7.0)	46.3	(7.5)
Physical activity, mean (SD)	1.6	(0.7)	1.5	(0.8)	1.7	(0.7)	1.6	(0.7)	1.6	(0.7)
Screen time, mean (SD)	2	(4.2)	1.8	(3.7)	2.1	(4.2)	2	(4.3)	2	(4.6)
COI: socioeconomic subscale, mean (SD)	-0.113	(0.260)	-0.131	(0.271)	-0.116	(0.257)	-0.106	(0.238)	-0.101	(0.272)
COI: education subscale, mean (SD)	-0.047	(0.068)	-0.047	(0.063)	-0.052	(0.063)	-0.05	(0.064)	-0.037	(0.079)
Near road, n (%)	264	(28)	83	(36)	77	(34)	55	(24)	43	(19)

Table 2-2. Distribution of CBCL scores in the externalizing domain in the analytic sample of CANDLE participants at age 4-6 years (n=943).

CBCL scale	**Continuous raw scores** [a]		**Above clinical threshold** [b]		**Above borderline-clinical threshold** [b]	
	Mean	(SD)	N	(%)	N	(%)
Externalizing score	9.18	(7.57)	45	(4.8)	73	(7.7)
Attention problems	2.20	(1.94)	30	(3.2)	64	(6.8)
Aggressive behavior problems	6.98	(6.17)	16	(1.7)	37	(3.9)

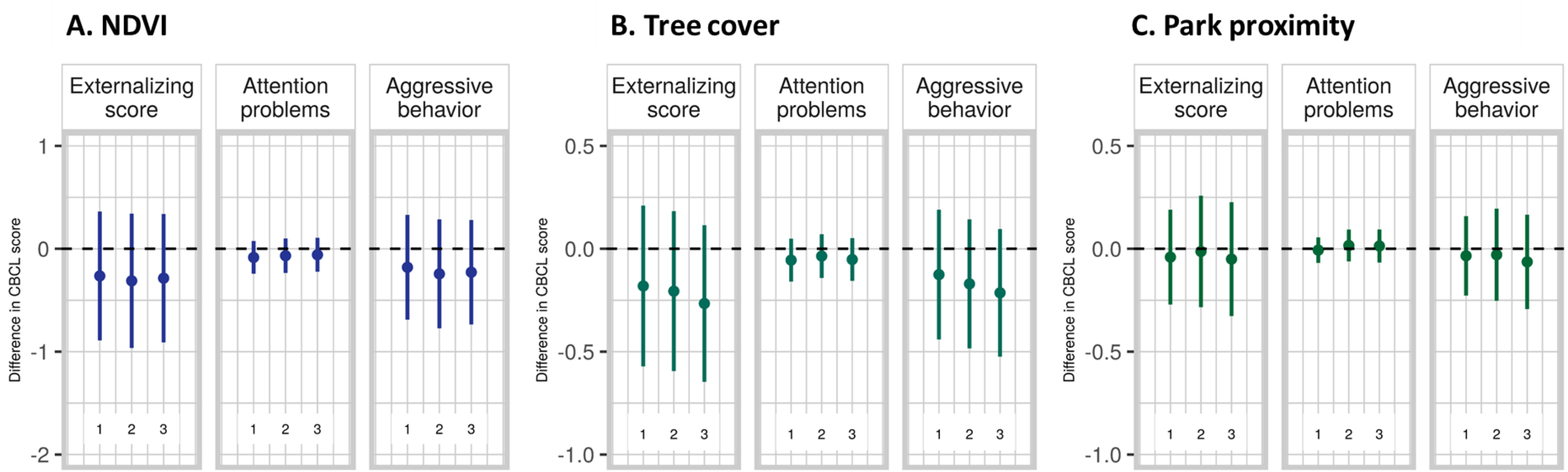
[a] The range of the externalizing score is 0-43 in this sample; the range of the attention problems scale is 0-9 and the range of the aggressive behavior problems scale is 0-36 in this sample.
[b] For the externalizing score, the clinical threshold is defined as a t-score greater than 63 and above the borderline-clinical threshold is defined as a t-score greater than or equal to 60. For the attention problems and aggressive behavior problems scales, above the clinical threshold is defined as a t-score greater than 69 and above the borderline-clinical threshold is defined as a t-score greater than or equal to 65.

Table 2-3. Distribution of green space exposures in the analytic sample (n=943).

Greenspace	**N**	**Mean**	**(SD)**	**Min**	**25th p.**	**Median**	**75th p.**	**Max**
NDVI [a]	919	0.592	(0.08)	0.249	0.543	0.596	0.647	0.789
Tree cover [b]	911	37.6	(12.0)	4.4	29.1	37.3	45.7	80.7
Park proximity [c]	934	795	(862)	3	282	549	987	8417

[a] NDVI was defined in a 300m buffer, weighted by the address history over the year prior to the age 4-6 study visit. Participants were excluded if they did not have a valid geocoded address for >25% of the exposure window.
[b] Tree cover was defined as the percentage of census block group, weighted by address history over year prior to age 4-6 study visit. Participants were excluded if they did not have a valid geocoded address for >25% of the exposure window.
[c] Park proximity was defined as Euclidean distance from the address at the age 4-6 study visit to the edge of the nearest park. Participants were excluded if the address at the age 4-6 study visit could not be geocoded or could only be geocoded to the zip code level.

Figure 2-2. Associations of residential green space with externalizing behavior.



Difference (95% confidence intervals) in externalizing score, attention problems scale, and aggressive behavior score are shown for (A) 0.1 unit higher NDVI in 300m buffer, (B) 10% higher tree cover in the census block group, and (C) 500m further distance to the nearest park. NDVI and tree cover exposures were calculated as a weighted average across all residential locations in the year prior to the age 4-6 study visit. Park proximity was calculated for the current address at the time of the age 4-6 study visit. Model 1 included only child sex and child age at outcome assessment as covariates in the model. Model 2 was additionally adjusted for maternal education, household income adjusted for household size, maternal race, socioeconomic COI scale and education COI opportunity, and residential stability. Model 3 was considered primary and further adjusted for maternal IQ, maternal depression, maternal smoking during pregnancy, childhood secondhand smoke exposure, preterm birth, and low birthweight.

Table 2-4. Sensitivity analyses for models of associations of residential green space with externalizing scores and attention problems.

	NDVI		**Tree cover**		**Park proximity**	
CBCL outcome	Externalizing score	Attention problems	Externalizing score	Attention problems	Externalizing score	Attention problems
Primary analysis [a]	-0.29 (-0.91, 0.34)	-0.06 (-0.22, 0.11)	-0.27 (-0.65, 0.11)	-0.05 (-0.16, 0.05)	-0.05 (-0.33, 0.23)	0.01 (-0.07, 0.09)
Sensitivity analyses						
Extended models [b]						
Sleep score	-0.29 (-0.89, 0.32)	-0.06 (-0.22, 0.10)	-0.24 (-0.61, 0.13)	-0.05 (-0.15, 0.06)	-0.04 (-0.30, 0.22)	0.02 (-0.06, 0.09)
Physical activity	-0.29 (-0.92, 0.33)	-0.06 (-0.22, 0.11)	-0.24 (-0.63, 0.14)	-0.05 (-0.15, 0.06)	-0.06 (-0.34, 0.21)	0.01 (-0.07, 0.09)
Screen time	-0.26 (-0.89, 0.37)	-0.05 (-0.22, 0.11)	-0.25 (-0.64, 0.13)	-0.05 (-0.15, 0.06)	-0.05 (-0.33, 0.22)	0.01 (-0.07, 0.09)
Near road residence	-0.29 (-0.92, 0.35)	-0.05 (-0.22, 0.12)	-0.27 (-0.65, 0.11)	-0.05 (-0.16, 0.05)	-0.05 (-0.33, 0.23)	0.02 (-0.06, 0.10)
Exposure windows [a]						
Age 1-4	-0.18 (-0.85, 0.48)	0 (-0.18, 0.18)	-0.24 (-0.66, 0.18)	0.01 (-0.10, 0.12)	-	-
All childhood	-0.12 (-0.84, 0.59)	0.02 (-0.17, 0.21)	-0.22 (-0.65, 0.20)	0.01 (-0.10, 0.13)	-	-
Current address	-0.34 (-0.90, 0.23)	-0.11 (-0.26, 0.05)	-0.22 (-0.60, 0.15)	-0.05 (-0.14, 0.05)	-	-
Longest address	-0.11 (-0.68, 0.47)	-0.01 (-0.17, 0.14)	-0.27 (-0.65, 0.11)	0 (-0.10, 0.10)		
Exposure buffers [a]						
100m	-0.05 (-0.56, 0.47)	-0.06 (-0.2, 0.08)	-	-	-	-
500m	-0.37 (-1.04, 0.29)	-0.07 (-0.24, 0.11)	-	-	-	-
1000m	-0.64 (-1.38, 0.11)	-0.05 (-0.25, 0.15)	-	-	-	-
By park size [a]						
Small parks	-	-	-	-	0.04 (-0.04, 0.12)	0.02 (0, 0.04)
Neighborhood parks	-	-	-	-	0.03 (-0.14, 0.21)	0.01 (-0.04, 0.06)
Community parks	-	-	-	-	0.12 (-0.09, 0.33)	0.05 (0, 0.11)

[a] Adjusted for the full set of covariates in model 3: child sex, child age, maternal education, household income adjusted for household size, maternal race, socioeconomic COI scale and education COI opportunity, residential stability, maternal IQ, maternal depression, maternal smoking during pregnancy, childhood secondhand smoke exposure, preterm birth, and low birthweight.
[b] Adjusted for all covariates in model 3, plus the additional covariate indicated in each row.

Table 2-5. Effect modification of associations of residential green space with externalizing scores and attention problems.

	NDVI		**Tree cover**		**Park proximity**	
CBCL outcome	Externalizing score	Attention problems	Externalizing score	Attention problems	Externalizing score	Attention problems
Neighborhood factors [a]						
Socioeconomic scale						
25th p.	-0.56 (-1.43, 0.31)	-0.1 (-0.3, 0.1)	-0.28 (-0.78, 0.23)	-0.05 (-0.18, 0.09)	-0.27 (-0.85, 0.32)	-0.02 (-0.18, 0.13)
50th p.	-0.25 (-0.85, 0.36)	-0.05 (-0.22, 0.11)	-0.27 (-0.64, 0.11)	-0.05 (-0.16, 0.05)	-0.14 (-0.51, 0.23)	0 (-0.10, 0.10)
75th p.	0.11 (-0.63, 0.85)	0 (-0.22, 0.23)	-0.25 (-0.73, 0.22)	-0.06 (-0.19, 0.07)	0 (-0.27, 0.28)	0.02 (-0.06, 0.11)
Interaction p-value	0.23	0.45	0.94	0.88	0.35	0.55
Individual characteristics [a]						
Child sex						
Boys	-0.48 (-1.41, 0.44)	-0.19 (-0.41, 0.03)	-0.34 (-0.86, 0.18)	-0.08 (-0.22, 0.06)	-0.05 (-0.43, 0.32)	-0.02 (-0.12, 0.09)
Girls	-0.1 (-0.9, 0.71)	0.06 (-0.16, 0.29)	-0.20 (-0.75, 0.35)	-0.03 (-0.18, 0.12)	-0.05 (-0.39, 0.3)	0.04 (-0.06, 0.15)
Interaction p-value	0.53	0.11	0.72	0.59	0.97	0.40

[a] Models were adjusted for child sex, child age, maternal education, household income adjusted for household size, maternal race, socioeconomic COI scale and education COI opportunity, residential stability, maternal IQ, maternal depression, maternal smoking during pregnancy, childhood secondhand smoke exposure, preterm birth, and low birthweight. Models additionally included a multiplicative interaction term between the green space exposure and the specified effect modifier.

Table 2-6. Odds ratios (95% confidence intervals) for externalizing scales above clinical and borderline-clinical thresholds.

CBCL scale	NDVI	Tree Cover	Park Proximity
Externalizing t-score			
Clinical threshold (>63)	0.77 (0.48, 1.25)	0.83 (0.62, 1.12)	0.98 (0.74, 1.30)
Borderline threshold (>=60)	0.95 (0.66, 1.38)	0.90 (0.70, 1.15)	0.99 (0.80, 1.22)
Attention problems t-score			
Clinical threshold (>69)	0.85 (0.42, 1.72)	0.88 (0.56, 1.38)	1.13 (0.78, 1.62)
Borderline threshold (>=65)	0.94 (0.64, 1.38)	0.96 (0.74, 1.25)	1.11 (0.88, 1.40)
Aggressive behavior t-score			
Clinical threshold (>69)	0.56 (0.25, 1.23)	0.56 (0.32, 0.97)	1.20 (0.80, 1.79)
Borderline threshold (>=65)	0.88 (0.53, 1.45)	0.84 (0.61, 1.16)	0.98 (0.70, 1.37)

[a] Models were adjusted for child sex, child age, maternal education, household income adjusted for household size, maternal race, socioeconomic COI scale and education COI opportunity, residential stability, maternal IQ, maternal depression, maternal smoking during pregnancy, childhood secondhand smoke exposure, preterm birth, and low birthweight.

Chapter 3. Associations of residential surrounding greenness, tree cover, and park proximity with internalizing behaviors in early childhood

Introduction

Natural environments may benefit mental health across the life course (10,11,114–116). Green space may be particularly relevant for mood disorders and higher greenspace has been found to be associated with reduced depressive symptoms and better general mental health in adolescents and adults (33,117–119). However, evidence in pediatric samples is more limited than that for older populations, despite trajectories of internalizing behavior problems potentially beginning early in childhood with lasting consequences for health and well-being (7,11). Exposure to natural environments may act through psychological pathways to promote positive affect or generate physiologic changes, including impacts on the autonomic nervous system and the hypothalamic-pituitary-adrenal axis (41,120,121). Prior research suggests potential effects on social contacts, stress reduction, and physical activity in adolescents (10,116,122). These pathways may also be important for relationships between green space and child mental health.

In this chapter, our objective was to assess the relationship between green space and internalizing behaviors in a sample of preschool-age children in Memphis, TN. We assessed a broadband internalizing behavior score in primary analyses. We investigated the four syndrome subscales that comprise the broadband internalizing score—anxious/depressed, emotionally reactive, somatic complaints, and withdrawn scales—in secondary analyses. We explored the robustness of observed associations to varying exposure windows and further adjustment for potential mediating variables, and examined effect modification by neighborhood SES and child sex.

Methods

Study population

The analytic sample in chapter 3 is the same as that used in chapter 2, drawn from the CANDLE cohort. Briefly, CANDLE is a community-based pregnancy cohort in Shelby County, TN, who enrolled women between 2006 to 2011. CANDLE is one of three cohort studies within the ECHO-PATHWAYS consortium. Mother-child dyads have been followed over time, with multiple study visits conducted during early childhood. The sample in this analysis was restricted to participants who completed the age 4-6 visit and who reported a current address within Shelby County, TN at the time of the study visit or for at least 75% of the year prior to the study visit.

Internalizing behaviors

Child behavior was assessed using Child Behavior Checklist (CBCL). For each behavior, mothers selected “not true”, “somewhat true”, or “very or often true”, which were coded as 0, 1, and 2, respectively. Each scale is calculated from a subset of the 99 behaviors reported on by mothers on the CBCL at the age 4-6 study visit. The primary outcome of interest was the broadband internalizing score. In secondary analyses, we explored the four syndrome scales in the internalizing domain: the emotionally reactive, anxious/depressed, somatic complaints, and withdrawn scales. In secondary analyses, each scale was also considered as a dichotomous outcome using clinical and borderline clinical thresholds.

Green space

As in chapter 2, we used three distinct measures of green space. These measures are described briefly here, with further details included in Appendix A. In this study, we calculated residential surrounding greenness using NDVI from the 30m resolution NASA Global Web-Enabled Landsat Data (GWELD) (64). We weighted NDVI by each of the addresses where the child had been living in the year prior to the date of outcome ascertainment, in buffers of 100m, 300m (considered primary), 500m, and 1000m. Tree cover was assessed over the same exposure window as NDVI, using data from the US Environmental Protection Agency (EPA) EnviroAtlas (66). This measure includes street trees, parks, urban forests, and single trees on various properties, which were derived from 1m resolution landcover data and aggregated to the census block group level. Park proximity was calculated for the address at the time of outcome assessment. Park data were obtained from a publicly available dataset compiled by the Trust for Public Land (67). Proximity was calculated as Euclidean distance from the geocoded home location to the closest park boundary.

The primary green space measures used as exposures in this analysis were NDVI in a 300m buffer and tree cover in the census block group, weighted by addresses in year prior to the age 4-6 visit. Sensitivity analyses include using the exposure window from age 1 to age 4, using the longest-lived address, using multiple buffer sizes of NDVI, and distance to larger neighborhood or community parks.

Covariates

Information on many potential confounders was collected at the age 4-6 study visit. Maternal education was reported in five categories (<HS, HS degree, technical school, college degree, or graduate/professional degree). Household income was reported in 8 categories ($0-$15,000 was the lowest category, then each of the next 6 categories were in increments of $10,000 and the highest category was $75,000 or more) at the age 4-6 visit. We converted reported income to a continuous variable by selecting the midpoint of each category; in the highest category the Pareto distribution was used to assign the income level. This continuous income variable was then adjusted for household size (number of adults and number of children) using the OECD formula

(99). Maternal race was considered as a proxy measure of socioeconomic resources based on racial residential segregation and included in the model as African-American/Black or White/other. Maternal depression was assessed using the Center for Epidemiological Studies-Depression scale (CES-D) at the age 4-6 study visit and modeled continuously. Maternal IQ percentile was assessed using the WASI short form. Child physical activity, screen time, and sleep habits were also reported at age 4-6. Physical activity was reported as how many times in a normal week the child engaged in vigorous physical activity (never or occasionally, once or twice per week, and three or more times per week). Screen time, including watching television and using a computer, was reported as the number of hours per day. Sleep habits were reported via the Children's Sleep Habits Questionnaire and a continuous total sleep score calculated. The Parenting Relationship Questionnaire (PRQ) was administered at the age 4-6 visit and the attachment score was used as a continuous variable.

Covariate information was also collected at multiple other time points in the CANDLE cohort. Maternal tobacco smoking during pregnancy was defined using both self-report or urinary cotinine greater than 200 ng/mL. Gestational age at birth and birthweight were obtained from medical records and dichotomized at standard thresholds (<37 weeks for preterm birth and <2500g for low birthweight).

Participant addresses were collected frequently throughout the study follow-up period, beginning at enrollment during pregnancy, providing a highly temporally resolved residential histories. Residential instability was calculated as number of changes of address between the child's birth to the age 4-6 study visit from reported residential history. Address histories were also used to characterize several neighborhood-level conditions. Neighborhood resources were operationalized using the socioeconomic and education opportunity domain scales of the

Childhood Opportunity Index (COI) (70). The socioeconomic scale is composed of several variables at the census tract level, including poverty rate, homeownership rate, median household income, and employment rate. The education scale includes factors related to early childhood educational opportunity, elementary education, and adult educational attainment. Some additional neighborhood measures were also considered in sensitivity analyses. Distance to major roadway (class A1, A2, or A3 road) at the current address was dichotomized at 150m to indicate a near-road residence.

Statistical analysis

Descriptive statistics were used to explore distributions of green space exposures, CBCL scores, and covariates. We used linear regression with robust standard errors to assess the association between green space and internalizing CBCL scores. Logistic regression was used to estimate associations with internalizing CBCL scores dichotomized at clinical and borderline-clinical thresholds.

We used a staged model approach to covariate adjustment. Model 1 was considered minimally-adjusted, with only child sex and child age at outcome assessment included as covariates in the model. Model 2 was additionally adjusted for maternal education, household income adjusted for household size, maternal race, COI socioeconomic and education scales, residential stability, and urbanicity. Model 3 was considered primary, and further adjusts for maternal IQ, maternal depression, PRQ attachment score, maternal smoking during pregnancy, preterm birth, and low birthweight. Sensitivity analyses explored extended models with additional adjustment for residence near a major roadway, child physical activity, screen time, or sleep. Effect modification was explored by including a multiplicative interaction term in the full model.

All analyses were run using R 3.6 (The R Foundation for Statistical Computing; Vienna, Austria).

Results

There were 943 children with at least one of the three primary green space exposure measures and CBCL scores at age 4-6 in CANDLE. In models 2 and 3, up to 49 and 91 participants were excluded due to missing covariates, primarily due to missing data on household income and maternal depression, respectively (see Chapter 2, Figure 2-1). Children were on average 4.3 (SD 0.4) years old at the time of assessment; 49% were boys (Table 3-1). In this sample, 65% of mothers identified as African-American/Black; 53% had at least a college or technical school degree. Mean household income adjusted for household size was $17,900 (SD 13,500). The mean internalizing domain score was 6.21 (SD 6.17) in this sample (Table 3-2); 5.3% of children in this sample had an internalizing score above the clinical threshold and 11.0% had an internalizing score above the borderline-clinical threshold.

The distribution of green space measures is illustrated in Figure 3-1. The median NDVI in a 300m buffer in this sample was 0.596 (IQR 0.104). Correlation between the measures is high (0.6) for tree cover and NDVI; park proximity was not correlated with either NDVI or tree canopy. 28% of children lived within 300m of the nearest park. Further details on the distribution of green space exposures can be found in Appendix A.

Higher levels of residential surrounding greenness were associated with lower internalizing scores (Figure 3-2). In the fully-adjusted model, a 0.1 unit higher NDVI exposure was associated with a 0.64 units lower internalizing score (95% CI: -1.24, -0.04; p=0.036). Higher NDVI was also associated with a lower anxious/depressed score (β -0.20; 95% CI: -0.40, 0.00; p=0.046) and a lower somatic complaints score (β -0.17; 95% CI: -0.34, 0.00; $p=0.046$) (Figure 3-3). Effect estimates tended to be consistent across the different covariate adjustment approaches in models 2 and 3 and in sensitivity analyses (Table 3-3). Differences in the internalizing score were similar for NDVI in larger buffer sizes, but attenuated for exposure in the 100m buffer. Effect estimates were similar for exposures weighted by the address history from age 1 to 4, but were attenuated with exposures calculated at a single address.

While effect estimates tended to be in the hypothesized direction (Figure 3-2), confidence intervals included the null in models of tree cover and internalizing behaviors (0.27 lower internalizing score per 10% higher tree cover; 95% CI: -0.60, 0.05; $p=0.098$). Results for the anxious/depressed scale were borderline statistically significant (Figure 3-3). A 10% higher tree canopy exposure was associated with a 0.10 units lower anxious/depressed score (95% CI: -0.21, 0.00; $p= 0.056$). Findings were consistent across models that further adjusted for additional covariates and in models with tree cover weighted by address history across various windows in early childhood (Table 3-3). No associations were observed between park proximity and internalizing scores in primary or sensitivity analyses (Figure 3-2 and Table 3-3).

Analysis of effect modification by neighborhood conditions suggested a larger association between NDVI and CBCL internalizing scores in neighborhoods with lower socioeconomic opportunity (Table 3-4). However, p-values for the interaction term were all greater than 0.05. No effect modification was observed for neighborhood conditions and tree canopy. No effect modification by child sex was observed.

When the internalizing score was considered as a binary outcome in secondary analyses, confidence intervals included the null (Table 3-5). For example, we estimated an odds ratio of 0.67 (95% CI: 0.42, 1.08) per 0.1 unit higher NDVI.

Discussion

Higher total residential surrounding greenness within 300m of the residential location, but not tree cover or distance to the nearest park, was associated with lower internalizing scores (fewer problems). In secondary analyses, these associations were most consistent for the anxious/depressed and somatic complaints syndrome scales. We observed slightly larger differences in internalizing scores when assessing exposures in larger buffer sizes. We conducted a number of sensitivity analyses in which additional covariates were added to the model or exposures were averaged by residential history across varying windows of time during early childhood. Results from these sensitivity analyses were generally consistent with conclusions from the primary analysis.

Several prior studies have investigated relationships between green space and internalizing behaviors in children. Madzia *et al*. observed lower anxiety t-scores per 0.1 unit higher NDVI in 800m (-1.83, 95% CI: -3.44, -0.22), as well as lower depression and somatization t-scores per 0.1 unit higher NDVI in 200m (-1.36 [95% CI: -2.16, -0.12] and -1.83 [95% CI: -3.22, -0.44], respectively) (35). These associations were identified among children at age 12 years, but not at age 7 years, using the Behavior Assessment System for Children, Second Edition (BASC-2) to assess behavior in a cohort in Cincinnati, OH. The distribution of residential surrounding greenness (mean NDVI of 0.55 [SD 0.10] in a 400m buffer) was similar to that in our study. A higher proportion of the Cincinnati cohort reported levels of behavioral problems above clinical thresholds—13%, 11%, and 15% above clinical thresholds for somatization, depression, and anxiety at age 7—than in the CANDLE cohort. These differences may reflect the use of different behavioral assessment tools, the younger age of our analytic sample, a lower underlying prevalence of behavioral problems, a greater hesitancy among parents to report behavior problems, or other differences in the selection of these sample populations. Amoly *et al*. similarly identified associations between green space and internalizing behaviors in a pediatric cohort, using the Strengths and Difficulties Questionnaire (SDQ) to assess child behaviors (36). This study of children ages 7 to 10 years old in Barcelona, Spain estimated a 4.3% lower emotion problems score (95% CI: -8.1, -0.1) per IQR higher residential surrounding greenness in a 500m buffer. This finding was supported by consistent results when refining the exposure metric to specify green space playing time.

Other studies provide more limited evidence for associations between green space and internalizing behavior, though effect estimates across these studies trended in the hypothesized direction. A study of children ages 5 to 6 years old in China examined surrounding greenness in 100m buffers, but only observed statistically significant associations with anxious/depressed CBCL scores when the exposure was weighted for greenness at both the residential and kindergarten locations (102). Studies including adolescents in the analytic sample observed some associations for greenness in larger buffer sizes. A study in Belgium found an association with CBCL internalizing score for NDVI in a 3km buffer, but not for NDVI in other buffer sizes (100). Comparing exposure in the highest tertile to the lowest tertile was suggestive of an association between higher greenness and lower internalizing CBCL score in a South Korean cohort, though confidence intervals included the null (101). NDVI exposures were generally lower in the South Korean cohort (median 0.29 [IQR 0.15]) than for CANDLE participants in

Memphis, TN. However, several published studies have not observed associations between green space and internalizing behaviors or only observed associations in some subgroups (43,45,46,123). Effect modification by child age suggested associations between neighborhood green space and internalizing behaviors among children ages 4-5 years old that weakened among those aged ≥10 years (37).

Studies of older children and adolescents have utilized self-reported measures of anxiety and depression, rather than parent-reported measures of internalizing behaviors. Two studies in adolescents identified associations between surrounding greenness in buffers of 1000m and 1250m and lower odds of depressive symptoms (33,118). However, in another cohort associations between surrounding greenness in buffers up to 800m were observed for symptoms of anxiety but not depression (34).

Few studies have examined tree cover and internalizing behaviors. While an array of potential benefits of tree canopy have been suggested in the literature, the quantitative evidence is still limited (8). We did not observe associations with tree canopy in the current study. One limitation of the tree cover exposure measure we used was that it was only available at the census block group geography so we were not able to explore associations across multiple sizes. The location of the trees may also be relevant; our exposure measure did not distinguish between trees lining the roadway versus wooded areas.

Access to green spaces has more commonly been operationalized as distance to the nearest urban green space or park in studies of outcomes in the internalizing domain. Distance to the nearest green space was associated with increased odds of a peer relationship problems score above the clinical threshold (OR 1.20 [95% CI: 1.02-1.40] per 500m increase in distance to nearest green space) in a study of 10-year-olds in Munich, Germany (104). In a cohort of 4 to 5 year old children, an association between proximity to city parks and the peer relationship problems scale was observed, but only among those with lower maternal education (43). Unlike these previous studies, we did not observe associations between proximity from parks and any subscales. One limitation of the current study is that we do not have a measure of neighborhood park use or whether the nearest park includes a playground; young children and their caregivers may not spend time in the park closest to the residence.

Multiple mechanisms have been proposed for relationships between green space and mental health. Stress Recovery Theory (SRT) posits that natural environments promote recovery from physiological stress (40,124). Experiments across multiple age groups suggest a relationship between natural environments and physiologic stress response measures (41,125). Observational literature also suggests associations with reported or perceived stress measures. Using geographic ecological momentary assessment in a sample of adolescents, urban green space was found to be associated with lower stress when participants were in locations other than their residence (126). Green space was associated with lower perceived stress in a sample of adolescents in California (127). The literature suggests that the stress recovery and attention restoration pathways may be particularly relevant for natural environments in the school setting for older children. We were unable to capture exposure in locations other than the residence in the current analysis, but future work using data from school-age and adolescence in the CANDLE cohort age may be able to examine this question with more refined exposure assessments.

Evidence of effect modification by SES at either individual or neighborhood levels in prior studies is mixed, with some estimating a stronger association with green space in the group with lower SES and others estimating a stronger association with green space in the group with higher SES (43–47). Though not statistically significant, the magnitude of the association between NDVI and internalizing behavior in this study was estimated to be larger for those in neighborhoods with lower socioeconomic opportunity. Further confirmation of this trend in future work could support green space as a resource to promote health equity.

This analysis was subject to several limitations. First, we relied on observational data in a cross-sectional study and thus we were unable to make a strong causal statement. We relied on parent-reported outcomes which, while well-validated for use in research populations and appropriate for this age, are nonetheless limited to the behavior problems parents are willing or able to report to research staff. Due to the particular urban setting in which the CANDLE cohort resides, there was limited variability in green space exposures. Any effects of green space may occur at lower levels and taper off at higher levels of exposure. We also did not have information about any child care or preschool locations where children may have spent most of their time, or information about the amount of time children actually spent outside in natural environments. As in much of the literature on green space and health, we relied on residential surrounding greenness as a proxy for exposure. While some researchers have identified relatively small activity spaces of children in their neighborhoods, we likely are not fully capturing all exposure to green space, which may result in some misclassification of the exposure. Perhaps most importantly, due to the observational nature of these data and despite the breadth of confounders we accounted for, the associations we observed may be explained by residual confounding, particularly due to factors related to residential selection into neighborhoods.

This study contributes an analysis from a well-characterized pediatric cohort in a US city to the growing literature on the relationship between green space and mental health. The location of the cohort in Memphis, TN provides a distinct geographic context for examining associations with natural environments. Few studies have explored relationships of natural environments to neurodevelopment in the US and fewer still have examined multiple forms of green space. We were able to examine multiple green space exposures using a highly-resolved address history across all of childhood. Furthermore, we adjusted for a suite of potential confounders at both the individual and neighborhood levels.

As the children in this cohort age into a life stage where more mental health concerns tend to emerge, future research might investigate relationships between early life green space exposures and trajectories of neurobehavioral development or the onset of depression and anxiety in later childhood and adolescence. Furthermore, as children develop more independence, larger activity spaces as well as different modes of interacting with natural environments may expand or alter exposures to green spaces. Analysis of school-age children may incorporate school-based exposures to further refine exposure assessment. Additionally, with rapid increases in computing power and advances in the field of green space exposure assessment, future studies may improve upon the current exposure measures.

Tables & Figures

Table 3-1. Characteristics of the analytic sample (n=943).

	Mean (SD) or N (%) [a]	
Boys, n (%)	465	(49)
Girls, n (%)	478	(51)
Child age in years, mean (SD)	4.3	(0.4)
Maternal race, n (%)		
African-American/Black	332	(35)
Not African-American/Black	611	(65)
Maternal education, n (%)		
<HS	52	(6)
HS	376	(40)
Technical school	124	(13)
College degree	237	(25)
Graduate/professional degree	147	(16)
Adjusted household income, mean (SD)	17903	(13514)
Maternal IQ, mean (SD)	40.3	(30.7)
Maternal depression score, mean (SD)	8.6	(7.2)
Maternal smoking during pregnancy, n (%)	81	(9)
Low birthweight, n (%)	66	(7)
Preterm birth, n (%)	83	(9)
PRQ Attachment score	52.9	(9.7)
Sleep score	46.6	(7.3)
Physical activity score	1.6	(0.7)
Screen time (hours/day)	2	(4.2)
Socioeconomic COI scale	-0.113	(0.260)
Education COI scale	-0.047	(0.068)
Residence within 150m of a major road	264	(28)

[a] Covariate data were missing for some participants: maternal education (7), adjusted household income (41), maternal IQ (10), maternal depression (9), low birthweight (6), preterm birth (5), PRQ attachment score (25).

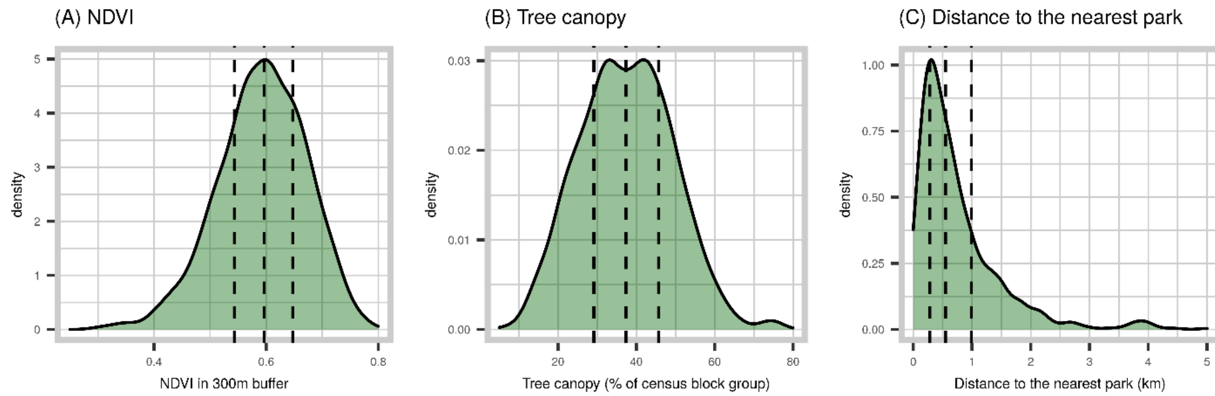
Table 3-2. Distribution of CBCL scores in the internalizing domain for CANDLE participants at age 4-6 (n=943).

CBCL scale	Continuous raw score [a]		Above clinical threshold [b]		Above borderline-clinical threshold [b]	
	Mean	(SD)	N	(%)	N	(%)
Internalizing score	6.21	(6.17)	50	(5.3)	104	(11.0)
Emotionally reactive syndrome scale	1.65	(2.00)	9	(1.0)	55	(5.8)
Anxious/depressive syndrome scale	1.80	(2.03)	9	(1.0)	36	(3.8)
Somatic complaints syndrome scale	1.33	(1.75)	19	(2.0)	54	(5.7)
Withdrawn syndrome scale	1.43	(1.83)	36	(3.8)	65	(6.9)

[a] The range of the internalizing score in this sample is 0-44, the range of the emotionally reactive scale is 0-15, the range of the anxious/depressed scale is 0-12, the range of the somatic complaints scale is 0-12, and the range of the withdrawn scale is 0-14.

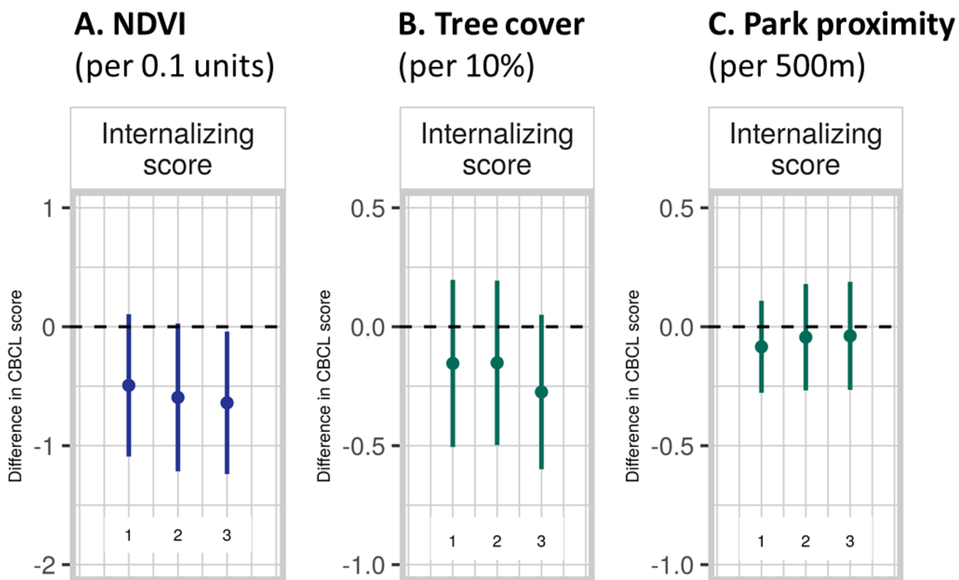
[b] For the internalizing score, the clinical threshold is defined as a t-score greater than 63 and above the borderline-clinical threshold is defined as a t-score greater than or equal to 60. For the syndrome subscales, above the clinical threshold is defined as a t-score greater than 69 and above the borderline-clinical threshold is defined as a t-score greater than or equal to 65.

Figure 3-1. Density plots of (A) NDVI in 300m, (B) tree canopy, and (C) distance to the nearest park, in the analytic sample.



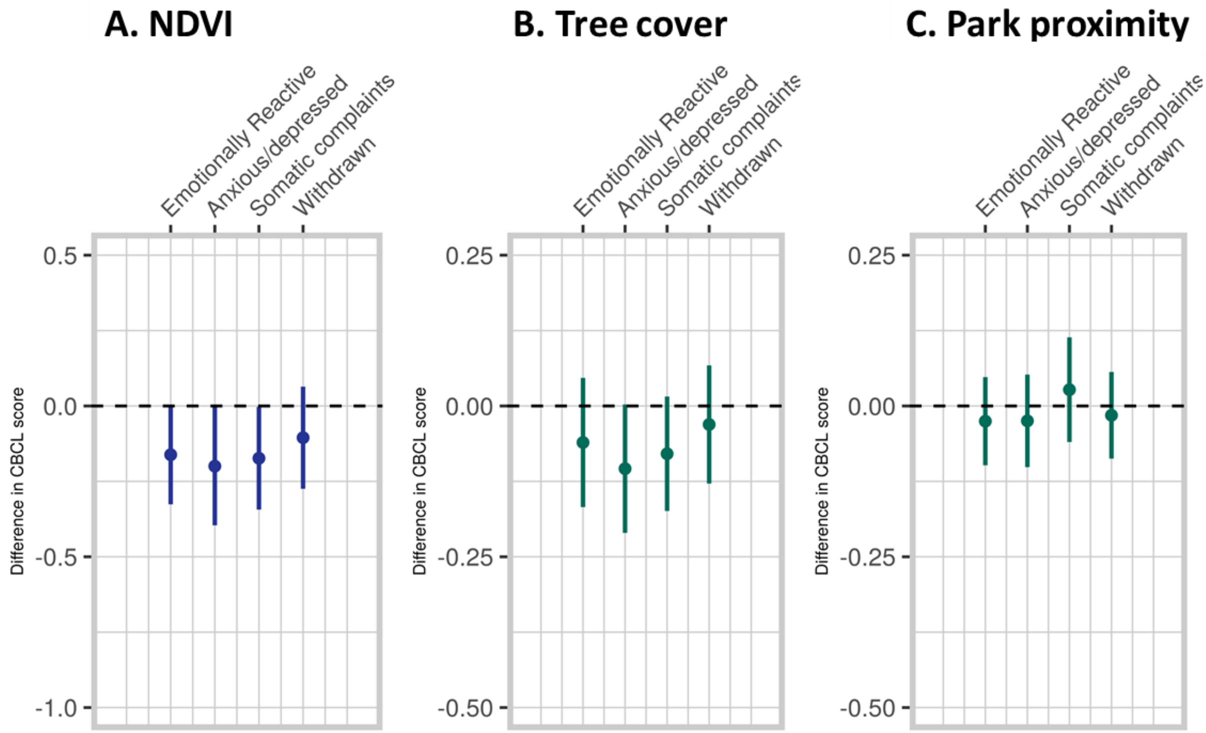
NDVI was defined in a 300m buffer, weighted by the address history over the year prior to the age 4-6 study visit. Participants were excluded if they did not have a valid geocoded address for >25% of the exposure window. Tree canopy was defined as the percentage of census block group, weighted by address history over year prior to age 4-6 study visit. Participants were excluded if they did not have a valid geocoded address for >25% of the exposure window. Park proximity was defined as Euclidean distance from the address at the age 4-6 study visit to the edge of the nearest park; participants were excluded if the address at the age 4-6 study visit could not be geocoded or could only be geocoded to the zip code level.

Figure 3-2. Residential green space and internalizing scores.



Difference (95% confidence intervals) are shown per (A) 0.1 unit higher NDVI, (B) 10% higher tree cover, and (C) 500m further distance to the nearest park. NDVI and tree cover exposures were averaged over the residential history across the year prior to the outcome assessment. Park proximity was calculated for the current address. Model 1 included only child sex and child age at outcome assessment as covariates in the model. Model 2 was additionally adjusted for maternal education, household income adjusted for household size, maternal race, socioeconomic COI scale and education COI opportunity, and residential stability. Model 3 was considered primary and further adjusted for maternal IQ, maternal depression, PRQ attachment score, maternal smoking during pregnancy, preterm birth, and low birthweight.

Figure 3-3. Residential green space and CBCL syndrome subscales in the internalizing domain.



Difference (95% confidence intervals) are shown per (A) 0.1 unit higher NDVI, (B) 10% higher tree cover, and (C) 500m further distance to the nearest park. Models were adjusted for child sex, child age at outcome assessment, maternal education, household income adjusted for household size, maternal race, socioeconomic COI scale and education COI opportunity, residential stability, maternal IQ, maternal depression, maternal smoking during pregnancy, preterm birth, low birthweight, and PRQ attachment score. P-values were less than 0.05 for NDVI and anxious/depressed and somatic complaints syndrome scales. P-values were less than 0.10 for NDVI and the emotionally reactive syndrome scale and for tree cover and the anxious/depressed scale.

Table 3-3. Sensitivity analyses of residential green space and broadband internalizing scores.

	NDVI	Tree cover	Park proximity
Primary analysis [a,b,c]	-0.64 (-1.24, -0.04)	-0.27 (-0.60, 0.05)	-0.04 (-0.27, 0.19)
Sensitivity analyses [a,b]			
Extended models [c]			
Sleep score	-0.64 (-1.23, -0.06)	-0.26 (-0.58, 0.06)	-0.03 (-0.24, 0.19)
Physical activity	-0.64 (-1.25, -0.04)	-0.26 (-0.58, 0.07)	-0.05 (-0.29, 0.18)
Screen time	-0.62 (-1.23, -0.02)	-0.27 (-0.59, 0.06)	-0.04 (-0.26, 0.19)
Near road residence	-0.68 (-1.30, -0.06)	-0.28 (-0.60, 0.05)	-0.04 (-0.27, 0.18)
Exposure windows			
Age 1-4	-0.58 (-1.14, -0.03)	-0.30 (-0.64, 0.05)	-
All childhood	-0.62 (-1.23, -0.02)	-0.35 (-0.71, 0.01)	-
Current address	-0.46 (-0.95, 0.02)	-0.23 (-0.55, 0.08)	-0.04 (-0.27, 0.19)
Longest address	-0.31 (-0.75, 0.12)	-0.24 (-0.52, 0.04)	0.02 (-0.19, 0.24)
Exposure buffer size [c]			
100m	-0.54 (-1.02, -0.06)	-	-
500m	-0.66 (-1.28, -0.04)	-	-
1000m	-0.65 (-1.28, -0.01)	-	-
By park size [c]			
Small parks	-	-	0.01 (-0.05, 0.08)
Neighborhood parks	-	-	0.03 (-0.11, 0.16)
Community parks	-	-	0.06 (-0.13, 0.25)

[a] Difference (95% confidence interval) in CBCL scores are shown per 0.1 unit higher NDVI, 10% higher tree cover, and 500m further distance to the nearest park.

[b] Models were adjusted for child sex, child age at outcome assessment, maternal education, household income adjusted for household size, maternal race, socioeconomic COI scale and education COI opportunity, residential stability, maternal IQ, maternal depression, maternal smoking during pregnancy, preterm birth, low birthweight, and PRQ attachment score.

[c] NDVI and tree cover exposure measures were a weighted average across all residential locations in the year prior to the age 4-6 study visit. Park proximity was calculated for the current address.

Table 3-4. Effect modification of associations between green space and internalizing scores.[a]

	NDVI	Tree cover	Park proximity
Neighborhood factors			
Socioeconomic opportunity			
25th p.	-0.83 (-1.76, 0.10)	-0.28 (-0.75, 0.18)	-0.12 (-0.64, 0.4)
50th p.	-0.61 (-1.17, -0.05)	-0.27 (-0.59, 0.05)	-0.07 (-0.39, 0.24)
75th p.	-0.36 (-1.00, 0.28)	-0.26 (-0.62, 0.1)	-0.02 (-0.24, 0.21)
Interaction p-value	0.43	0.93	0.70
Individual characteristics			
Child sex			
Boys	-0.94 (-1.90, 0.03)	-0.28 (-0.74, 0.19)	-0.04 (-0.35, 0.27)
Girls	-0.36 (-1.03, 0.30)	-0.27 (-0.73, 0.19)	-0.04 (-0.32, 0.25)
Interaction p-value	0.33	0.72	0.99

[a] Effect modification was assessed by including a multiplicative interaction term in the model; p-values shown are for the interaction term. Differences in CBCL scores (95% confidence intervals) are shown for a 0.1 unit higher NDVI at the 25th percentile, median, and 75th percentile of the effect modifier. Models were adjusted for child sex, child age at outcome assessment, maternal education, household income adjusted for household size, maternal race, socioeconomic COI scale and education COI opportunity, residential stability, maternal IQ, maternal depression, maternal smoking during pregnancy, preterm birth, low birthweight, and PRQ attachment score

Table 3-5. Odds ratios (95% CI) for associations between green space and high internalizing problem scores based on clinical and borderline-clinical thresholds. [a]

CBCL scale	NDVI	Tree Canopy	Park Proximity
Internalizing T-score			
Clinical threshold (>63)	0.67 (0.42, 1.08)	1.05 (0.75, 1.45)	0.99 (0.93, 1.05)
Borderline threshold (≥60)	0.75 (0.56, 1.02)	0.97 (0.78, 1.21)	0.99 (0.96, 1.02)

[a] Models were adjusted for child sex, child age at outcome assessment, maternal education, household income adjusted for household size, maternal race, socioeconomic COI scale and education COI opportunity, residential stability, maternal IQ, maternal depression, maternal smoking during pregnancy, preterm birth, low birthweight, and PRQ attachment score. Odds ratios and 95% confidence intervals are shown per 0.1 higher NDVI, per 10% higher tree canopy, and per 500m further distance to the nearest park.

Conclusion

In this study, we utilized data from the age 4-6 year study visit in the socio-demographically diverse CANDLE cohort, a prospective longitudinal pregnancy cohort in Memphis, TN. We first examined the distribution of green space by other neighborhood and individual characteristics in Chapter 1. Overall greenness and tree canopy were predicted by homeownership rate in the neighborhood, as well as by early childhood educational resources in the neighborhood. We then analyzed the relationships between green space and externalizing behaviors in Chapter 2 and between green space and internalizing behaviors in Chapter 3. Unlike much of the prior literature, we did not observe associations between green space and externalizing behaviors or attention problems in Chapter 2. However, we did estimate a protective effect of green space exposures for internalizing behavior problems in Chapter 3. These associations between green space and internalizing behavior were generally robust to further adjustment for variables potentially on the causal pathway. We did not observe statistically significant effect modification by neighborhood SES, though the magnitude of effect estimates suggested a stronger relationship between green space and child behaviors in lower SES neighborhoods.

In Chapter 1, we took a novel approach to exploring relationships between green space and other neighborhood features. The importance of homeownership rate, early childhood education, and racial residential segregation indicators in these models suggest future research questions regarding equitable access to green space with potential implications for local policy and design of green spaces in Memphis. These results also suggest that neighborhood-level characteristics across multiple domains may be important to consider in the context of confounding in epidemiologic analyses of green space. Prior studies, if they adjusted for any neighborhood-level covariates, have typically adjusted only for a smaller set of indicators specifically related socioeconomic resources, such as some of the indicators within the socioeconomic domain of the COI. However, indicators within the education domain were also highly weighted in the WQS index; given the associations between educational opportunities and healthy child development, these may be important confounders that are not often included in epidemiologic analyses of environmental factors and child behavior and mental health.

The analyses presented in chapters 2 and 3 benefited from several strengths of both exposure and outcome assessment. Exposure assessment in this study improved upon approaches in prior studies by using source data for NDVI at a higher spatial resolution and including measures of tree cover and park proximity. Using 30m resolution data, we were able to examine NDVI in buffer sizes from 100m and larger. Furthermore, the residential history was collected by frequent contact over early childhood. Prior work suggests that early childhood is an important exposure window for neighborhood contexts, including green space (30,128). In chapters 2 and 3, we were able to explore multiple exposure windows in sensitivity analyses based on the residential history. Exposure assessment in this study suffered from several limitations as well. The spatial scale for tree cover was limited by data availability to the census block group level and these administrative boundaries may not be the most relevant aggregation areas for effects on child health. Incorporating measures of time spent in green space or weighting exposure by the school and home locations may reduce exposure measurement error in future studies.

One strength of this study was that we were able to use a continuous measure to assess behavior across a range of severity. In young children, most developmental phenomena are better detected across a continuum of symptoms. These may be more appropriately assessed using continuous scales, as a categorical diagnosis may only capture the most severe behavior problems. One limitation of our approach is the reliance on parent-reported outcomes. While this approach is well-validated for young children, outcome assessments in future studies may be improved by using objective assessments of executive functioning or self-report of internalizing problems later in childhood.

A major limitation in many studies of green space and health is the potential for biased estimates due to residual confounding. A strength of this study was the adjustment for a number of covariates previously cited as potential sources of residual confounding. For example, prior studies have indicated lack of data on parent mental health, birth outcomes, and neighborhood resources, as a limitation. The CANDLE study has collected a large suite of variables on the mother-child dyads in the cohort and we were able to adjust for these confounders. Despite our efforts to address this concern, we cannot rule out the possibility that the observed relationships are due to residual confounding, particularly by socioeconomic factors.

The findings of this study may not be generalizable to other urban settings. The NDVI and tree cover levels in Memphis are higher than observed in many other studies. This may be one reason we did not observe an association with externalizing scores or attention problems. Though exposure variability is limited in analyses within both chapter 2 and 3, different mechanisms may be operating for externalizing and internalizing outcomes and these mechanisms may be more or less relevant in an urban setting with high levels of greenery.

In our study, we were explored extended models adjusting for potential mediators. While the number of tests conducted was large, increasing the potential for Type 1 errors, and we did not adjust for multiple comparisons, results remained generally consistent across sensitivity analyses. This literature would also be strengthened by formal mediation analyses.

Causal inference in observational studies of residential green space and health is limited by study design and potential confounding by selection into neighborhoods. Among young children, the evidence relating natural environments to child behavioral and mental health is largely observational. Experimental studies are typically restricted to short-term exposures. Large-scale experiments of residential neighborhoods have included provided housing vouchers for families and evaluating large-scale greening interventions in residential and school settings (32). These approaches aim to address confounding by characteristics determining selection into neighborhoods and strengthen the evidence for a causal relationship between residential green space and health.

This study contributes to the larger body of literature on green space and pediatric behavioral and mental health, with the potential to inform both clinical practice and urban planning. For the age group considered in our study, current recommendations for initial treatment include behavioral therapy rather than medication (129). Additionally, the potential for behavioral and mental health problems to result in impaired functioning and subsequent negative health outcome across the life course and the large resulting cost at individual and societal levels, suggests that preventive measures are warranted. While this study in isolation does not test whether prescribed time spent in green space is associated with these outcomes, we used measures of availability and access to green space as a proxy for nature contact. Further work specifically testing the effect of park prescriptions is needed to provide insight into the efficacy of these recommendations in clinical practice (130). This literature may also inform policy and urban design; young children should be considered in designing green spaces and facilitating access to these spaces. Even at an early age these green spaces may promote healthy child development, as was observed in chapter 3. Additionally, the more equitable access to parks compared to other measures of green space in Memphis suggests that parks may be a venue for improving access to natural elements, through tree planting and design within existing park spaces. Recent efforts and investments in Memphis have included improving access to nature features such as trees within existing park spaces in historically underserved communities (131).

Cities have re-examined the design and use of outdoor public spaces for social gatherings and physical activity during the COVID-19 pandemic. The COVID-19 pandemic has also highlighted the role of parks and urban green spaces as neighborhood resources for physical and mental health, and their inequitable distribution (132,133). Findings from this study contribute to the rapidly growing literature regarding green space and child behavior, and moving forward urban green spaces may offer opportunities to positively affect children's health and development.

Appendix A. Descriptive analysis of green space measures

This appendix contains an overview of green space measures most commonly utilized in the observational epidemiological literature. It also expands on details provided in chapters 1 through 3 on the methods employed to assess green space, as well as results from descriptive analyses of these various methods in the CANDLE cohort.

No gold standard exists for the assessment of exposure to the natural environment in epidemiologic research and existing research across urban, suburban, and rural settings has measured human contact with nature in a variety of ways (3). In large epidemiologic studies, green space exposure is commonly operationalized as either the quantity of green space within a buffer around a residential location or as the proximity of the residence to the nearest green space (134). The quantity of green space in buffers may represent a cumulative exposure that reflects both intentional and unintentional contact by individuals with those environments (2). Prior research examining green space and health has primarily focused on measures of overall greenness based on satellite imagery. Furthermore, increasing availability of these data at finer spatial resolution and access to expanded processing capabilities has accelerated the use of satellite imagery for this purpose.

However, several authors argued that selection of an exposure metric should be motivated by hypotheses regarding the relevant mechanism by which the natural environment may influence health, as exposures may be differentially associated with particular health outcomes (3,4). While the above-mentioned residence-based exposures may influence health through pathways operating at the neighborhood level, for individuals (i.e., to address psychological, physiological and behavior mechanisms), the use of these measures serves as a proxy for an individual's specific types of contact with those environments. For example, public parks may be particularly relevant for promoting physical activity or social connections.

In this study, three approaches were taken to operationalize green space exposures: (1) residential surrounding greenness assessed using the Normalized Difference Vegetation Index (NDVI), (2) the percentage of the census block group covered by tree canopy, and (3) proximity of the nearest park.

Normalized Difference Vegetation Index (NDVI)

NDVI measures the total vegetation in an area from an overhead view and has frequently been used as a measure of residential greenness (3,135). The color bands in the satellite images are processed based on the wavelengths of light reflected by vegetation. Chlorophyll in plant leaves strongly absorbs visible light (0.4 to 0.7 μm) for photosynthesis, while near-infrared (NIR) light (0.7 to 1.1 μm) is reflected. NDVI is calculated using the following (136):

$$\text{NDVI} = (\text{reflectance of NIR} - \text{reflectance of visible light}) \,/\, (\text{reflectance of NIR} + \text{reflectance of visible light})$$

In any pixel, NDVI can range from -1 to 1. Water results in an NDVI value less than 0. Due to our interest in natural environments and as is commonly done in the literature on the health impacts of greenness, we excluded any pixels less than zero (2). In and around Shelby County, this primarily excludes the Mississippi river and its tributaries. The resulting NDVI scale ranges from 0 to 1. Healthy vegetation reflects more NIR light than unhealthy or sparse vegetation resulting in an NDVI closer to 1; low values reflect built impervious resulting in an NDVI close to 0.

A challenge in obtaining time-resolved measures of NDVI from satellite imagery is missing data, arising in part due to the timing of satellites passing over a specified region. Landsat programs collect data on the same area approximately every 16 days. Cloud cover interference in these images further reduces available ground-level imagery. In prior epidemiological literature, researchers have focused primarily on the maximum vegetation period in order to increase variability in exposure. This is often implemented by selecting images from a single cloud-free day in spring or summer (35,36,43,102). Others have utilized data from all images collected during either a summer month or across a full year, using standardized algorithms to select the maximum value across the specified time frame (101,103).

In this study, NDVI exposures were calculated using NASA Global Web-Enabled Landsat Data (GWELD) v031 data (64). The GWELD data products are derived from imagery collected via the Landsat program of the US Geological Survey (USGS), including available images from Landsat 4 and 5 Thematic Mapper (TM) and Landsat 7 Enhanced Thematic Mapper Plus (ETM+). The satellite imagery is calibrated using top-of-atmosphere reflectance, projected to a map projection that aligns with the MODIS satellite products, and implemented an algorithm to select the highest NDVI in each pixel. Data are available at a 30m resolution. Data for this analysis were downloaded through the NASA and USGS Land Processes Distributed Active Archive Center (LP DAAC). The raw HDF data files were converted to TIF files in the MODIS map projection using the HEG conversion tool for Linux provided by the LP DAAC. The TIF files were then processed using raster-based tools in R (The R Foundation for Statistical Computing; Vienna, Austria) to select the layer containing the calculated NDVI value. Prior work has utilized a range of spatial scales of greenness exposure, both in the spatial scale of the underlying satellite imagery data and in the spatial scale that those data are aggregated to in order to calculate an exposure. One strength of the WELD NDVI data source is the availability of data at the 30m resolution, which improves upon the commonly used, coarser, 250m data (3).

The temporal scale of the NDVI exposure used in the work presented here was driven by data availability. Primary exposure measures were derived using the WELD 2011 annual dataset, which was calculated from December 1st, 2010 through November 30th, 2011. The oldest children in the CANDLE cohort turned 4 starting in 2011 and thus this year was selected due to its proximity to outcome ascertainment in this cohort. However, several potential approaches were explored in descriptive analyses. The finest time-scale provided by this data source was monthly, but for the Memphis area there was substantial missing data for several months in 2011. January 2011 was selected as a month representative of the low vegetation season and compared to July 2011 which was selected as a month representative of the high vegetation season. Data for January, July, and the full year are shown in Figure A1. NDVI was generally high in Shelby County, TN. Spatial patterns of NDVI in Shelby County, TN were similar in both low and high vegetation seasons and in the annual surface used to calculate primary exposure metrics (Figure A1). As expected, July tended to be highly correlated with and had a similar distribution to the 2011 annual measure (Table A1). The 2011 annual measure was selected as the primary exposure to maximize exposure contrasts and to increase the likelihood that we captured greenness due to seasonal vegetation.

Figure A1. Maps of NDVI raster data in Shelby County, TN from (A) January 2011, (B) July 2011, and (C) 2011 Annual.

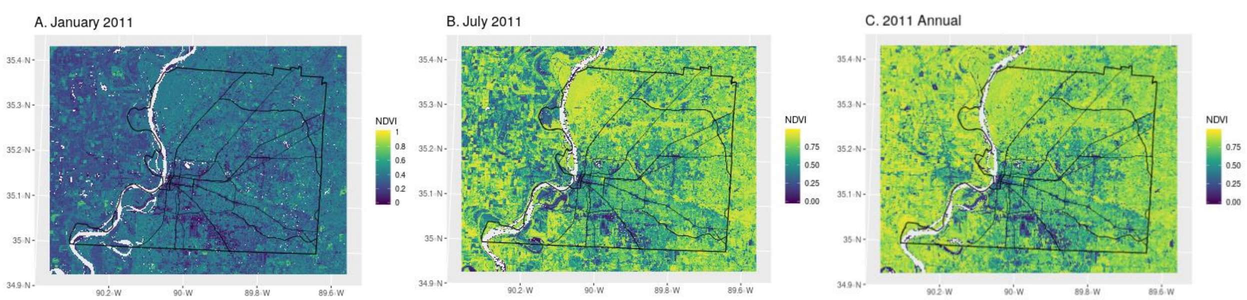


Table A1. Distribution of NDVI exposures derived using satellite imagery from (1) annual 2011 data, (2) January 2011, and (3) July 2011, within in 300m buffer, weighted by address history over the year prior to the age 4-6 visit.

	Mean	SD	Min	Q1	Median	Q3	Max
Annual	0.592	0.080	0.249	0.543	0.596	0.647	0.789
January	0.332	0.049	0.154	0.301	0.336	0.368	0.461
July	0.581	0.081	0.243	0.532	0.582	0.640	0.797

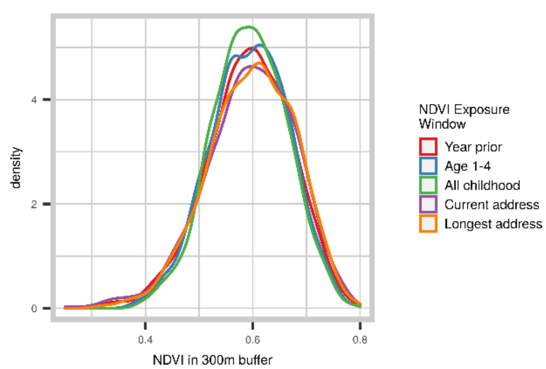
Residential greenspace measures, including NDVI, are often calculated within a straight-line buffer around the residence or using a road network buffer, and buffers of various radii have been used in the literature (3,135). A distance of 300m has often been used in policy and urban planning, as well as in prior epidemiologic studies, particularly when accessibility to greenspace is of interest (65). A smaller buffer likely reflects more private rather than public greenspace; distances of 50m or 100m have been used to capture the most immediate neighborhood which may be particularly important for psychological and physiological mechanisms. Larger buffers such as 500m to 1000m have been justified based on 5-10 minute walking distances for adults (137). We used a 300m buffer in the primary analysis and investigated additional buffer sizes (100, 500, and 1000m) in sensitivity analyses (Table A2). Correlations between the primary 300m buffer and the 100m, 500m, and 1000m buffers were 0.77, 0.91, and 0.66, respectively.

Table A2. Distribution of NDVI exposure in multiple buffer sizes, derived from annual 2011 satellite imagery and weighted by address history over the year prior to the Age 4 study visit.

Buffer size	Mean	SD	Min	Q1	Median	Q3	Max
100m	0.579	0.095	0.202	0.513	0.587	0.652	0.831
300m (primary)	0.592	0.080	0.249	0.543	0.596	0.647	0.789
500m	0.598	0.075	0.242	0.552	0.598	0.649	0.805
1000m	0.606	0.074	0.334	0.560	0.608	0.654	0.835

No clear consensus exists in the literature regarding the most relevant window for developing children’s exposure to green space. Developmental milestones such as walking and other gross motor skills likely influence the ways in which children are exposed to or interact with natural environments. Prior studies have investigated various windows during childhood with some suggestion that early childhood may be particularly important window for exposure (128). Others have averaged exposure measures across all of childhood or for some shorter period of time immediately prior to outcome ascertainment (21). In this study, given the highly resolved residential history, NDVI was averaged across the residential history during multiple windows in early childhood (Figure A2). We used a weighted average of NDVI calculated from each of the addresses where the child had been living in the year prior to the date of outcome ascertainment, in the primary analysis. In sensitivity analyses, we also examined lifelong exposures averaged from birth to outcome assessment, NDVI averaged by the address history between age 1 and age 4, and at the address where the child had lived the longest. Correlations between the prior exposure window of the year prior to the age 4-6 study visit and each additional window ranged from 0.71 to 0.91. Exposures from age 1 to age 4 and across all of childhood (birth to age 4-6 study visit) were highly correlated. Exposure at the current address and the longest childhood address were correlated at 0.59.

Figure A2. Density plot of NDVI in varying exposure windows.

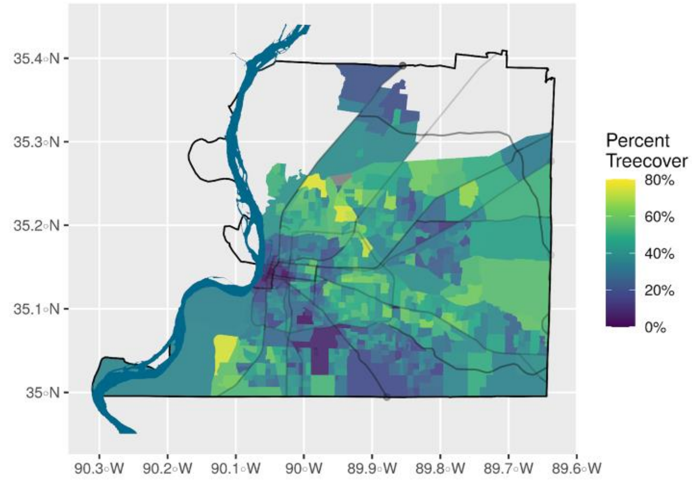


Tree Cover

While green space is often treated as a homogenous exposure when assessed using NDVI, different types of natural environments may be of varying importance for health benefits (138,139). The structure or form of greenspace may be especially relevant for some mechanistic pathways or health outcomes. Studies that have examined an exposure in finer detail have primarily focused on tree cover (8). Furthermore, understanding the components of natural environments that are most associated with improved health outcomes may guide urban planning and intervention. Such advances in exposure measurement might be informed by other disciplines, including ecology, anthropology, or landscape architecture. In particular, work in these disciplines has hypothesized that features such as biodiversity or structure may be important characteristics of natural environments for human health (138).

Tree canopy data were downloaded from the US Environmental Protection Agency (EPA) EnviroAtlas (66). Tree coverage includes street trees, trees in park areas, urban forests, and single trees on various properties. This metric was derived from 1m resolution land cover data and reported as percent of land covered by trees within each census block group, which is shown in Figure A3. In Tennessee, where natural vegetation has not been converted to agriculture, the region is primarily deciduous forests. The most common tree species in urban areas across Tennessee are Chinese privet, Virginia pine, and eastern red cedar, though oaks, maples, dogwood, and poplar trees are also common (53).

Figure A3. Map of tree cover in census block groups in Memphis, TN.



Exposures were matched to participant addresses by the census block group identifier. In primary analyses, percent tree cover was weighted by address history over the year prior to outcome ascertainment at the age 4-6 visit. In the CANDLE cohort, average percent of the block group covered with tree canopy is 36.6% (SD 12.8) for the address history over the primary exposure window (Table A4). In sensitivity analyses, percent tree canopy was weighted by address history over several additional exposure windows. Spearman correlations between these exposure windows ranged from 0.6 – 0.9.

Table A4. Distribution of percent tree canopy at the census block group level across multiple exposure windows in the CANDLE cohort.

Exposure Window	Mean	SD	Min	Q1	Median	Q3	Max
Year prior to age 4-6 study visit	36.6	12.8	1.0	27.6	36.9	45.3	80.7
Age 1 to age 4	36.9	12.1	0.4	27.9	37.3	45.3	81.1
All childhood addresses	36.7	11.6	2.1	28.1	36.5	44.6	81.1
Current (age 4-6 visit) address	37.8	12.5	4.4	28.3	37.8	46.8	81.2
Longest address	37.9	12.9	4.4	28.3	38.5	46.8	81.2

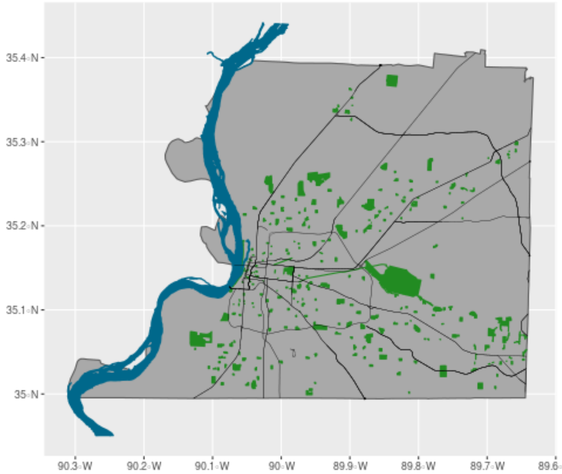
Other sources of land cover information were explored, including the Multi-Resolution Land Characteristics (MRLC) Consortium's 2011 30m resolution National Landcover Dataset compiled by the US Geological Survey (140). Green space types in the MRLC dataset include forests (deciduous, evergreen and mixed), shrubland, and grassland. However, within Memphis, the vast majority of pixels were classified as low-, medium-, or high-density development. Most participants lived near few or no pixels that were designated specifically as any green space type rather than some level of mixed development. Given this limited variability, this data source was not used as an exposure in this study.

Park Proximity

Proximity to public parks was also used as an exposure in these analyses. This measure may be a better proxy for nature contact, particularly in a pediatric study population, because parks may facilitate nature contact and provide space for physical activity (141). Additionally, this measure is restricted to public areas rather than including private greenspace, reducing concern about whether or not participants actually have access to the greenspace.

Park boundaries were derived from ParkServe data compiled by the Trust for Public Land, downloaded on 10 March 2020 (67). There are 336 parks in Shelby County with 14,812 total acreage, which includes publicly owned local city, county, and state parks; school parks with a joint-use agreement with local government; and privately owned parks that are managed for full public use (Figure A4). Home Owner Association parks, golf courses, and cemeteries were not included.

Figure A4. Map of parks in Memphis, TN.



We calculated Euclidean distance from the geocoded participant residence (latitude and longitude) and the nearest boundary of a park polygon. In sensitivity analyses, we restricted the set of parks used in this calculation by size. Based on prior literature, we defined small parks (sometimes referred to as mini-parks or pocket parks) as those less than 2 acres, neighborhood parks as those 2-20 acres, and community parks as those >20 acres (142). Neighborhood parks may be more likely to include facilities and are typically intended to serve residents within approximately a 1-mile radius around the park. In Shelby County, there are 52 small parks, 194 neighborhood parks, and 90 community parks.

In primary analyses of park proximity and child behavior, park proximity to the address at the time of outcome ascertainment was used (Table A5). At the current address, 26.8% of the study population lived within 300m of a park. In sensitivity analyses, we also examined park proximity to the address at which the child had lived the longest. Park proximity at the current address was correlated with park proximity at the longest-lived address (0.60).

Table A5. Distribution of park proximity exposure metrics derived from ParkServe data.[a]

Exposure	Mean	SD	Min	Q1	Median	Q3	Max
All parks							
Distance to current address (m)	801	876	3	282	550	988	8417
Distance to longest-lived address (m)	811	862	0	306	560	997	8356
By park size at current address							
Small parks	4730	3093	73	2107	4408	6976	17075
Neighborhood parks	1381	1346	9	465	977	1826	8439
Community parks	1375	1098	3	595	1147	1840	9120
By urbanicity at current address							
Only for participants in urban census tracts	751	736	3	280	543	968	5450

[a] Participants included here who were included in analyses chapters 2 and 3 (N=934).

One limitation of the park exposure is that actual use of parks may largely be determined by factors other than just the natural environment characteristics of the park, including availability of park facilities such as playgrounds, as well as perceived safety of the park. However, a strength of using parks as an exposure is the relevancy for urban planning and investment in greening initiatives in public spaces.

Comparison of green space measures

Spearman correlations between the primary green space measures in chapter 1 (at the current address) and in chapters 2 and 3 (averaged over the year prior to outcome ascertainment at the age 4-6 visit) are shown in Table A6.

Table A6. Spearman correlations between primary green space measures at the current (age 4-6 visit) address and averaged across the residential history over the prior year.

		NDVI		Tree canopy	
		At current address	Averaged over prior year	At current address	Averaged over prior year
NDVI	At current address	1			
	Averaged over prior year	0.91	1		
Tree canopy	At current address	0.59	0.56	1	
	Averaged over prior year	0.55	0.60	0.92	1
Park Proximity	At current address	-0.01	-0.001	0.003	0.02

References

1. Hartig T, Mitchell R, de Vries S, Frumkin H. Nature and Health. Annu Rev Public Health. 2014 Mar 18;35(1):207–28.

2. Ekkel ED, de Vries S. Nearby green space and human health: Evaluating accessibility metrics. Landsc Urban Plan. 2017 Jan;157:214–20.

3. Labib SM, Lindley S, Huck JJ. Spatial dimensions of the influence of urban green-blue spaces on human health: A systematic review. Environ Res. 2020 Jan;180:108869.

4. Jarvis I, Koehoorn M, Gergel SE, van den Bosch M. Different types of urban natural environments influence various dimensions of self-reported health. Environ Res. 2020 Jul;186:109614.

5. Taylor L, Hochuli DF. Defining greenspace: Multiple uses across multiple disciplines. Landsc Urban Plan. 2017 Feb;158:25–38.

6. Markevych I, Schoierer J, Hartig T, Chudnovsky A, Hystad P, Dzhambov AM, et al. Exploring pathways linking greenspace to health: Theoretical and methodological guidance. Environ Res. 2017 Oct;158:301–17.

7. Mygind L, Kurtzhals M, Nowell C, Melby PS, Stevenson MP, Nieuwenhuijsen M, et al. Landscapes of becoming social: A systematic review of evidence for associations and pathways between interactions with nature and socioemotional development in children. Environ Int. 2021 Jan;146:106238.

8. Wolf KL, Lam ST, McKeen JK, Richardson GRA, van den Bosch M, Bardekjian AC. Urban Trees and Human Health: A Scoping Review. Int J Environ Res Public Health. 2020 Jun 18;17(12):4371.

9. Chawla L. Benefits of Nature Contact for Children. J Plan Lit. 2015 Nov;30(4):433–52.

10. Tillmann S, Tobin D, Avison W, Gilliland J. Mental health benefits of interactions with nature in children and teenagers: a systematic review. J Epidemiol Community Health. 2018 Oct;72(10):958–66.

11. Vanaken G-J, Danckaerts M. Impact of Green Space Exposure on Children's and Adolescents' Mental Health: A Systematic Review. Int J Environ Res Public Health. 2018 Nov 27;15(12):2668.

12. Achenbach TM, Ivanova MY, Rescorla LA, Turner LV, Althoff RR. Internalizing/Externalizing Problems: Review and Recommendations for Clinical and Research Applications. J Am Acad Child Adolesc Psychiatry. 2016 Aug;55(8):647–56.

13. Erskine HE, Norman RE, Ferrari AJ, Chan GCK, Copeland WE, Whiteford HA, et al. Long-Term Outcomes of Attention-Deficit/Hyperactivity Disorder and Conduct Disorder: A

Systematic Review and Meta-Analysis. J Am Acad Child Adolesc Psychiatry. 2016 Oct;55(10):841–50.

14. Althoff RR, Verhulst FC, Rettew DC, Hudziak JJ, van der Ende J. Adult Outcomes of Childhood Dysregulation: A 14-year Follow-up Study. J Am Acad Child Adolesc Psychiatry. 2010 Nov;49(11):1105-1116.e1.

15. Willcutt EG. The Prevalence of DSM-IV Attention-Deficit/Hyperactivity Disorder: A Meta-Analytic Review. Neurotherapeutics. 2012 Jul;9(3):490–9.

16. Landis TD, Garcia AM, Hart KC, Graziano PA. Differentiating Symptoms of ADHD in Preschoolers: The Role of Emotion Regulation and Executive Function. J Atten Disord. 2020 Jan 6;108705471989685.

17. Skogan AH, Zeiner P, Egeland J, Urnes A-G, Reichborn-Kjennerud T, Aase H. Parent ratings of executive function in young preschool children with symptoms of attention-deficit/-hyperactivity disorder. Behav Brain Funct. 2015 Dec;11(1):16.

18. Diamond A. Executive Functions. Annu Rev Psychol. 2013 Jan 3;64(1):135–68.

19. Thygesen M, Engemann K, Holst GJ, Hansen B, Geels C, Brandt J, et al. The Association between Residential Green Space in Childhood and Development of Attention Deficit Hyperactivity Disorder: A Population-Based Cohort Study. Environ Health Perspect. 2020 Dec;128(12):127011.

20. Dadvand P, Nieuwenhuijsen MJ, Esnaola M, Forns J, Basagaña X, Alvarez-Pedrerol M, et al. Green spaces and cognitive development in primary schoolchildren. PNAS. 2015;112(26):7937–42.

21. Dadvand P, Tischer C, Estarlich M, Llop S, Dalmau-Bueno A, López-Vicente M, et al. Lifelong Residential Exposure to Green Space and Attention: A Population-based Prospective Study. Environ Health Perspect. 2017 Sep 22;125(9):097016.

22. Kaplan S, Berman MG. Directed Attention as a Common Resource for Executive Functioning and Self-Regulation. Perspect Psychol Sci. 2010 Jan;5(1):43–57.

23. Ohly H, White MP, Wheeler BW, Bethel A, Ukoumunne OC, Nikolaou V, et al. Attention Restoration Theory: A systematic review of the attention restoration potential of exposure to natural environments. J Toxicol Environ Health Part B. 2016 Oct 2;19(7):305–43.

24. Stevenson MP, Schilhab T, Bentsen P. Attention Restoration Theory II: a systematic review to clarify attention processes affected by exposure to natural environments. J Toxicol Environ Health Part B. 2018 May 19;21(4):227–68.

25. Christiansen L, Beck MM, Bilenberg N, Wienecke J, Astrup A, Lundbye-Jensen J. Effects of Exercise on Cognitive Performance in Children and Adolescents with ADHD: Potential Mechanisms and Evidence-based Recommendations. J Clin Med. 2019 Jun 12;8(6):841.

26. Newman NC, Ryan P, LeMasters G, Levin L, Bernstein D, Hershey GKK, et al. Traffic-Related Air Pollution Exposure in the First Year of Life and Behavioral Scores at 7 Years of Age. Environ Health Perspect. 2013 Jun;121(6):731–6.

27. Liu J, Chen X, Lewis G. Childhood internalizing behaviour: analysis and implications: Childhood internalizing behaviour. J Psychiatr Ment Health Nurs. 2011 Dec;18(10):884–94.

28. Whalen DJ, Sylvester CM, Luby JL. Depression and Anxiety in Preschoolers. Child Adolesc Psychiatr Clin N Am. 2017 Jul;26(3):503–22.

29. Carneiro A, Dias P, Soares I. Risk Factors for Internalizing and Externalizing Problems in the Preschool Years: Systematic Literature Review Based on the Child Behavior Checklist 1½–5. J Child Fam Stud. 2016 Oct;25(10):2941–53.

30. Minh A, Muhajarine N, Janus M, Brownell M, Guhn M. A review of neighborhood effects and early child development: How, where, and for whom, do neighborhoods matter? Health Place. 2017 Jul;46:155–74.

31. Humphrey JL, Root ED. Spatio-temporal neighborhood impacts on internalizing and externalizing behaviors in U.S. elementary school children: Effect modification by child and family socio-demographics. Soc Sci Med. 2017 May;180:52–61.

32. Leventhal T, Brooks-Gunn J. Moving to Opportunity: an Experimental Study of Neighborhood Effects on Mental Health. Am J Public Health. 2003 Sep;93(9):1576–82.

33. Bezold CP, Banay RF, Coull BA, Hart JE, James P, Kubzansky LD, et al. The relationship between surrounding greenness in childhood and adolescence and depressive symptoms in adolescence and early adulthood. Ann Epidemiol. 2018 Apr;28(4):213–9.

34. Hartley K, Perazzo J, Brokamp C, Gillespie GL, Cecil KM, LeMasters G, et al. Residential surrounding greenness and self-reported symptoms of anxiety and depression in adolescents. Environ Res. 2021 Mar;194:110628.

35. Madzia J, Ryan P, Yolton K, Percy Z, Newman N, LeMasters G, et al. Residential Greenspace Association with Childhood Behavioral Outcomes. J Pediatr. 2019 Apr;207:233–40.

36. Amoly E, Dadvand P, Forns J, López-Vicente M, Basagaña X, Julvez J, et al. Green and Blue Spaces and Behavioral Development in Barcelona Schoolchildren: The BREATHE Project. Environ Health Perspect. 2014 Dec;122(12):1351–8.

37. Feng X, Astell-Burt T. Residential Green Space Quantity and Quality and Child Well-being: A Longitudinal Study. Am J Prev Med. 2017 Nov;53(5):616–24.

38. Nordbø ECA. Neighborhood green spaces, facilities and population density as predictors of activity participation among 8-year-olds: a cross-sectional GIS study based on the Norwegian mother and child cohort study. 2019;22.

39. Jennings V, Bamkole O. The Relationship between Social Cohesion and Urban Green Space: An Avenue for Health Promotion. Int J Environ Res Public Health. 2019 Feb 4;16(3):452.

40. Ulrich RS, Simons RF, Losito BD, Fiorito E, Miles MA, Zelson M. Stress recovery during exposure to natural and urban environments. J Environ Psychol. 1991 Sep;11(3):201–30.

41. Mygind L, Kjeldsted E, Hartmeyer R, Mygind E, Stevenson M, Quintana D, et al. Effects of Public Green Space on Acute Psychophysiological Stress Response: A Systematic Review and Meta-Analysis of the Experimental and Quasi-Experimental Evidence. Environ Behav. 2019;53(2):184–226.

42. Rigolon A, Browning MHEM, McAnirlin O, Yoon H (Violet). Green Space and Health Equity: A Systematic Review on the Potential of Green Space to Reduce Health Disparities. Int J Environ Res Public Health. 2021 Mar 4;18(5):2563.

43. Balseviciene B, Sinkariova L, Grazuleviciene R, Andrusaityte S, Uzdanaviciute I, Dedele A, et al. Impact of Residential Greenness on Preschool Children's Emotional and Behavioral Problems. Int J Environ Res Public Health. 2014 Jun 27;11(7):6757–70.

44. Flouri E, Midouhas E, Joshi H. The role of urban neighbourhood green space in children's emotional and behavioural resilience. J Environ Psychol. 2014 Dec;40:179–86.

45. Richardson EA, Pearce J, Shortt NK, Mitchell R. The role of public and private natural space in children's social, emotional and behavioural development in Scotland: A longitudinal study. Environ Res. 2017 Oct;158:729–36.

46. McEachan RRC, Yang TC, Roberts H, Pickett KE, Arseneau-Powell D, Gidlow CJ, et al. Availability, use of, and satisfaction with green space, and children's mental wellbeing at age 4 years in a multicultural, deprived, urban area: results from the Born in Bradford cohort study. Lancet Planet Health. 2018 Jun;2(6):e244–54.

47. Feng X, Astell-Burt T. The Relationship between Neighbourhood Green Space and Child Mental Wellbeing Depends upon Whom You Ask: Multilevel Evidence from 3083 Children Aged 12–13 Years. Int J Environ Res Public Health. 2017 Feb 27;14(3):235.

48. Palmer FB, Anand KJS, Graff JC, Murphy LE, Qu Y, Völgyi E, et al. Early Adversity, Socioemotional Development, and Stress in Urban 1-Year-Old Children. J Pediatr. 2013 Dec;163(6):1733-1739.e1.

49. Völgyi E, Carroll K, Hare M, Ringwald-Smith K, Piyathilake C, Yoo W, et al. Dietary Patterns in Pregnancy and Effects on Nutrient Intake in the Mid-South: The Conditions Affecting Neurocognitive Development and Learning in Early Childhood (CANDLE) Study. Nutrients. 2013 May 3;5(5):1511–30.

50. ECHO-PATHWAYS [Internet]. Available from: https://deohs.washington.edu/echo/

51. Sontag-Padilla L, Burns R, Shih R, Griffin B, Martin L, Chandra A, et al. The Urban Child Institute CANDLE Study: Methodological Overview and Baseline Sample Description [Internet]. RAND Corporation; 2015 [cited 2021 Apr 24]. Available from: http://www.rand.org/pubs/research_reports/RR1336.html

52. United States Census Bureau. 2010 American Community Survey (ACS) [Internet]. Available from: https://data.census.gov/cedsci/

53. Nowak DJ, Cumming AB, Twardus D, Hoehn RE, Oswalt CM, Brandeis TJ. Urban forests of Tennessee, 2009 [Internet]. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station; 2012 [cited 2021 Apr 19] p. SRS-GTR-149. Report No.: SRS-GTR-149. Available from: https://www.fs.usda.gov/treesearch/pubs/40246

54. Astell-Burt T, Feng X, Mavoa S, Badland HM, Giles-Corti B. Do low-income neighbourhoods have the least green space? A cross-sectional study of Australia's most populous cities. BMC Public Health. 2014 Dec;14(1):292.

55. Schüle SA. Relationship between neighbourhood socioeconomic position and neighbourhood public green space availability: An environmental inequality analysis in a large German city applying generalized linear models. Int J Hyg Environ Health. 2017;8.

56. Nesbitt L, Meitner MJ, Girling C, Sheppard SRJ, Lu Y. Who has access to urban vegetation? A spatial analysis of distributional green equity in 10 US cities. Landsc Urban Plan. 2019 Jan;181:51–79.

57. Ferguson M, Roberts HE, McEachan RRC, Dallimer M. Contrasting distributions of urban green infrastructure across social and ethno-racial groups. Landsc Urban Plan. 2018 Jul;175:136–48.

58. Dai D. Racial/ethnic and socioeconomic disparities in urban green space accessibility: Where to intervene? Landsc Urban Plan. 2011;11.

59. Gerrish E, Watkins SL. The relationship between urban forests and income: A meta-analysis. Landsc Urban Plan. 2018 Feb;170:293–308.

60. Engelberg JK, Conway TL, Geremia C, Cain KL, Saelens BE, Glanz K, et al. Socioeconomic and race/ethnic disparities in observed park quality. BMC Public Health. 2016 Dec;16(1):395.

61. Moore LV, Diez Roux AV, Evenson KR, McGinn AP, Brines SJ. Availability of Recreational Resources in Minority and Low Socioeconomic Status Areas. Am J Prev Med. 2008 Jan;34(1):16–22.

62. Jenkins G, Yuen H, Rose E, Maher A, Gregory K, Cotton M. Disparities in Quality of Park Play Spaces between Two Cities with Diverse Income and Race/Ethnicity Composition: A Pilot Study. Int J Environ Res Public Health. 2015 Jul 14;12(7):8009–22.

63. Wolch JR, Byrne J, Newell JP. Urban green space, public health, and environmental justice: The challenge of making cities 'just green enough.' Landsc Urban Plan. 2014 May;125:234–44.

64. Roy D, Zhang H. GWELD [Internet]. NASA Global Web-Enabled Landsat Data Annual Global 30 m V031 [Data set]. NASA EOSDIS Land Processes DAAC. 2019 [cited 2020 Oct 15]. Available from: https://doi.org/10.5067/MEaSUREs/GWELD/GWELDYR.031

65. Annerstedt van den Bosch M, Mudu P, Uscila V, Barrdahl M, Kulinkina A, Staatsen B, et al. Development of an urban green space indicator and the public health rationale. Scand J Public Health. 2016 Mar;44(2):159–67.

66. United States Environmental Protection Agency. EnviroAtlas [Internet]. EnviroAtlas. Community Metrics: Percent Tree Cover. [cited 2020 Nov 9]. Available from: enviroatlas.epa.gov/enviroatlas

67. The Trust for Public Land. ParkServe [Internet]. ParkServe Data Downloads. [cited 2020 Mar 10]. Available from: www.tpl.org/parkserve/downloads

68. Kersten EE, Adler NE, Gottlieb L, Jutte DP, Robinson S, Roundfield K, et al. Neighborhood Child Opportunity and Individual-Level Pediatric Acute Care Use and Diagnoses. Pediatrics. 2018 May;141(5):e20172309.

69. Aris IM, Rifas-Shiman SL, Jimenez MP, Li L-J, Hivert M-F, Oken E, et al. Neighborhood Child Opportunity Index and Adolescent Cardiometabolic Risk. Pediatrics. 2021 Feb;147(2):e2020018903.

70. Noelke C, McArdle N, Baek M, Huntington N, Huber R, Hardy E, et al. Childhood Opportunity Index 2.0 Technical Documentation [Internet]. 2020. Available from: diversitydatakids.org/research-library/research-brief/how-we-built-it

71. Acevedo-Garcia D, McArdle N, Hardy EF, Crisan UI, Romano B, Norris D, et al. The Child Opportunity Index: Improving Collaboration Between Community Development And Public Health. Health Aff (Millwood). 2014 Nov;33(11):1948–57.

72. Kramer MR. Residential Segregation and Health. 2018;49.

73. Getis, Ord - 1992 - The Analysis of Spatial Association by Use of Distance Statistics.pdf.

74. Jones MR, Diez-Roux AV, Hajat A, Kershaw KN, O'Neill MS, Guallar E, et al. Race/Ethnicity, Residential Segregation, and Exposure to Ambient Air Pollution: The Multi-Ethnic Study of Atherosclerosis (MESA). Am J Public Health. 2014 Nov;104(11):2130–7.

75. Ard K. By all measures: an examination of the relationship between segregation and health risk from air pollution. Popul Environ. 2016 Sep;38(1):1–20.

76. Czarnota J, Gennings C, Wheeler DC. Assessment of Weighted Quantile Sum Regression for Modeling Chemical Mixtures and Cancer Risk. Cancer Inform. 2015 Jan;14s2:CIN.S17295.

77. Carrico C, Gennings C, Wheeler DC, Factor-Litvak P. Characterization of Weighted Quantile Sum Regression for Highly Correlated Data in a Risk Analysis Setting. J Agric Biol Environ Stat. 2015 Mar;20(1):100–20.

78. Wheeler DC, Czarnota J, Jones RM. Estimating an area-level socioeconomic status index and its association with colonoscopy screening adherence. Green J, editor. PLOS ONE. 2017 Jun 8;12(6):e0179272.

79. Wheeler DC, Jones RM, Schootman M, Nelson EJ. Explaining variation in elevated blood lead levels among children in Minnesota using neighborhood socioeconomic variables. Sci Total Environ. 2019;8.

80. Casey J, James P, Cushing L, Jesdale B, Morello-Frosch R. Race, Ethnicity, Income Concentration and 10-Year Change in Urban Greenness in the United States. Int J Environ Res Public Health. 2017 Dec 10;14(12):1546.

81. Riley CB, Gardiner MM. Examining the distributional equity of urban tree canopy cover and ecosystem services across United States cities. Nedkov S, editor. PLOS ONE. 2020 Feb 11;15(2):e0228499.

82. Landry SM, Chakraborty J. Street Trees and Equity: Evaluating the Spatial Distribution of an Urban Amenity. Environ Plan. 2009;41:2651–70.

83. Krieger N, Williams DR, Moss NE. Measuring Social Class in US Public Health Research: Concepts, Methodologies, and Guidelines. Annu Rev Public Health. 1997 May;18(1):341–78.

84. Dobbins D, McCready M, Rackas L. Unequal Access: Barriers to Early Childhood Education for Boys of Color. :17.

85. Lin J, Wang Q, Li X. Socioeconomic and spatial inequalities of street tree abundance, species diversity, and size structure in New York City. Landsc Urban Plan. 2021 Feb;206:103992.

86. Baró F, Camacho DA, Pérez Del Pulgar C, Triguero-Mas M, Anguelovski I. School greening: Right or privilege? Examining urban nature within and around primary schools through an equity lens. Landsc Urban Plan. 2021 Apr;208:104019.

87. Nardone A, Rudolph KE, Morello-Frosch R, Casey JA. Redlines and Greenspace: The Relationship between Historical Redlining and 2010 Greenspace across the United States. Environ Health Perspect. 2021 Jan;129(1):017006.

88. Hoffman JS, Shandas V, Pendleton N. The Effects of Historical Housing Policies on Resident Exposure to Intra-Urban Heat: A Study of 108 US Urban Areas. Climate. 2020 Jan 13;8(1):12.

89. Volin E, Ellis A, Hirabayashi S, Maco S, Nowak DJ, Parent J, et al. Assessing macro-scale patterns in urban tree canopy and inequality. Urban For Urban Green. 2020 Nov;55:126818.

90. Wen M, Zhang X, Harris CD, Holt JB, Croft JB. Spatial Disparities in the Distribution of Parks and Green Spaces in the USA. Ann Behav Med. 2013 Feb;45(S1):18–27.

91. Rigolon A. Parks and young people: An environmental justice study of park proximity, acreage, and quality in Denver, Colorado. Landsc Urban Plan. 2017 Sep;165:73–83.

92. Williams TG, Logan TM, Zuo CT, Liberman KD, Guikema SD. Parks and safety: a comparative study of green space access and inequity in five US cities. Landsc Urban Plan. 2020 Sep;201:103841.

93. Choi D, Park K, Rigolon A. From XS to XL Urban Nature: Examining Access to Different Types of Green Space Using a 'Just Sustainabilities' Framework. 2020;25.

94. Schüle SA, Hilz LK, Dreger S, Bolte G. Social Inequalities in Environmental Resources of Green and Blue Spaces: A Review of Evidence in the WHO European Region. Int J Environ Res Public Health. 2019 Apr 4;16(7):1216.

95. Gascon M, Vrijheid M, Nieuwenhuijsen MJ. The Built Environment and Child Health: An Overview of Current Evidence. Curr Environ Health Rep. 2016 Sep;3(3):250–7.

96. de Greeff JW, Bosker RJ, Oosterlaan J, Visscher C, Hartman E. Effects of physical activity on executive functions, attention and academic performance in preadolescent children: a meta-analysis. J Sci Med Sport. 2018 May;21(5):501–7.

97. Taylor AF, Kuo FE, Sullivan WC. Views of nature and self-discipline: Evidence from inner city children. J Environ Psychol. 2002 Mar;22(1–2):49–63.

98. Hudziak JJ, Copeland W, Stanger C, Wadsworth M. Screening for DSM-IV externalizing disorders with the Child Behavior Checklist: a receiver-operating characteristic analysis. J Child Psychol Psychiatry. 2004 Oct;45(7):1299–307.

99. What are Equivalence Scales? OECD Project on Income Distribution and Poverty [Internet]. Available from: https://www.oecd.org/social/inequality.htm

100. Bijnens EM, Derom C, Thiery E, Weyers S, Nawrot TS. Residential green space and child intelligence and behavior across urban, suburban, and rural areas in Belgium: A longitudinal birth cohort study of twins. Markevych I, editor. PLOS Med. 2020 Aug 18;17(8):e1003213.

101. Lee M, Kim S, Ha M. Community greenness and neurobehavioral health in children and adolescents. Sci Total Environ. 2019 Jul;672:381–8.

102. Liao J, Yang S, Xia W, Peng A, Zhao J, Li Y, et al. Associations of exposure to green space with problem behaviours in preschool-aged children. Int J Epidemiol. 2020 Jun 1;49(3):944–53.

103. Andrusaityte S, Grazuleviciene R, Dedele A, Balseviciene B. The effect of residential greenness and city park visiting habits on preschool Children's mental and general health in Lithuania: A cross-sectional study. Int J Hyg Environ Health. 2020 Jan;223(1):142–50.

104. Markevych I, Tiesler CMT, Fuertes E, Romanos M, Dadvand P, Nieuwenhuijsen MJ, et al. Access to urban green spaces and behavioural problems in children: Results from the GINIplus and LISAplus studies. Environ Int. 2014 Oct;71:29–35.

105. Younan D, Tuvblad C, Li L, Wu J, Lurmann F, Franklin M, et al. Environmental Determinants of Aggression in Adolescents: Role of Urban Neighborhood Greenspace. J Am Acad Child Adolesc Psychiatry. 2016 Jul;55(7):591–601.

106. Markevych I, Tesch F, Datzmann T, Romanos M, Schmitt J, Heinrich J. Outdoor air pollution, greenspace, and incidence of ADHD: A semi-individual study. Sci Total Environ. 2018 Nov;642:1362–8.

107. Yang B-Y, Zeng X-W, Markevych I, Bloom MS, Heinrich J, Knibbs LD, et al. Association Between Greenness Surrounding Schools and Kindergartens and Attention-Deficit/Hyperactivity Disorder in Children in China. JAMA Netw Open. 2019 Dec 18;2(12):e1917862.

108. Donovan GH, Michael YL, Gatziolis D, Mannetje A 't, Douwes J. Association between exposure to the natural environment, rurality, and attention-deficit hyperactivity disorder in children in New Zealand: a linkage study. Lancet Planet Health. 2019 May;3(5):e226–34.

109. Flouri E, Papachristou E, Midouhas E. The role of neighbourhood greenspace in children's spatial working memory. Br J Educ Psychol. 2019 Jun;89(2):359–73.

110. Reuben A, Arseneault L, Belsky DW, Caspi A, Fisher HL, Houts RM, et al. Residential neighborhood greenery and children's cognitive development. Soc Sci Med. 2019 Jun;230:271–9.

111. Faber Taylor A, Kuo FE. Children With Attention Deficits Concentrate Better After Walk in the Park. J Atten Disord. 2009 Mar;12(5):402–9.

112. Schutte AR, Torquati JC, Beattie HL. Impact of Urban Nature on Executive Functioning in Early and Middle Childhood. Environ Behav. 2017 Jan;49(1):3–30.

113. Stevenson MP, Dewhurst R, Schilhab T, Bentsen P. Cognitive Restoration in Children Following Exposure to Nature: Evidence From the Attention Network Task and Mobile Eye Tracking. Front Psychol. 2019 Feb 5;10:42.

114. Gascon M, Triguero-Mas M, Martínez D, Dadvand P, Forns J, Plasència A, et al. Mental Health Benefits of Long-Term Exposure to Residential Green and Blue Spaces: A Systematic Review. Int J Environ Res Public Health. 2015 Apr 22;12(4):4354–79.

115. Fong K, Hart JE, James P. A Review of Epidemiologic Studies on Greenness and Health: Updated Literature Through 2017. 2018;19.

116. Zhang Y, Mavoa S, Zhao J, Raphael D, Smith M. The Association between Green Space and Adolescents' Mental Well-Being: A Systematic Review. Int J Environ Res Public Health. 2020 Sep 11;17(18):6640.

117. Norwood MF, Lakhani A, Fullagar S, Maujean A, Downes M, Byrne J, et al. A narrative and systematic review of the behavioural, cognitive and emotional effects of passive nature exposure on young people: Evidence for prescribing change. Landsc Urban Plan. 2019 Sep;189:71–9.

118. Bezold CP, Banay RF, Coull BA, Hart JE, James P, Kubzansky LD, et al. The Association Between Natural Environments and Depressive Symptoms in Adolescents Living in the United States. J Adolesc Health. 2018 Apr;62(4):488–95.

119. Alcock I, White MP, Wheeler BW, Fleming LE, Depledge MH. Longitudinal Effects on Mental Health of Moving to Greener and Less Green Urban Areas. Environ Sci Technol. 2014 Jan 21;48(2):1247–55.

120. Triguero-Mas M, Donaire-Gonzalez D, Seto E, Valentín A, Martínez D, Smith G, et al. Natural outdoor environments and mental health: Stress as a possible mechanism. Environ Res. 2017 Nov;159:629–38.

121. Bratman GN, Young G, Mehta A, Lee Babineaux I, Daily GC, Gross JJ. Affective Benefits of Nature Contact: The Role of Rumination. Front Psychol. 2021 Mar 10;12:643866.

122. Dadvand P, Hariri S, Abbasi B, Heshmat R, Qorbani M, Motlagh ME, et al. Use of green spaces, self-satisfaction and social contacts in adolescents: A population-based CASPIAN-V study. Environ Res. 2019 Jan;168:171–7.

123. Van Aart CJC, Michels N, Sioen I, De Decker A, Bijnens EM, Janssen BG, et al. Residential landscape as a predictor of psychosocial stress in the life course from childhood to adolescence. Environ Int. 2018 Nov;120:456–63.

124. Berto R. The Role of Nature in Coping with Psycho-Physiological Stress: A Literature Review on Restorativeness. Behav Sci. 2014 Oct 21;4(4):394–409.

125. Li D, Sullivan WC. Impact of views to school landscapes on recovery from stress and mental fatigue. Landsc Urban Plan. 2016 Apr;148:149–58.

126. Mennis J, Mason M, Ambrus A. Urban greenspace is associated with reduced psychological stress among adolescents: A Geographic Ecological Momentary Assessment (GEMA) analysis of activity space. Landsc Urban Plan. 2018 Jun;174:1–9.

127. Franklin M, Yin X, McConnell R, Fruin S. Association of the Built Environment With Childhood Psychosocial Stress. JAMA Netw Open. 2020 Oct 21;3(10):e2017634.

128. Jimenez MP, Oken E, Gold DR, Luttmann-Gibson H, Requia WJ, Rifas-Shiman SL, et al. Early life exposure to green space and insulin resistance: An assessment from infancy to early adolescence. Environ Int. 2020 Sep;142:105849.

129. Wolraich ML, Hagan JF, Allan C, Chan E, Davison D, Earls M, et al. Clinical Practice Guideline for the Diagnosis, Evaluation, and Treatment of Attention-Deficit/Hyperactivity Disorder in Children and Adolescents. Pediatrics. 2019 Oct;144(4):e20192528.

130. Kondo MC, Oyekanmi KO, Gibson A, South EC, Bocarro J, Hipp JA. Nature Prescriptions for Health: A Review of Evidence and Research Opportunities. Int J Environ Res Public Health. 2020 Jun 12;17(12):4213.

131. Studio Gang Architects. Memphis Riverfront Concept [Internet]. 2017 [cited 2021 Jun 21]. Available from: https://www.memphisriverparks.org

132. Jackson SB, Stevenson KT, Larson LR, Peterson MN, Seekamp E. Outdoor Activity Participation Improves Adolescents' Mental Health and Well-Being during the COVID-19 Pandemic. Int J Environ Res Public Health. 2021 Mar 3;18(5):2506.

133. Mitra R, Moore SA, Gillespie M, Faulkner G, Vanderloo LM, Chulak-Bozzer T, et al. Healthy movement behaviours in children and youth during the COVID-19 pandemic: Exploring the role of the neighbourhood environment. Health Place. 2020 Sep;65:102418.

134. Jarvis I, Gergel S, Koehoorn M, van den Bosch M. Greenspace access does not correspond to nature exposure: Measures of urban natural space with implications for health research. Landsc Urban Plan. 2020 Feb;194:103686.

135. James P, Banay RF, Hart JE, Laden F. A Review of the Health Benefits of Greenness. Curr Epidemiol Rep. 2015 Jun;2(2):131–42.

136. Robinson N, Allred B, Jones M, Moreno A, Kimball J, Naugle D, et al. A Dynamic Landsat Derived Normalized Difference Vegetation Index (NDVI) Product for the Conterminous United States. Remote Sens. 2017 Aug 21;9(8):863.

137. Smith G, Cirach M, Swart W, Dėdelė A, Gidlow C, van Kempen E, et al. Characterisation of the natural environment: quantitative indicators across Europe. Int J Health Geogr. 2017 Dec;16(1):16.

138. Wheeler BW, Lovell R, Higgins SL, White MP, Alcock I, Osborne NJ, et al. Beyond greenspace: an ecological study of population general health and indicators of natural environment type and quality. Int J Health Geogr. 2015;14:17.

139. Reid CE, Clougherty JE, Shmool JLC, Kubzansky LD. Is all urban green space the same? A comparison of the health benefits of trees and grass in New York city. Int J Environ Res Public Health. 2017;14(11).

140. Homer CG, Dewitz J, Yang L, Jin S, Danielson P, Xian, et al. Completion of the 2011 National Land Cover Database for the conterminous United States – representing a decade of land cover change information. Photogramm Eng Remote Sens. 2015;81:345–53.

141. Schipperijn J, Ekholm O, Stigsdotter UK, Toftager M, Bentsen P, Kamper-JÃ F. Factors influencing the use of green space: Results from a Danish national representative survey. Landsc Urban Plan. 2010;8.

142. Cohen DA, Han B, Nagel CJ, Harnik P, McKenzie TL, Evenson KR, et al. The First National Study of Neighborhood Parks. Am J Prev Med. 2016 Oct;51(4):419–26.