Cities & Health

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rcah20

Types of greenspace and adolescent mental health and well-being in metropolitan London

Marie A. E. Mueller, Emily Midouhas & Eirini Flouri

To cite this article: Marie A. E. Mueller, Emily Midouhas & Eirini Flouri (2023): Types of greenspace and adolescent mental health and well-being in metropolitan London, Cities & Health, DOI: 10.1080/23748834.2023.2175410

To link to this article: https://doi.org/10.1080/23748834.2023.2175410

CITIES & HEALTH
https://doi.org/10.1080/23748834.2023.2175410

ORIGINAL SCHOLARSHIP

OPEN ACCESS Check for updates

Types of greenspace and adolescent mental health and well-being in metropolitan London

Marie A. E. Mueller[a], Emily Midouhas[b] and Eirini Flouri[b]

[a]Department of Epidemiology and Applied Clinical Research, UCL Division of Psychiatry, University College London, London, UK; [b]Department of Psychology and Human Development, UCL Institute of Education, University College London, London, UK

ABSTRACT

The evidence suggests a link between greenspace and adolescent mental health. One limitation is the typically crude measure of greenspace quantity or greenness. We explored the roles of different types of greenspace in the mental health of 10- to 15-year-old adolescents living in London, using data from Understanding Society, a UK household longitudinal study. We used data on 1,879 adolescents from waves 1-8 (2009-2018). As some adolescents had observations at multiple waves, 4,217 observations were included. Mental health and well-being measures were Strengths and Difficulties Questionnaire scores, self-esteem, and happiness. Proportions of green land cover, parks & gardens, natural & semi-natural urban greenspaces, outdoor sports facilities, and total green land use were measured in 500 m around postcodes. We ran linear regressions, stratified by age, adjusted for confounders, and accounting for Understanding Society's complex sampling design. We did not find consistent results across analyses, but we identified patterns worth exploring further: older adolescents (13-15 years) seemed to 'benefit' more from greenspace than younger adolescents (10-12 years); and parks & gardens and outdoor sports facilities seemed to be most 'beneficial'. Overall, however, no clear conclusions can be drawn, and findings need to be confirmed in future studies.

ARTICLE HISTORY
Received 9 June 2022
Accepted 28 January 2023

KEYWORDS
Greenspace; park; neighbourhood; urban; adolescence; mental health

Introduction

Most cases of mental disorders start before the age of 24 years (Kessler *et al.* 2005, de Girolamo *et al.* 2012). This makes adolescence a critical period for prevention and intervention, and it is important to identify modifiable factors that are either risk factors or positive (protective and promotive) factors for adolescent mental health and well-being. Greenspace is a factor of the physical environment that is thought to have a positive effect on mental health and well-being, for example, by supporting attention restoration and stress recovery, and by encouraging physical and social activities (Markevych *et al.* 2017). Therefore, greenspace is a promising factor for protecting and promoting adolescent mental health and well-being. The focus of the present study is to explore the role of neighbourhood greenspace in the mental health and well-being of adolescents living in London, the largest city in the UK. We focus on greenspace in the residential neighbourhood which we consider a key area of exposure. Because children and adolescents in the UK tend to have limited independent mobility (Shaw *et al.* 2015), they may spend much of their spare time in the immediate areas around their homes. This makes modifiable factors of the residential neighbourhood, such as greenspace, a promising target for prevention and intervention.

In 2017, about 13% of children and adolescents between ages 5 and 19 years in England had a mental disorder, and adolescents were more likely to have a mental disorder than younger children. In 11- to 16-year-olds, the rate was over 14% (NHS Digital 2018). Considering that this number only includes adolescents who fulfil criteria for a clinical diagnosis, we must assume that considerably more suffer from mental health problems (Ford and Parker 2016). As early mental health problems have a great impact on both the individual (e.g. well-being and life chances) and society (e.g. costs for health, educational, or criminal justice systems; Suhrcke *et al.* 2007, Christensen *et al.* 2020), it is important to identify the factors that may influence adolescent mental health and well-being. Greenspace is a physical environment factor that may have a positive effect.

Greenspace can impact mental health via three main pathways: 'reducing harm' (mitigation), 'restoring capacities' (restoration), and 'building capacities' (instoration; Markevych *et al.* 2017). First, greenspace

CONTACT Marie A. E. Mueller marie.mueller.16@ucl.ac.uk Department of Epidemiology and Applied Clinical Research, UCL Division of Psychiatry, University College London, Maple House, 149 Tottenham Ct Rd, London W1T 7NF, United Kingdom
Supplemental data for this article can be accessed online at https://doi.org/10.1080/23748834.2023.2175410

can mitigate the harmful effects of environmental stressors, such as air pollution and noise, by reducing levels of these stressors in the environment (Nguyen *et al.* 2015, Li *et al.* 2020), and by moderating their harmful effects (Dzhambov *et al.* 2018). Second, greenspace can promote the restoration of both cognitive and affective resources (Stevenson *et al.* 2018, Bratman *et al.* 2021). Third, greenspace can encourage engagement in physical and social activities. In addition to these often-mentioned pathways, there are other ways in which greenspace can affect health, for example, by affecting the human microbiome (Mills *et al.* 2017). Therefore, in theory, a relationship between neighbourhood greenspace and adolescent mental health is plausible. However, the studies investigating this relationship have some important limitations. We will now review the evidence to date, drawing from research on children and adolescents. The review will highlight that one of the main limitations of the existing literature is the usually crude measurement of neighbourhood greenspace or greenness. One of the main contributions of the present study is the inclusion of measures of different types of green spaces (i.e. parks and gardens, natural and semi-natural urban greenspaces, and outdoor sports facilities), adding nuance to our understanding of the association between greenspace and adolescent mental health and well-being.

Review of the literature

Most of the studies to date have investigated the association of the availability of greenspace in, or the average greenness of, the neighbourhood with mental health and well-being. Many studies used the Normalised Difference Vegetation Index (NDVI) to assess the average greenness of a given area. Studies using the NDVI typically find a positive link between greenness and mental health. For example, Amoly *et al.* (2015) found that higher levels of greenness around the home were associated with lower levels of emotional and behavioural problems in 7- to 10-year-old children from Barcelona (Spain). Although some of the associations found depended on the buffer size used in the analysis (i.e. 100 m, 250 m, and 500 m), overall, the link between greenness and mental health seemed to be robust. Bezold *et al.* (2018) found a similar association between residential greenness and depressive symptoms in 12- to 18-year-old adolescents from the United States: higher levels of greenness were associated with lower odds of depressive symptoms. The researchers measured greenness in both a 250 m buffer and a 1,250 m buffer around the home, but the negative association with depressive symptoms was only found in the 1,250-m-buffer analysis. This might be explained by the fact that adolescents are typically allowed to roam in a wider radius around their home than younger children. In a study on both children (7 years old) and adolescents (12 years old), Madzia *et al.* (2019) found associations of residential greenness (measured in 200 m, 400 m, and 800 m around the home) with internalising and externalising problems. Again, results depended on the buffer size used in the analysis; however, overall, the study found a negative association of residential greenness with externalising symptoms (e.g. hyperactivity) in 7-year-olds, and with internalising symptoms (e.g. depression) in 12-year-olds, which suggests that the same exposure can have different effects in different age groups.

In addition to the three studies above, studies also found associations of residential greenness with depressive symptoms, psychological distress (Mavoa *et al.* 2019, Wang *et al.* 2019), and aggressive behaviour (Younan *et al.* 2016) in adolescents; and with attention deficit/hyperactivity disorder (ADHD) later in life (Thygesen *et al.* 2020). There have been some notable exceptions to this pattern, however. Some studies did not find a clear link between residential greenness and children's and adolescents' mental health (Markevych *et al.* 2014, Maes *et al.* 2021), and one study found a negative association, though only in socio-economically advantaged children (Balseviciene *et al.* 2014).

Rather than using the NDVI, some studies used measures of green land cover or green land use to assess the link between greenspace availability and mental health. These studies suggest links between quantity of parkland and perceived stress in adolescents (Feda *et al.* 2015) as well as internalising and externalising symptoms from early childhood to early adolescence (Feng and Astell-Burt 2017). One study that is particularly relevant in the context of the present investigation (and will be described in more detail below) found that higher levels of woodland were associated with lower levels of mental health problems in adolescents living in London (Maes *et al.* 2021). This association appeared to be stronger for larger buffers (250 m and 500 m) than for smaller buffers (50 m and 100 m). Again, there have been null findings too (Flouri *et al.* 2014, Richardson *et al.* 2017, Mueller *et al.* 2019, Mueller and Flouri 2020, 2021).

Some studies have investigated dimensions of exposure other than quantity. There is some evidence that the proximity to greenspaces (Balseviciene *et al.* 2014, Markevych *et al.* 2014), the use of/time spent in green spaces (Flouri *et al.* 2014, Amoly *et al.* 2015), and the quality of green spaces (Feng and Astell-Burt 2017) may also play a role in child and adolescent mental health.

A few studies to date have investigated the roles of different types of greenery or greenspace. For example, Richardson *et al.* (2017) used measures of the quantity of public parks and the quantity of (any) natural space (including all public and private natural spaces). However, neither was associated with children's mental health. Markevych *et al.* (2014) investigated the association of proximity of urban green spaces (i.e. cemeteries, gardens, parks, and plant nurseries) with the mental health of 10-year-old children from Munich, Germany. In a sensitivity analysis, they also tested the effects of proximity of forests (rather than urban green spaces) and residential greenness. They found a significant effect only for proximity of urban green spaces (not proximity of forests or residential greenness). Probably the most relevant study in the context of this paper was published by Maes *et al.* (2021). They investigated the association of natural space with the mental health and well-being of adolescents (9 to 15 years) living in London. In separate models, they included measures of natural space; green space and blue space; and grassland and woodland (defined as vegetation lower and higher than 1 m, respectively). They found that higher levels of woodland (but not grassland) were associated with lower levels of emotional and behavioural problems. Higher levels of grassland (but not woodland) were associated with lower odds of low overall well-being.

Maes *et al.*'s (2021) study suggests that woodland may be more important for adolescent mental health than grassland. On the contrary, Markevych *et al.*'s (2014) study suggests that forests may not be as relevant as urban green spaces (e.g. parks) for the mental health of young adolescents. Of course, one must note several differences between the two studies. For example, Maes *et al.* (2021) used a measure of quantity, whereas Markevych *et al.* (2014) used a measure of proximity; Maes *et al.* (2021) defined woodlands as vegetation higher than 1 m, whereas Markevych *et al.* (2014) used land use data to define forests; adolescents in Markevych *et al.*'s (2014) study were younger; finally, the two studies took place in different cities (of vastly different population sizes) and, indeed, countries. Although we cannot know how much each of these differences contributes to the findings, it is plausible that several factors play a role, including the measure of exposure and the adolescents' age. In fact, it is important to keep in mind that many factors may play a role in the association between neighbourhood greenspace and adolescent mental health and well-being. For example, sex and education may affect adolescents' visits to, and usage of, green spaces (Bloemsma *et al.* 2018), and other environmental factors, such as private garden access, may moderate the association between neighbourhood greenspace and adolescent well-being (Mueller and Flouri 2021). Therefore, although not the focus of the present study, it is important to keep in mind that the association between neighbourhood greenspace and adolescent mental health and well-being is complex, and that factors at several levels (e.g. individual, family, and neighbourhood) may play a role in it.

The present study

We carried out this study to address some of the limitations described above. We investigated the role of different types of neighbourhood greenspace in the mental health (i.e. emotional and behavioural problems) and well-being (i.e. self-esteem and happiness) of 10- to 15-year-old adolescents living in London, UK. We defined the residential neighbourhood as a circular buffer with a radius of 500 m around a postcode centroid. To explore the role of different types of greenspace, we distinguished between parks and gardens; natural and semi-natural urban greenspaces; and outdoor sports facilities. We also included a measure of total green land cover. We tested the associations of the different exposures with mental health and well-being separately for each of the six age groups. In addition, we ran several sensitivity analyses to assess different neighbourhood and exposure definitions. Note that we use the term 'exposure' to refer to the main variable of interest (i.e. greenspace) that may influence outcomes (i.e. mental health and well-being). In other disciplines, 'exposure' may be referred to as 'factor' or 'independent variable'. It is important to highlight that exposure variables in our study do not necessarily capture adolescents' true exposure to greenspace. In other words, our exposure variables are proxies for adolescents' true exposure.

Methods

Study area

In this project, we focused on the London region which consists of Greater London and the City of London. London is located in the southeast of England and has 33 local government districts: 32 Greater London boroughs and the City of London. These 33 districts are divided into 4,835 Lower Layer Super Output Areas (LSOAs). A LSOA is a unit of Census geography in the UK with a minimum population of 1,000, a maximum population of 3,000, and an average population of 1,500. In this study, we used two levels of geography for the calculation of green land cover and green land use variables: LSOAs and postcodes. For exposures at LSOA level, we used LSOA boundaries from the 2011 Census. For exposures at postcode level, we calculated circular buffers with a radius of 500 m around postcodes. In a sensitivity analysis, we used buffers with radii of 300 m and 1,000 m. Note that in the Study Variables section below, when we describe how we calculated our exposure variables, we will describe this in separate sub-sections for LSOAs and postcodes.

Study sample

We used data from Understanding Society (University of Essex, Institute for Social and Economic Research 2020, 2021, 2022). Understanding Society is the UK Household Longitudinal Study (UKHLS) and has data on the members of approximately 40,000 UK households at wave 1 (2009-2011). Up to this day, households have been followed for ten waves. More information about the UKHLS data and study design is provided in the UKHLS user guide (Institute for Social and Economic Research 2020). In this study, we used data from waves 1 to 8 (2009-2018). In particular, we used data from the youth dataset, based on self-reports from 10- to 15-year-olds. To these data, we linked information on parents, families, and neighbourhoods. Mental health and well-being outcomes considered in this study – Strengths and Difficulties Questionnaire (SDQ) scores, self-esteem, and happiness – were measured at waves 1, 3, 5, and 7 (SDQ); waves 2, 4, 6, and 8 (self-esteem); and waves 1 to 8 (happiness). Across waves 1 to 8, there were 32,404 observations clustered in 12,675 adolescents. Our analytic sample included those adolescents who lived in London, had valid data on at least one of the three study outcomes for at least one wave, and had a non-zero study weight ($n = 1{,}879$). Key differences between analytic and non-analytic samples are summarised in Table 1 and described under Descriptive Statistics in the Results section. As some of the 1,879 adolescents had data for more than one wave, there were 4,217 observations included in the analysis. Note, however, that numbers depended on the outcome measured (because outcomes were measured at different waves). The mean age of the study sample was 12.42 years, and around 51% were female. For more information on the study sample, please see Table 1. Of the 4,217 observations, 74% had complete data, and 26% had missing data on at least one variable. To minimise bias and retain as many observations as possible, we used multiple imputation by chained equations (which is described in the Statistical Analysis section below).

Study variables

In the following section, we will describe outcomes (dependent variables) and exposures (independent variables). Our outcomes were mental health and well-being; our exposures were green land use and green land cover (described in two separate sections below). According to the two levels of geography used in this study, we calculated percentages of green land use and green land cover for both LSOAs and postcodes. After describing outcomes and exposures, we will describe neighbourhood-level confounders, and family- and child-level confounders. Note that outcomes, and family- and child-level variables, were taken from Understanding Society. Exposures and other neighbourhood-level variables were taken from other sources, as described in detail below.

Outcomes (mental health and well-being)

Mental health was measured with the self-reported Strengths and Difficulties Questionnaire (SDQ) at waves 1, 3, 5, and 7. The SDQ is a validated, widely used index of emotional and behavioural difficulties (i.e. emotional symptoms, conduct problems, hyperactivity and inattention, and peer relationship problems; Goodman 1997, Goodman *et al*. 1998). Each of the SDQ subscales includes five items which are rated on a scale ranging from 0 ('not true') to 2 ('certainly true'). Example items for *emotional symptoms* are, 'I worry a lot', and, 'I have many fears'; example items for *conduct problems* are, 'I get very angry and often lose my temper', and, 'I fight a lot'; example items for *hyperactivity and inattention* are, 'I am restless', and, 'I am easily distracted'; and example items for *peer relationship problems* are, 'I am usually on my own', and, 'I get on better with adults than with people my age'. The scores for each subscale range between 0 and 10, and the 20 items of the four subscales can be combined into a total difficulties score ranging from 0 to 40. In the analytic sample at wave 1, the Cronbach's alpha values were 0.65 (emotional symptoms), 0.61 (conduct problems), 0.64 (hyperactivity and inattention), 0.53 (peer relationship problems), and 0.76 (total difficulties). Due to the scale's low internal consistency, results regarding peer relationship problems should be taken with caution.

Mental well-being was measured with two scales: self-esteem and happiness. Self-esteem was measured with eight items based on the Rosenberg self-esteem scale (Rosenberg 1965) at waves 2, 4, 6, and 8. Example items are, 'I feel I have a number of good qualities', and, 'I am a likeable person'. Each item was rated on a scale from 1 ('strongly disagree') to 4 ('strongly agree'). The self-esteem scale score is the mean of the eight items. The Cronbach's alpha value for the self-esteem scale was 0.76 at wave 2. Happiness was measured with six items at waves 1 to 8: 'How do you feel about (a) your schoolwork, (b) your appearance, (c) your family, (d) your friends, (e) your school, and (f) your life as a whole?' This scale has been used in previous studies as a measure of mental well-being (Bannink *et al*. 2016, Kelly *et al*. 2016, 2018, Mueller

Table 1. Bias analysis between analytic and non-analytic samples.

	Analytic sample (n = 4,217)		Non-analytic sample (n = 28,187)		Test
Continuous variables	n	M (SD)	n	M (SD)	F
SDQ CP (0-10)	2,281	2.13 (1.73)	14,119	2.21 (1.81)	1.54
SDQ ES (0-10)	2,281	2.64 (2.14)	14,118	2.88 (2.26)	7.28 **
SDQ HA (0-10)	2,279	3.70 (2.18)	14,115	4.06 (2.34)	22.18 ***
SDQ PP (0-10)	2,279	1.61 (1.54)	14,120	1.82 (1.70)	16.01 ***
SDQ TD (0-40)	2,276	10.09 (5.29)	14,106	10.97 (5.80)	17.79 ***
Self-esteem (1-4)	1,881	3.17 (0.44)	13,819	3.11 (0.44)	8.65 **
Happiness (1-7)	4,202	5.89 (0.87)	28,072	5.81 (0.85)	6.61 *
Green land cover [%][1]	4,123	38.54 (11.79)	-	-	-
Green land use [%][1]	4,123	8.69 (9.25)	-	-	-
Parks/gardens [%][1]	4,123	4.85 (6.97)	-	-	-
Natural/semi-natural spaces [%][1]	4,123	2.06 (6.10)	-	-	-
Outdoor sports facilities [%][1]	4,123	1.79 (3.93)	-	-	-
Air pollution [mean NO_2][1]	4,123	35.32 (5.65)	-	-	-
Area deprivation [Carstairs z-score]	4,217	1.90 (3.49)	23,358	−0.32 (3.04)	135.32 ***
Maternal psychological distress (0-36)	3,203	11.76 (5.96)	24,613	11.96 (5.90)	0.75
Age [years]	4,217	12.42 (1.69)	28,187	12.54 (1.69)	9.59 **
Categorical variables	n	%	n	%	F[2]
University education (mother)	1,560	41.42	10,308	40.05	0.29
Family owns its home	1,805	48.38	19,498	66.45	35.44 ***
Intact family structure	2,826	64.18	18,073	62.77	0.26
Ethnicity White	1,182	53.58	23,262	90.65	359.93 ***
Ethnicity Mixed	534	10.23	1,031	2.85	82.07 ***
Ethnicity Indian	310	6.29	824	1.76	54.23 ***
Ethnicity Pakistani and Bangladeshi	824	7.64	1,980	2.75	47.27 ***
Ethnicity Black or Black British	1,097	17.49	692	1.03	567.74 ***
Ethnicity Other	270	4.77	340	0.96	61.23 ***
Female	2,120	50.75	14,025	49.50	0.36

Note. M = mean; SD = standard deviation; CP = conduct problems; ES = emotional symptoms; HA = hyperactivity/inattention; PP = peer problems; TD = total difficulties. Data are taken from waves 1 to 8. Sample sizes refer to observations (not individuals). Some individuals have multiple observations across waves, and these multiple observations are included in the descriptive statistics above. Descriptive statistics by wave and by age group differ slightly, but the overall descriptive statistics in this table give an appropriate overview of the sample characteristics. Ns are unweighted. Ms, SDs, and %s are weighted. [1] Values are for 500 m buffers around postcodes. [2] Design-based F statistic (i.e. corrected weighted Chi^2 statistic). * $p < .05$, ** $p < .01$, *** $p < .001$.

and Flouri 2021). Each item was rated on a scale ranging from 1 ('not at all happy') to 7 ('completely happy'). The happiness scale score was the mean of the six items. The Cronbach's alpha value for the happiness scale was 0.76 at wave 1.

Exposures (green land use and green land cover)

We used two types of data to measure greenspace in the neighbourhood: green land use and green land cover. Both types of data, and how they were used to calculate percentages of green land use and green land cover at LSOA and postcode levels, will be described in detail below. The main difference between the green land use and green land cover variables is their focus on different types of information. The green land use variables are based on open space data and include clearly defined and delineated open spaces in London. They distinguish between different types of open (green) spaces and, therefore, provide information on the function of these spaces. The green land cover variable is based on a combination of satellite imagery and land use data. The green land cover variable does not distinguish between different types of green spaces or vegetation and, therefore, does not provide information on the function of spaces. However, it captures even small areas of vegetation and provides information about how much of London's area is covered with (any) vegetation. It can be used as an indicator of greenness. Because green land use and green land cover variables capture different aspects of greenspace (which may be associated with adolescent mental health and well-being via different pathways), we included both as indicators of neighbourhood greenspace. We will now describe how we calculated green land use and green land cover variables at LSOA and postcode levels.

Green land use

We used data from Greenspace Information for Greater London (GiGL; 'Greenspace Information for Greater London CIC – GIGL', 2022). GiGL works with the Greater London Authority (GLA) and the London boroughs to curate and share data on London's natural environments. The GiGL open space dataset includes information on 12,781 open spaces (version 2020/2021). The open spaces are categorised into 11 categories that are based on the 2002 Planning Policy Guidance 17 (PPG17; 'Planning Policy Guidance 17:

Planning for open space, sport and recreation', 2012): parks & gardens; natural & semi-natural urban greenspaces; green corridors; outdoor sports facilities; amenity; children & teenagers; allotments, community gardens, & city farms; cemeteries & churchyards; other urban fringe; civic spaces; and other. For a detailed description of categories and subcategories, please visit the GiGL website. In addition to information on the type of open space, the dataset includes information on other attributes, such as type of access (e.g. free, de facto, or restricted).

Arguably, not all of the 12,781 open spaces are relevant for adolescents. To decide what open spaces to include, we considered the general functions of greenspace proposed by Markevych *et al.* (2017): reducing harm, restoring capacities, and building capacities. In theory, most types of open space would support at least one of the three pathways. However, many open spaces in London are not public, as they have restricted or no access (e.g. private woodlands, allotments, and equestrian centres). Therefore, adolescents may not use these spaces. We therefore included only those open spaces that could be accessed and used by adolescents (i.e. we excluded open spaces with restricted or no access). Applying this criterion, we retained 5,845 open spaces. Please see Figure 1 for maps of open spaces (top panel) and 'free access' open spaces (bottom panel) in London.

We further reduced open spaces included in this study to three main categories: parks and gardens; natural and semi-natural urban greenspaces; and outdoor sports facilities. These are the largest open space categories, and all three are associated with different levels of greenery and different functions. Natural spaces tend to be the most natural and biodiverse. Parks and gardens are more formal and function as spaces for recreation. Outdoor sports facilities are the least natural and function mainly as spaces for activity and exercise. A similar approach to using GiGL data was taken by Houlden *et al.* (2021) who investigated the role of greenspace in the mental well-being of adults living in London. Focusing on the three open space categories further reduced the number of open spaces to 2,521 (i.e. 1,327 parks or gardens; 522 natural or semi-natural urban greenspaces; and 672 outdoor sports facilities). Note that the excluded 3,324 open spaces were open spaces from the other eight categories listed above (e.g. amenity, cemeteries & churchyards, and civic spaces).

Green land use at LSOA level. We calculated three proportions for each LSOA, using R (in RStudio). We used functions from the *tidyverse* package (Wickham *et al.* 2019) and the *sf* package (Pebesma 2018) to clean data, make geometries topologically valid, and calculate exposures. GiGL open space data and LSOA data shared the same Coordinate Reference System (CRS), the British National Grid (a projected CRS). We calculated the intersections of open spaces and LSOAs, and the area of these intersections (in m^2). We used this approach for each of the three subsets to calculate how much area of each LSOA was covered with each of the three types of space (in %): parks and gardens; natural and semi-natural urban greenspaces; and outdoor sports facilities. Further, we calculated a green land use measure that combined the three types of space into one green land use variable. Please see Figure 2 (top panel) for a visualisation of the proportions of 'free access' green land use for London LSOAs.

Green land use at postcode level. In addition to calculating exposures at LSOA level, we calculated exposures at postcode level. For each adolescent in our study sample, we had a postcode grid reference which we used as a proxy for their home. Using the *sf* package in R, we calculated circular buffers around the grid references and spatially intersected these buffers with open spaces. We used a 500 m radius to calculate circular buffers. In the literature, buffer sizes often range between 100 m and 1,000 m (although there are also studies using smaller or larger buffers). We decided to use a 500 m radius for two reasons. First, smaller radii were associated with high numbers of zero proportions of green land use. Second, larger radii may be too large to represent adolescents' actual activity spaces. As it is still unclear what radius best approximates actual exposure, and whether green spaces may have a different relevance within different distances, we also ran sensitivity analyses using 300 m and 1,000 m buffers (reported in the supplementary material).

Green land cover

In addition to GiGL data, we used London Green and Blue Cover data, provided by the GLA ('London Green and Blue Cover – London Datastore', 2018). Green and blue cover data combine 2016 NDVI and land use data, providing information on London's total natural space (including all areas of public and private green and blue space). Each green and blue area is captured as a polygon. For this study, we used data on green cover only, capturing even small areas of vegetation (such as trees, private gardens, and green roofs) and, therefore, measuring the overall greenness of an area. The green cover data are a good complement to the GiGL data (which give information on the type of space but do not provide information on greenness). For more information about the London Green and Blue Cover data, please visit the London Datastore.

Green land cover at LSOA level. To calculate the proportion of green land cover for each LSOA, we first had to manage the large size of the data. After

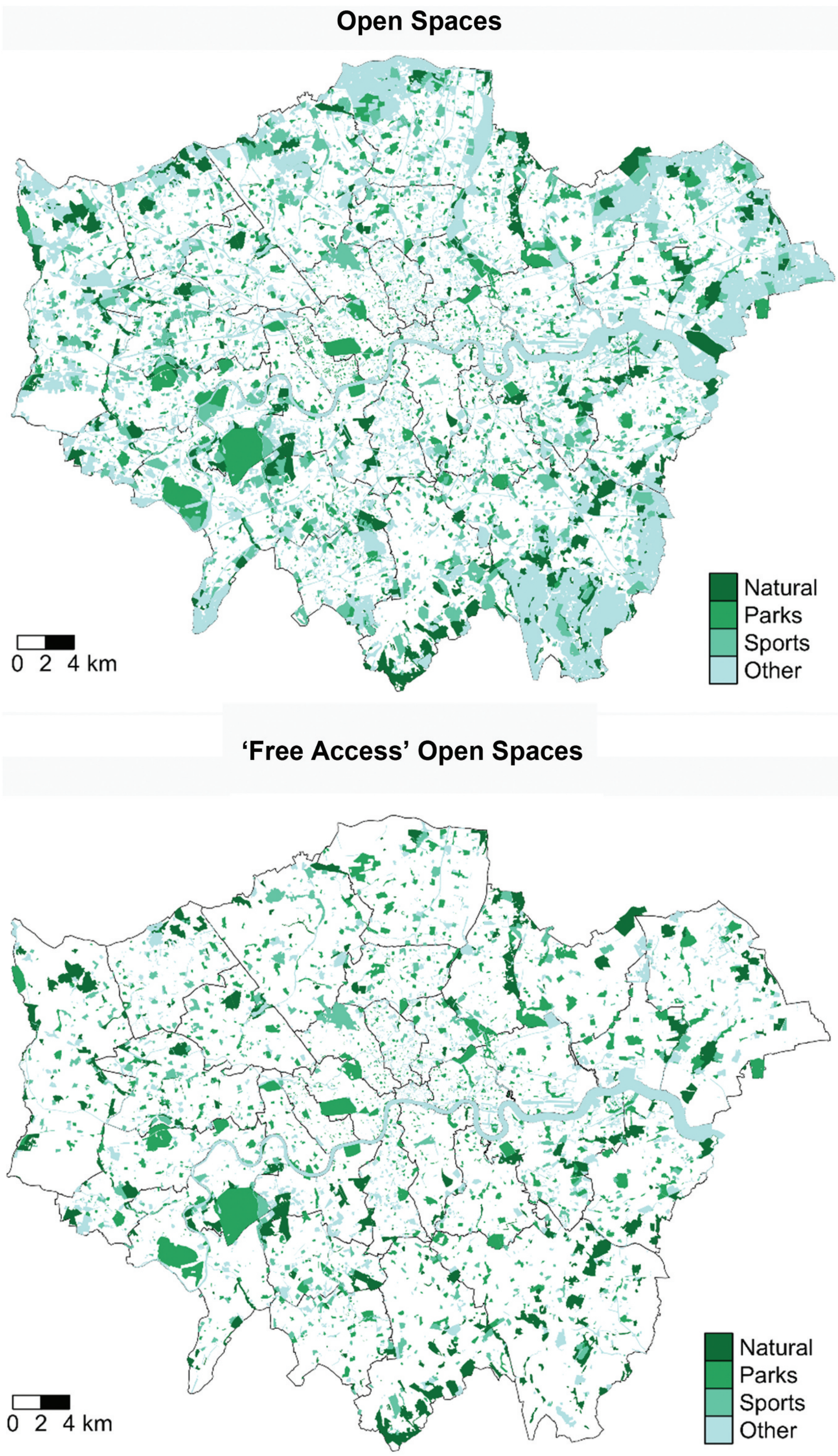
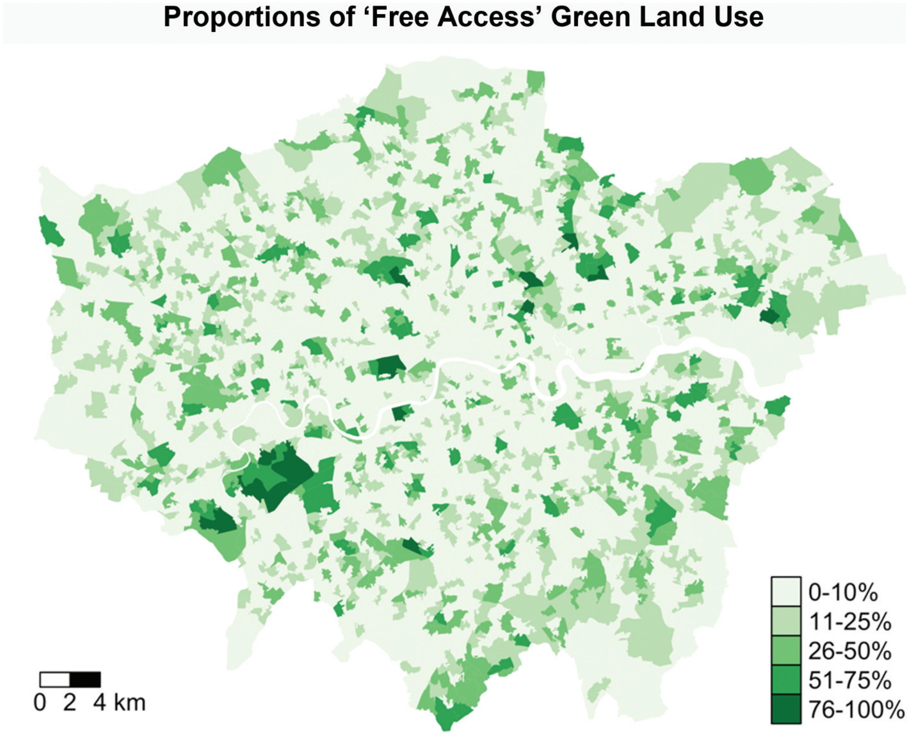


Figure 1. Maps of all GiGL open spaces (top) and GiGL open spaces with free or de facto access (bottom). Natural: natural and semi-natural urban greenspaces; Parks: parks and gardens; Sports: outdoor sports facilities; Other: all other open spaces. Maps display GiGL data [2020]. Maps contain National Statistics data © Crown copyright and database right [2015]. Maps contain Ordnance Survey data © Crown copyright and database right [2015].

importing a shapefile, we simplified its geometries using a function of the *rmapshaper* package (Teucher and Russell 2022). Simplifying geometries (i.e. polygons) makes the spatial object easier to work with. Simplifying the geometry of a polygon changes its area. As we were interested in the proportion of green land cover for each LSOA, we wanted to minimise changes in area as much as possible. By keeping 10% of the original vertices, we found a good balance between accuracy and object size. After simplification, we spatially intersected green cover polygons with LSOAs and calculated how much area of each LSOA was covered with green land cover (in %). Please see Figure 2 (bottom panel)



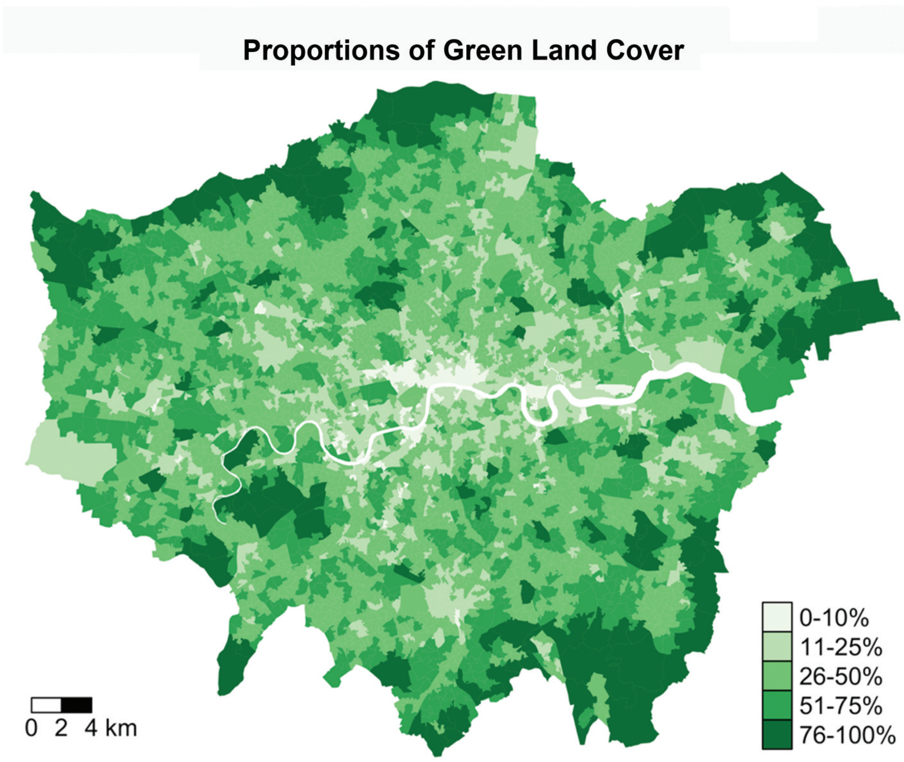


Figure 2. Proportions of 'free access' green land use (top) and green land cover (bottom) at LSOA level. The top map is derived from GiGL data [2020]. The bottom map contains Verisk Analytics GeoInformation Group UKMap data. Both maps contain National Statistics data © Crown copyright and database right [2015]. Both maps contain Ordnance Survey data © Crown copyright and database right [2015, 2019].

for a visualisation of the proportions of green land cover for London LSOAs.

Green land cover at postcode level. As for green land use, we calculated proportions of green land cover at postcode level, using circular buffers with a radius of 500 m. We spatially intersected these buffers with the green polygons of the London Green and Blue Cover data and calculated the area of green cover for each circular buffer (in %). As for green land use, we decided to also run analyses for 300 m and 1,000 m buffers (reported in the supplementary material).

Neighbourhood-level confounders

We included three neighbourhood-level confounders: air pollution, deprivation, and LSOA size. These variables are spatially correlated with greenspace, so it was important to adjust our statistical models for them. We will now describe each of the three variables in turn.

Air pollution was measured with nitrogen dioxide (NO_2) data provided by the GLA and Transport for London (TfL) for the years 2010, 2013, and 2016 ('Air Quality Data – London Datastore', 2022). NO_2 data (alongside other pollutants) are provided as annual mean concentrations (µg/m^3), modelled using the London Air Quality Toolkit (LAQT) model. The LAQT model uses a kernel modelling technique to describe the dispersion from emission sources (i.e. road transport; aviation; river; rail; industry; gas heating; domestic and commercial fuels; biomass burning; cooking emissions; and other sources). The contributions of these sources were summed and mapped on a 20 m by 20 m grid. Model results were validated by evaluating modelled data against fixed site measurements. Using the modelled data, we calculated the average annual mean NO_2 concentration for each LSOA (for the LSOA analysis) and each 500 m buffer (for the postcode analysis), calculating the mean of all 20 m by 20 m squares (for each LSOA and 500 m buffer). We linked the data of all three years (2010, 2013, and 2016) to UKHLS waves. Depending on when UKHLS data were collected, we linked air pollution data from 2010 (waves 1 and 2), 2013 (wave 3, 4, and 5), or 2016 (waves 6, 7, and 8).

Area deprivation was measured with the 2011 Carstairs Deprivation Index at LSOA level (Carstairs *et al.* 1989, Wheeler 2019). The Carstairs Index is the sum of the z-scores of four unweighted Census variables: proportions of low social class households; households with no car or van; overcrowded households; and male unemployment. The Carstairs Deprivation Index reflects the level of socio-economic deprivation at LSOA level, with higher scores indicating higher levels of deprivation.

LSOA size was measured as area in km^2. LSOAs on the outskirts of London tend to be larger and to have higher proportions of greenspace than LSOAs in the centre of London. We included LSOA area (in km^2) as a confounder in the LSOA analysis. Note that we did not include LSOA size in the postcode analysis.

Family- and child-level covariates

Family-level covariates were maternal mental health, maternal education, home ownership, and family structure. Maternal mental health was measured with the 12-item version of the General Health Questionnaire (GHQ). The GHQ scale score ranges from 0 to 36, with higher scores indicating higher psychological distress. Maternal education was measured with a binary variable (whether the mother has a University education). Home ownership (whether the family owns their home) and family structure (whether the child lives with two natural parents) were also measured with binary variables. Child-level covariates were sex (male/female) and ethnicity (White; Mixed; Indian; Pakistani and Bangladeshi; Black; and Other). As we were using data of multiple waves in one analysis (as will be described below), we also included a categorical variable for wave.

Statistical analysis

All analyses were run in Stata 16. To assess the effect of exposures on adolescent mental health and well-being, we fitted five linear regression models for each outcome (i.e. one for each of the five exposures). We adjusted each model for air pollution, deprivation, LSOA area (in the LSOA-level analysis), maternal mental health, maternal education, home ownership, family structure, sex, ethnicity, and wave. We stratified each model by age (i.e. ran each model for the six age groups separately). The reason for stratifying by age will be described below. All models accounted for the complex sampling design of the UKHLS (clustering and stratification) and for selective attrition (using study-specific weights). Noteworthy, because the land use (not land cover) variables were skewed with a large number of zeros and a few extreme values, we transformed these exposure variables, using a cube root transformation (which reduces the impact of extreme values). Also note that, for the postcode analysis, we only included adolescents whose 500 m buffers were fully within London (i.e. did not overlap with London's outer boundary).

A difficulty of this project was to make the most of the youth data. In Understanding Society, an individual is considered a youth when they are between 10 and 15 years old. This means that, depending on age at study entry, some individuals will never be considered 'youth'; some individuals will be considered 'youth' at one wave; and some individuals will be considered 'youth' at multiple waves. Due to this study design feature, youth data are not suitable for longitudinal analyses (and there are no longitudinal weights available). To avoid using data of only one wave, it is possible to pool the data of the eight cross-sectional datasets into one cross-sectional analysis. However, because some individuals were a youth at multiple waves, they contributed more than one observation to the analysis and, thus, observations were not independent. To address this, we pooled the data of the eight waves and ran separate models for each age group (i.e. 10, 11, 12, 13, 14, and 15 years). This ensured that each model included only one observation per individual and also allowed us to assess age-specific effects.

As some covariates had missing data, and under the assumption that missing data were missing at random (MAR), we imputed missing data using multiple imputation by chained equations (MICE; Raghunathan *et al.* 2001). For each analysis (i.e. outcome-age combination), we generated 25 imputed datasets and used Rubin's combination rules to pool the individual estimates into a single set of multiply imputed estimates (Rubin 1987). Around 74% of adolescents had complete data. The highest proportion of missingness was for maternal psychological distress (around 24%). Note that these numbers are averages; the exact amount of missingness differed depending on the age investigated in a given analysis. It should also be noted that the assumption that missing data were MAR is probably not true for maternal psychological distress. However, because a complete-case analysis would make our sample selective, and because sample sizes for individual analyses were already relatively small, we decided to use MICE to retain cases with missing data in our analysis.

To test for the robustness of results, we ran a series of sensitivity analyses. First, we ran analyses for buffer sizes of 300 m and 1,000 m. Using multiple buffer sizes is a common approach to assess whether results generalise to smaller (more proximal) and larger (more distal) exposure areas. Second, we assessed exposure by distance to the nearest greenspace, thereby focusing on accessibility (not availability). We assessed whether using distances (to the closest park or garden; natural or semi-natural urban greenspace; and outdoor sports facility) as exposures would lead to different results than using availability (i.e. proportions in 500 m around the home). Finally, we transformed (raw) green land use data into binary variables (comparing adolescents with 0% green land use in their neighbourhood with adolescents with at least some green land use), and variables with three categories (i.e. 'zero', 'some' [>0 AND < median], and 'more' [> median] green land use). This allowed us to test for potential non-linear effects.

Results

In this section, we report results based on the 500m buffer analysis. Results for the LSOA analysis and for sensitivity analyses are reported in the supplementary material (Tables S2.1 to S5.7). Note that analyses testing for non-linear effects did not add insight or clarity. For parsimony, these results are omitted from both main and supplementary analysis sections.

Descriptive statistics

Table 1 displays descriptive statistics for our analytic sample (i.e. observations across eight waves). Therefore, $n = 4,217$ does not refer to individuals but to observations. As adolescents in the analytic sample may not be representative of adolescents in the non-analytic sample, we ran a bias analysis to compare analytic and non-analytic samples (also shown in Table 1). Compared with adolescents in the non-analytic sample, those in the analytic sample had lower scores on the SDQ (i.e. better mental health), had higher scores on self-esteem and happiness (i.e. better mental well-being), and lived in more deprived areas. Also, there were differences in ethnicity: adolescents in the non-analytic sample were more likely to be 'White' (91%) than adolescents in the analytic sample (54%).

Table 2 displays correlations between outcomes and exposures. SDQ outcomes correlated positively with each other and negatively with self-esteem and happiness. Green land cover and green land use variables were inter-correlated positively. Interestingly, parks and gardens were negatively correlated with natural and semi-natural urban green spaces, and outdoor sports facilities, whereas the correlation between natural and semi-natural urban green spaces and outdoor sports facilities was positive. There were only three significant correlations between outcomes and exposures: both green land cover and green land use were negatively correlated with conduct problems, whereas outdoor sports facilities were positively correlated with happiness.

Model results

In this section, we describe the regression model results for the 500m buffer analysis. As mentioned above, results from LSOA analysis and sensitivity analyses (except for tests for non-linear effects) can be found in the supplementary material (Tables S2.1 to S5.7). All model results based on the 500m buffer analysis can be found in Tables 3–9. Each of the seven tables shows results for one of the seven outcomes considered in this study. Each table is separated by age group, and, for each age group, it displays estimates for each of the five exposures. Overall, there are only few statistically significant associations. However, unlike in the correlation analysis above, these associations are not all 'positive'; some suggest a 'negative' association (i.e. that more greenspace may be associated with poorer mental health). We summarise our results briefly below. For an overview, please see Tables 3–9. All regression results reported in text and in tables are taken from fully adjusted models (including neighbourhood-, family-, and child-level variables). Note, however, that only coefficients for the main exposures of interest (i.e. five greenspace variables) are reported.

For conduct problems, we found negative associations with green land use in 13-year-olds ($b = -0.286$, $p = .032$), and with outdoor sports facilities in 14-year-

Table 2. Correlations between outcomes and exposures (*n* = 4,217).

	SDQ CP	SDQ ES	SDQ HA	SDQ PP	SDQ TD	Self-esteem	Happiness	Green LC[1]	Green LU[1]	P/G[1]	N/SN UG[1]
SDQ ES	0.260 ***										
SDQ HA	0.513 ***	0.303 ***									
SDQ PP	0.200 ***	0.355 ***	0.179 ***								
SDQ TD	0.707 ***	0.717 ***	0.754 ***	0.574 ***							
Self-esteem	-	-	-	-	-						
Happiness	−0.374 ***	−0.397 ***	−0.367 ***	−0.293 ***	−0.521 ***	0.538 ***					
Green LC[1]	−0.058 **	−0.010	0.029	−0.025	−0.020	−0.033	0.013				
Green LU[1]	−0.049 *	−0.017	−0.012	0.005	−0.028	−0.016	0.023	0.541 ***			
P/G[1]	−0.022	−0.006	−0.014	−0.010	−0.019	−0.027	−0.004	0.266 ***	0.669 ***		
N/SN UG[1]	−0.036	0.002	0.009	0.032	0.002	0.000	0.002	0.343 ***	0.491 ***	−0.137 ***	
OSF[1]	−0.027	−0.030	−0.014	−0.010	−0.030	0.011	0.058 ***	0.306 ***	0.440 ***	−0.074 ***	0.074 ***

Note: CP = conduct problems; ES = emotional symptoms; HA = hyperactivity/inattention; PP = peer problems; TD = total difficulties; LC = land cover; LU = land use; P/G = parks/gardens; N/SN UG = natural/semi-natural urban greenspaces; OSF = outdoor sports facilities. Data are taken from waves 1 to 8. The sample size refers to observations (not individuals). Some individuals have multiple observations across waves, and these multiple observations are included in the correlations above. The sample size used to establish a given correlation depends on the variables involved in that correlation; the smallest sample size is *n* = 1,839. There are no correlations between self-esteem and SDQ scales because these outcomes were measured at different waves. Values of 0.000 represent values > 0 AND < 0.001. [1] Exposures are measured in 500 m buffers around postcodes. * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 3. Regression results for conduct problems (500m buffer analysis).

	b	SE	95% CI	p
15 years (*n* = 365)				
Green land cover	−0.015	0.013	[−0.041, 0.010]	0.227
Green land use	0.024	0.127	[−0.229, 0.278]	0.850
Parks/gardens	0.151	0.105	[−0.058, 0.360]	0.154
Natural/semi-natural spaces	−0.092	0.111	[−0.314, 0.129]	0.410
Outdoor sports facilities	0.052	0.128	[−0.203, 0.307]	0.685
14 years (*n* = 349)				
Green land cover	−0.004	0.014	[−0.032, 0.024]	0.774
Green land use	−0.195	0.172	[−0.537, 0.148]	0.261
Parks/gardens	−0.071	0.129	[−0.329, 0.186]	0.583
Natural/semi-natural spaces	−0.084	0.184	[−0.450, 0.281]	0.647
Outdoor sports facilities	−0.291	0.140	[−0.570, −0.012]	0.041
13 years (*n* = 378)				
Green land cover	−0.019	0.018	[−0.056, 0.017]	0.295
Green land use	−0.286	0.131	[−0.547, −0.026]	0.032
Parks/gardens	−0.105	0.108	[−0.321, 0.110]	0.334
Natural/semi-natural spaces	−0.130	0.132	[−0.393, 0.133]	0.327
Outdoor sports facilities	−0.250	0.166	[−0.581, 0.081]	0.137
12 years (*n* = 392)				
Green land cover	0.004	0.009	[−0.014, 0.023]	0.659
Green land use	0.153	0.185	[−0.215, 0.522]	0.410
Parks/gardens	0.111	0.101	[−0.089, 0.312]	0.272
Natural/semi-natural spaces	0.044	0.176	[−0.306, 0.395]	0.802
Outdoor sports facilities	−0.055	0.135	[−0.324, 0.214]	0.686
11 years (*n* = 368)				
Green land cover	−0.010	0.012	[−0.033, 0.013]	0.407
Green land use	−0.217	0.132	[−0.480, 0.047]	0.106
Parks/gardens	0.027	0.136	[−0.243, 0.297]	0.841
Natural/semi-natural spaces	−0.250	0.198	[−0.643, 0.144]	0.211
Outdoor sports facilities	−0.261	0.192	[−0.643, 0.121]	0.178
10 years (*n* = 375)				
Green land cover	0.001	0.013	[−0.025, 0.027]	0.959
Green land use	−0.027	0.247	[−0.519, 0.464]	0.912
Parks/gardens	−0.120	0.163	[−0.445, 0.205]	0.464
Natural/semi-natural spaces	0.075	0.137	[−0.199, 0.348]	0.583
Outdoor sports facilities	0.134	0.181	[−0.227, 0.496]	0.461

Note. b = coefficient; SE = standard error; CI = confidence interval. Estimates are taken from separate models (i.e. one model for each age-exposure combination). Estimates are pooled estimates of 25 imputed datasets. The green land cover variable is based on raw data [%], whereas the green land use variables are based on cube root transformed data [$\sqrt[3]{\%}$]. The size of the green land cover coefficient should therefore not be compared to the size of a green land use coefficient.

olds (b = −0.291, p = .041). ***For emotional symptoms,*** we found negative associations with green land use (b = −0.312, p = .004) and parks and gardens (b = −0.290, p = .035) in 13-year-olds, and with outdoor sports facilities in 15-year-olds (b = −0.359, p = .032). Noteworthy, we also found positive associations with green land use (b = 0.264, p = .014) and parks and gardens (b = 0.290, p = .010) in 15-year-olds (suggesting that more greenspace could also be linked to more problems). ***For hyperactivity and inattention,*** we found negative associations with green land use (b = −0.322, p = .014) and parks and gardens (b = −0.299, p = .015) in 13-year-olds, and with outdoor sports facilities in 14-year-olds (b = −0.589, p = .005). However, we also found positive associations with green land cover in 10- and 11-year-olds (b = 0.037,

Table 4. Regression results for emotional symptoms (500m buffer analysis).

	b	SE	95% CI	p
15 years (*n* = 365)				
Green land cover	0.013	0.010	[−0.006, 0.032]	0.172
Green land use	0.264	0.105	[0.056, 0.473]	0.014
Parks/gardens	0.290	0.111	[0.070, 0.510]	0.010
Natural/semi-natural spaces	0.318	0.189	[−0.059, 0.695]	0.097
Outdoor sports facilities	−0.359	0.164	[−0.686, −0.032]	0.032
14 years (*n* = 349)				
Green land cover	0.009	0.015	[−0.022, 0.040]	0.562
Green land use	−0.329	0.256	[−0.838, 0.181]	0.203
Parks/gardens	0.048	0.185	[−0.319, 0.416]	0.794
Natural/semi-natural spaces	−0.334	0.304	[−0.938, 0.270]	0.274
Outdoor sports facilities	−0.228	0.233	[−0.692, 0.236]	0.332
13 years (*n* = 378)				
Green land cover	−0.025	0.019	[−0.063, 0.013]	0.198
Green land use	−0.312	0.106	[−0.522, −0.102]	0.004
Parks/gardens	−0.290	0.135	[−0.558, −0.021]	0.035
Natural/semi-natural spaces	−0.105	0.137	[−0.379, 0.169]	0.449
Outdoor sports facilities	0.084	0.199	[−0.313, 0.480]	0.676
12 years (*n* = 392)				
Green land cover	−0.006	0.014	[−0.034, 0.021]	0.643
Green land use	0.030	0.217	[−0.401, 0.461]	0.889
Parks/gardens	0.081	0.145	[−0.206, 0.368]	0.576
Natural/semi-natural spaces	0.018	0.218	[−0.416, 0.451]	0.936
Outdoor sports facilities	0.007	0.137	[−0.265, 0.280]	0.957
11 years (*n* = 368)				
Green land cover	0.002	0.011	[−0.021, 0.024]	0.876
Green land use	0.011	0.202	[−0.391, 0.412]	0.957
Parks/gardens	−0.094	0.139	[−0.371, 0.183]	0.503
Natural/semi-natural spaces	0.047	0.213	[−0.377, 0.471]	0.826
Outdoor sports facilities	0.028	0.267	[−0.502, 0.558]	0.916
10 years (*n* = 375)				
Green land cover	0.032	0.019	[−0.006, 0.070]	0.096
Green land use	0.137	0.305	[−0.470, 0.744]	0.654
Parks/gardens	0.068	0.190	[−0.312, 0.447]	0.724
Natural/semi-natural spaces	0.238	0.424	[−0.607, 1.083]	0.576
Outdoor sports facilities	−0.252	0.228	[−0.706, 0.202]	0.273

Note. b = coefficient; SE = standard error; CI = confidence interval. Estimates are taken from separate models (i.e. one model for each age-exposure combination). Estimates are pooled estimates of 25 imputed datasets. The green land cover variable is based on raw data [%], whereas the green land use variables are based on cube root transformed data [$\sqrt[3]{\%}$]. The size of the green land cover coefficient should therefore not be compared to the size of a green land use coefficient.

p = .003; b = 0.028, p = .009), and with parks and gardens in 12-year-olds (b = 0.235, p = .039). ***For peer relationship problems,*** we found a negative association with parks and gardens in 13-year-olds (b = −0.161, p = .031), but also positive associations with green land use (b = 0.244, p = .045) and natural and semi-natural urban greenspaces (b = 0.439, p = .008) in 10-year-olds. ***For total difficulties,*** we found negative associations with green land use (b = −0.932, p = .008) and parks and gardens (b = −0.856, p = .007) in 13-year-olds, and with outdoor sports facilities in 14-year-olds (b = −1.186, p = .007). However, we also found a positive association with green land cover in 10-year-olds (b = 0.091, p = .007). ***For self-esteem,*** we found a positive association with natural and semi-natural urban greenspaces in 15-year-olds (b = 0.088, p = .009). ***Finally, for happiness,*** we found positive associations with green land use (b = 0.095, p = .021) and outdoor sports facilities (b = 0.127, p = .011) in 14-year-olds, and with green land use in 15-year-olds (b = 0.098, p = .026).

The mixed nature of the results makes it difficult to clearly summarise the main findings. Indeed, one of the main conclusions that can be drawn is that, across outcomes and age groups, there is no consistent pattern of findings. However, what can be observed is that the direction of findings seems to be different in younger and older adolescents. Except for the positive associations of green land use, and parks and gardens with emotional symptoms in 15-year-old adolescents, similar ‘negative’ links (i.e. suggesting a negative link between greenspace and mental health) were only found in 10- to 12-year-olds, especially for social and externalising-type problems. In contrast, in 13- to 15-year-olds, we typically found ‘positive’ links between greenspace and mental health and well-being (except for the two ‘negative’ associations mentioned above). It seems that, on average, older urban adolescents benefit from green spaces in their neighbourhoods, while younger urban adolescents living in greener neighbourhoods show more peer and externalising problems. To test whether this pattern was not an artefact driven by a few extreme observations, we ran an additional sensitivity analysis (not mentioned above; see Table S1 in the supplementary material). We grouped 10- to 12-year-olds and 13- to 15-year-olds and, for each wave, ran models for each of the two groups. For waves 1 and 3, we found ‘positive’

Table 5. Regression results for hyperactivity and inattention (500m buffer analysis).

	b	SE	95% CI	p
15 years (*n* = 365)				
Green land cover	0.005	0.011	[−0.017, 0.028]	0.630
Green land use	0.162	0.137	[−0.111, 0.435]	0.240
Parks/gardens	0.062	0.143	[−0.223, 0.347]	0.665
Natural/semi-natural spaces	0.033	0.127	[−0.220, 0.287]	0.793
Outdoor sports facilities	−0.005	0.161	[−0.325, 0.315]	0.975
14 years (*n* = 349)				
Green land cover	0.004	0.017	[−0.029, 0.037]	0.827
Green land use	−0.024	0.260	[−0.541, 0.492]	0.925
Parks/gardens	0.233	0.204	[−0.172, 0.638]	0.256
Natural/semi-natural spaces	−0.208	0.470	[−1.141, 0.726]	0.660
Outdoor sports facilities	−0.589	0.204	[−0.994, −0.183]	0.005
13 years (*n* = 378)				
Green land cover	−0.021	0.019	[−0.058, 0.017]	0.272
Green land use	−0.322	0.128	[−0.577, −0.067]	0.014
Parks/gardens	−0.299	0.120	[−0.538, −0.061]	0.015
Natural/semi-natural spaces	−0.137	0.169	[−0.473, 0.198]	0.418
Outdoor sports facilities	0.128	0.172	[−0.215, 0.471]	0.459
12 years (*n* = 392)				
Green land cover	0.007	0.012	[−0.018, 0.031]	0.584
Green land use	0.183	0.182	[−0.178, 0.544]	0.316
Parks/gardens	0.235	0.112	[0.012, 0.458]	0.039
Natural/semi-natural spaces	−0.118	0.158	[−0.433, 0.196]	0.456
Outdoor sports facilities	0.135	0.137	[−0.137, 0.407]	0.327
11 years (*n* = 368)				
Green land cover	0.028	0.011	[0.007, 0.049]	0.009
Green land use	0.077	0.132	[−0.186, 0.340]	0.561
Parks/gardens	−0.068	0.120	[−0.307, 0.171]	0.573
Natural/semi-natural spaces	0.136	0.190	[−0.242, 0.514]	0.476
Outdoor sports facilities	0.057	0.213	[−0.367, 0.481]	0.790
10 years (*n* = 375)				
Green land cover	0.037	0.012	[0.013, 0.061]	0.003
Green land use	0.126	0.202	[−0.276, 0.528]	0.534
Parks/gardens	0.117	0.154	[−0.191, 0.424]	0.451
Natural/semi-natural spaces	0.057	0.268	[−0.477, 0.592]	0.832
Outdoor sports facilities	0.084	0.150	[−0.214, 0.382]	0.575

Note. b = coefficient; SE = standard error; CI = confidence interval. Estimates are taken from separate models (i.e. one model for each age-exposure combination). Estimates are pooled estimates of 25 imputed datasets. The green land cover variable is based on raw data [%], whereas the green land use variables are based on cube root transformed data [$\sqrt[3]{\%}$]. The size of the green land cover coefficient should therefore not be compared to the size of a green land use coefficient.

associations in 13- to 15-year-olds. In wave 7, we found a 'positive' association in 13- to 15-year-olds, and several 'negative' associations in 10- to 12-year-olds. For wave 5, however, we found a 'negative' association in 13- to 15-year-olds and a 'positive' association in 10- to 12-year-olds. Although the sensitivity analysis seems to support the finding that greenspace may be more beneficial for older adolescents (except for the wave 5 results), it should be noted that specific associations were not the same across analyses.

In addition to the 500m buffer analysis, we ran models using exposures at different scales: LSOA, 300 m buffer, and 1,000 m buffer. Further, we ran models using distances to the closest green land use; park or garden; natural or semi-natural urban greenspace; and outdoor sports facility. We will not describe these results in detail, but all tables can be found in the supplementary material (Tables S2.1 to S5.7). None of the sensitivity analyses provided a clearer pattern of associations than that of the main analysis. Findings remained mixed and inconsistent across analyses. The results of LSOA and 300-m-buffer analyses were similar to the results of the main analysis. The results of 1,000-m-buffer and distance analyses were slightly different and showed generally fewer significant associations.

Discussion

In this study, we investigated the associations of different types of greenspace with the mental health and well-being of adolescents (10- to 15-year-olds) living in London, UK. We distinguished between green land cover; green land use; parks and gardens; natural and semi-natural urban green spaces; and outdoor sports facilities. Across age groups (10 to 15 years) and outcome domains (mental health and well-being), we did not find consistent results that would allow for clear conclusions about what types of greenspace may be most beneficial for adolescents. However, we did make a few observations that we will now discuss, as they provide interesting insights and raise questions for future research. We will also discuss the limitations of our study.

Table 6. Regression results for peer relationship problems (500m buffer analysis).

	b	SE	95% CI	p
15 years (*n* = 365)				
Green land cover	−0.008	0.006	[−0.019, 0.004]	0.209
Green land use	0.073	0.089	[−0.103, 0.249]	0.410
Parks/gardens	0.083	0.090	[−0.096, 0.262]	0.358
Natural/semi-natural spaces	0.152	0.079	[−0.006, 0.310]	0.060
Outdoor sports facilities	−0.115	0.110	[−0.335, 0.105]	0.300
14 years (*n* = 349)				
Green land cover	0.001	0.010	[−0.018, 0.020]	0.912
Green land use	0.000	0.178	[−0.355, 0.355]	0.999
Parks/gardens	−0.152	0.106	[−0.362, 0.059]	0.156
Natural/semi-natural spaces	0.153	0.131	[−0.108, 0.414]	0.246
Outdoor sports facilities	−0.078	0.106	[−0.288, 0.132]	0.461
13 years (*n* = 378)				
Green land cover	−0.007	0.008	[−0.022, 0.009]	0.407
Green land use	−0.011	0.106	[−0.222, 0.199]	0.917
Parks/gardens	−0.161	0.073	[−0.306, −0.016]	0.031
Natural/semi-natural spaces	0.099	0.127	[−0.155, 0.352]	0.440
Outdoor sports facilities	0.188	0.104	[−0.020, 0.395]	0.075
12 years (*n* = 392)				
Green land cover	0.009	0.007	[−0.006, 0.024]	0.217
Green land use	0.054	0.090	[−0.125, 0.234]	0.549
Parks/gardens	−0.019	0.077	[−0.173, 0.134]	0.804
Natural/semi-natural spaces	0.160	0.109	[−0.057, 0.377]	0.146
Outdoor sports facilities	0.122	0.090	[−0.058, 0.302]	0.181
11 years (*n* = 368)				
Green land cover	−0.004	0.010	[−0.023, 0.015]	0.677
Green land use	0.161	0.154	[−0.144, 0.467]	0.297
Parks/gardens	0.040	0.113	[−0.186, 0.266]	0.725
Natural/semi-natural spaces	0.010	0.188	[−0.365, 0.384]	0.959
Outdoor sports facilities	0.161	0.254	[−0.344, 0.666]	0.527
10 years (*n* = 375)				
Green land cover	0.021	0.011	[−0.000, 0.043]	0.055
Green land use	0.244	0.120	[0.006, 0.483]	0.045
Parks/gardens	0.016	0.095	[−0.174, 0.206]	0.865
Natural/semi-natural spaces	0.439	0.162	[0.116, 0.763]	0.008
Outdoor sports facilities	0.125	0.162	[−0.198, 0.449]	0.443

Note. b = coefficient; SE = standard error; CI = confidence interval. Estimates are taken from separate models (i.e. one model for each age-exposure combination). Estimates are pooled estimates of 25 imputed datasets. Values of 0.000 represent values > 0 AND < 0.001; values of −0.000 represent values < 0 AND > −0.001. The green land cover variable is based on raw data [%], whereas the green land use variables are based on cube root transformed data [$\sqrt[3]{\%}$]. The size of the green land cover coefficient should therefore not be compared to the size of a green land use coefficient.

Main observations

The first observation is that there were only few associations of green land cover with adolescent mental health, and all of these associations were 'negative': higher levels of green land cover were associated with more mental health problems (i.e. more hyperactivity and inattention in 10- and 11-year-olds). This finding is not in line with the generally positive associations between the NDVI and mental health reported in previous studies. It also is not in line with studies suggesting that merely viewing nature (Taylor *et al.* 2007) or actively noticing nature (Passmore and Holder 2017), as opposed to intentionally visiting nature, may have benefits for well-being. Indeed, greener areas may offer opportunities for incidental exposure, and viewing and noticing nature, and higher levels of greenness may also have more indirect effects on health and well-being, for example, by reducing levels of environmental stressors (Markevych *et al.* 2017). Therefore, it would certainly be plausible that the greenness of an area is associated with adolescents' mental health and well-being, but we did not find this association in our study. This may have to do, in part, with the specifics of our green land cover variable. The green land cover variable in this study was based on a combination of NDVI and land use data, capturing the greenness of an area. The advantage of this measure is that it captures any type of greenery (such as trees) rather than only designated open spaces (such as parks). However, this also means that it captures greenery that may not be beneficial for adolescents (such as green roofs). Furthermore, any type of green land cover was categorised as 'green', so the measure does not distinguish between dense and sparse vegetation (as the raw NDVI does). Therefore, although the measure is a proxy for the greenness of an area, it does not tell us much about the quality of this greenness. Unlike the green land use measures used in this study, the green land cover measure also does not capture the usability of the greenery. All this may explain why we did not find a positive association of green land cover with mental health.

It is unclear why more green land cover may be linked to more hyperactivity and inattention problems in 10- and 11-year-olds. One explanation could be that young adolescents may be more restricted in their independent mobility and may therefore spend more time in their own neighbourhoods. If their

Table 7. Regression results for total difficulties (500m buffer analysis).

	b	SE	95% CI	p
15 years (n = 365)				
Green land cover	−0.004	0.029	[−0.062, 0.054]	0.883
Green land use	0.524	0.345	[−0.163, 1.211]	0.133
Parks/gardens	0.587	0.329	[−0.066, 1.241]	0.078
Natural/semi-natural spaces	0.411	0.354	[−0.293, 1.115]	0.248
Outdoor sports facilities	−0.427	0.418	[−1.259, 0.406]	0.311
14 years (n = 349)				
Green land cover	0.010	0.031	[−0.052, 0.071]	0.756
Green land use	−0.548	0.530	[−1.601, 0.506]	0.304
Parks/gardens	0.059	0.353	[−0.644, 0.761]	0.869
Natural/semi-natural spaces	−0.473	0.876	[−2.215, 1.269]	0.591
Outdoor sports facilities	−1.186	0.430	[−2.041, −0.330]	0.007
13 years (n = 378)				
Green land cover	−0.072	0.056	[−0.184, 0.041]	0.208
Green land use	−0.932	0.341	[−1.611, −0.252]	0.008
Parks/gardens	−0.856	0.309	[−1.472, −0.239]	0.007
Natural/semi-natural spaces	−0.273	0.386	[−1.042, 0.495]	0.481
Outdoor sports facilities	0.149	0.417	[−0.682, 0.980]	0.721
12 years (n = 392)				
Green land cover	0.014	0.029	[−0.043, 0.071]	0.632
Green land use	0.421	0.477	[−0.527, 1.369]	0.380
Parks/gardens	0.408	0.296	[−0.179, 0.996]	0.171
Natural/semi-natural spaces	0.104	0.514	[−0.918, 1.125]	0.841
Outdoor sports facilities	0.210	0.335	[−0.455, 0.874]	0.533
11 years (n = 368)				
Green land cover	0.016	0.028	[−0.039, 0.072]	0.557
Green land use	0.033	0.413	[−0.790, 0.855]	0.937
Parks/gardens	−0.094	0.306	[−0.703, 0.514]	0.759
Natural/semi-natural spaces	−0.057	0.600	[−1.251, 1.137]	0.925
Outdoor sports facilities	−0.014	0.789	[−1.584, 1.555]	0.985
10 years (n = 375)				
Green land cover	0.091	0.033	[0.025, 0.156]	0.007
Green land use	0.480	0.666	[−0.846, 1.806]	0.473
Parks/gardens	0.081	0.363	[−0.644, 0.805]	0.825
Natural/semi-natural spaces	0.810	0.782	[−0.747, 2.368]	0.303
Outdoor sports facilities	0.092	0.405	[−0.715, 0.899]	0.821

Note. b = coefficient; SE = standard error; CI = confidence interval. Estimates are taken from separate models (i.e. one model for each age-exposure combination). Estimates are pooled estimates of 25 imputed datasets. The green land cover variable is based on raw data [%], whereas the green land use variables are based on cube root transformed data [$\sqrt[3]{\%}$]. The size of the green land cover coefficient should therefore not be compared to the size of a green land use coefficient.

neighbourhoods have high levels of green land cover which may not be free for them to use (e.g. agriculture, private woodlands, or golf courses), they may not have the opportunity to step outside to play and be active. This, in turn, may lead to higher levels of hyperactivity and inattention. Although we found similar findings in the 300 m and 1,000 m sensitivity analyses, our explanation for these findings is, of course, speculative. In fact, the finding that higher levels of parks and gardens were associated with higher levels of hyperactivity and inattention in 12-year-olds is not in line with our explanation (because parks and gardens are 'usable' spaces). Therefore, the role of green land cover and green land use in early adolescence needs further investigation. For example, it would be possible that high levels of green land cover and greenspace are proxies for other environmental characteristics that may have negative effects on adolescent mental health and well-being, such as a lack of facilities valued by adolescents (e.g. shops). In other words, if living in a greener neighbourhood means missing out on other opportunities, this may impact the mental health and well-being of adolescents negatively.

The second observation is that there was a pattern of positive associations (i.e. more greenspace was associated with better mental health) in older adolescents, and negative associations (i.e. more greenspace was associated with poorer mental health) in younger adolescents. Older adolescents (13 to 15 years) seemed to 'benefit' from green land use (especially parks and gardens and outdoor sports facilities) across outcomes, whereas younger adolescents (10 to 12 years) seemed to 'dis-benefit' from green land use and, as mentioned above, green land cover (however, only in terms of hyperactivity and inattention and peer relationship problems). The only exception of this were positive associations of green land use and parks and gardens with emotional symptoms in 15-year-olds (suggesting that more parks and gardens were associated with more emotional symptoms). The differences between age groups were generally supported in sensitivity analyses, especially in the 300-m-buffer analysis and in the wave-specific sensitivity analysis. Noteworthy, the positive associations of green land use with mental health problems in 15-year-olds were also supported and, indeed, extended in

Table 8. Regression results for self-esteem (500m buffer analysis).

	b	SE	95% CI	p
15 years (*n* = 293)				
Green land cover	−0.000	0.004	[−0.008, 0.007]	0.912
Green land use	0.067	0.071	[−0.074, 0.208]	0.347
Parks/gardens	0.013	0.052	[−0.091, 0.117]	0.808
Natural/semi-natural spaces	0.088	0.033	[0.023, 0.154]	0.009
Outdoor sports facilities	0.024	0.043	[−0.061, 0.109]	0.579
14 years (*n* = 306)				
Green land cover	0.001	0.002	[−0.002, 0.005]	0.417
Green land use	0.040	0.024	[−0.008, 0.088]	0.100
Parks/gardens	0.019	0.016	[−0.013, 0.051]	0.243
Natural/semi-natural spaces	0.013	0.033	[−0.054, 0.080]	0.698
Outdoor sports facilities	0.022	0.027	[−0.033, 0.076]	0.428
13 years (*n* = 338)				
Green land cover	−0.003	0.002	[−0.008, 0.001]	0.177
Green land use	−0.054	0.050	[−0.154, 0.046]	0.284
Parks/gardens	−0.050	0.031	[−0.112, 0.011]	0.109
Natural/semi-natural spaces	−0.010	0.074	[−0.158, 0.139]	0.897
Outdoor sports facilities	−0.024	0.045	[−0.113, 0.065]	0.596
12 years (*n* = 298)				
Green land cover	−0.003	0.003	[−0.010, 0.003]	0.342
Green land use	−0.045	0.036	[−0.116, 0.026]	0.213
Parks/gardens	0.005	0.030	[−0.055, 0.065]	0.863
Natural/semi-natural spaces	−0.052	0.054	[−0.160, 0.057]	0.342
Outdoor sports facilities	−0.005	0.062	[−0.129, 0.120]	0.940
11 years (*n* = 336)				
Green land cover	−0.001	0.002	[−0.005, 0.003]	0.661
Green land use	0.007	0.028	[−0.048, 0.063]	0.788
Parks/gardens	−0.024	0.020	[−0.064, 0.016]	0.230
Natural/semi-natural spaces	0.025	0.022	[−0.020, 0.070]	0.279
Outdoor sports facilities	0.016	0.039	[−0.062, 0.093]	0.689
10 years (*n* = 264)				
Green land cover	−0.001	0.003	[−0.008, 0.005]	0.694
Green land use	0.016	0.051	[−0.086, 0.119]	0.749
Parks/gardens	0.013	0.041	[−0.069, 0.095]	0.751
Natural/semi-natural spaces	−0.033	0.076	[−0.185, 0.119]	0.662
Outdoor sports facilities	0.082	0.046	[−0.010, 0.173]	0.079

Note. b = coefficient; SE = standard error; CI = confidence interval. Estimates are taken from separate models (i.e. one model for each age-exposure combination). Estimates are pooled estimates of 25 imputed datasets. Values of −0.000 represent values < 0 AND > −0.001. The green land cover variable is based on raw data [%], whereas the green land use variables are based on cube root transformed data [$\sqrt[3]{\%}$]. The size of the green land cover coefficient should therefore not be compared to the size of a green land use coefficient.

sensitivity analyses. In both the LSOA analysis and the 300-m analysis (but not the 1,000-m analysis), more parks and gardens were associated with more conduct problems in 15-year-olds. Similarly, a further distance to the nearest park or garden was linked to lower levels of conduct problems. This suggests that living in close proximity to a park or garden in London is linked to more conduct problems in older adolescents. In 13- and 14-year-olds, however, availability and proximity of parks and gardens and outdoor sports facilities seemed to be linked to fewer mental health problems.

A third observation, which is related to the previous observation, is that green land use was positively associated with mental well-being (i.e. self-esteem and happiness) only in older adolescents (14- and 15-year-olds), and this was generally supported in sensitivity analyses (especially in the LSOA analysis and in the 300-m-buffer analysis). This suggests that older adolescents living in greener areas in London not only show fewer mental health problems but also more well-being. This is important to note because mental health and well-being are, while related, not the same constructs.

Although individual associations were not robust, the pattern of associations described above suggests that the role of green land use is not the same across age groups: older adolescents seem to 'benefit' more from green land use in their neighbourhoods than younger adolescents, both in terms of mental health and well-being. This is in line with studies suggesting that effects differ by age. For example, Madzia *et al.* (2019) found that neighbourhood greenness affected 7-year-old children differently than 12-year-olds, Feng and Astell-Burt (2017) found that neighbourhood greenspace quantity and quality had different effects across childhood (4-5 to 12-13 years), and Bezold *et al.* (2018) found stronger effects in middle school students than in high school students.

From a theoretical perspective, it is plausible that greenspace affects children and adolescents at different ages differently. This is because, across childhood and adolescence, individuals go through several developmental stages, experiencing biological and social changes. In adolescence, individuals become more independent, enter puberty, shift their focus from

Table 9. Regression results for happiness (500m buffer analysis).

	b	SE	95% CI	p
15 years (*n* = 663)				
Green land cover	0.002	0.004	[−0.005, 0.009]	0.561
Green land use	0.098	0.044	[0.012, 0.185]	0.026
Parks/gardens	0.057	0.049	[−0.039, 0.153]	0.240
Natural/semi-natural spaces	0.075	0.088	[−0.099, 0.248]	0.397
Outdoor sports facilities	0.031	0.050	[−0.068, 0.130]	0.540
14 years (*n* = 663)				
Green land cover	0.004	0.004	[−0.004, 0.013]	0.318
Green land use	0.095	0.041	[0.015, 0.176]	0.021
Parks/gardens	−0.012	0.042	[−0.095, 0.071]	0.771
Natural/semi-natural spaces	0.116	0.071	[−0.023, 0.255]	0.102
Outdoor sports facilities	0.127	0.049	[0.029, 0.225]	0.011
13 years (*n* = 725)				
Green land cover	0.005	0.003	[−0.002, 0.011]	0.173
Green land use	0.013	0.054	[−0.095, 0.120]	0.812
Parks/gardens	−0.006	0.040	[−0.084, 0.073]	0.888
Natural/semi-natural spaces	0.003	0.084	[−0.162, 0.168]	0.972
Outdoor sports facilities	0.048	0.068	[−0.086, 0.181]	0.482
12 years (*n* = 699)				
Green land cover	−0.009	0.005	[−0.018, 0.000]	0.060
Green land use	−0.008	0.064	[−0.133, 0.117]	0.902
Parks/gardens	−0.028	0.052	[−0.130, 0.074]	0.583
Natural/semi-natural spaces	−0.022	0.071	[−0.162, 0.117]	0.754
Outdoor sports facilities	0.038	0.069	[−0.099, 0.175]	0.582
11 years (*n* = 710)				
Green land cover	−0.005	0.003	[−0.012, 0.001]	0.091
Green land use	−0.076	0.042	[−0.159, 0.006]	0.069
Parks/gardens	−0.059	0.035	[−0.128, 0.010]	0.091
Natural/semi-natural spaces	−0.033	0.069	[−0.170, 0.104]	0.634
Outdoor sports facilities	0.060	0.052	[−0.042, 0.162]	0.246
10 years (*n* = 644)				
Green land cover	−0.006	0.003	[−0.013, 0.001]	0.101
Green land use	−0.058	0.048	[−0.153, 0.036]	0.226
Parks/gardens	−0.040	0.039	[−0.117, 0.038]	0.313
Natural/semi-natural spaces	−0.096	0.069	[−0.232, 0.040]	0.164
Outdoor sports facilities	0.048	0.052	[−0.054, 0.151]	0.355

Note. b = coefficient; SE = standard error; CI = confidence interval. Estimates are taken from separate models (i.e. one model for each age-exposure combination). Estimates are pooled estimates of 25 imputed datasets. Values of 0.000 represent values > 0 AND < 0.001. The green land cover variable is based on raw data [%], whereas the green land use variables are based on cube root transformed data [$\sqrt[3]{\%}$]. The size of the green land cover coefficient should therefore not be compared to the size of a green land use coefficient.

parents to peers, and start taking more risks (Christie and Viner 2005). Therefore, an individual will show different needs and interests at the age of 10 or 11 years than at the age of 13 or 14 years. Young adolescents may still be dependent on their parents to take them outside, whereas older adolescents will be allowed to move around their neighbourhoods in a wider radius and unsupervised. With age, adolescents spend more time away from home and with their peers. Public spaces, such as green spaces, may therefore become more important in older adolescence. Our findings suggest that parks & gardens and outdoor sports facilities may be especially 'beneficial' for the mental health and well-being of older adolescents. This makes sense because these spaces, unlike natural urban green spaces, offer features (such as benches, playgrounds, and sports fields) that attract adolescents. Adolescents report that they use green spaces mostly for social and physical activities (Bloemsma *et al.* 2018), and parks & gardens and outdoor sports facilities offer opportunities for exactly these activities (more so than natural spaces, such as woodlands or nature reserves).

Why parks and gardens were associated with more emotional problems and (in the sensitivity analysis) conduct problems in 15-year-olds is unclear. This finding could suggest a change in older adolescents (e.g. changing interests), or it could have to do with specific characteristics of parks in London, so the association may be confounded. Living in close proximity to parks does not suggest that these parks have a high quality. In fact, urban parks can be littered or dominated by anti-social behaviour. Older adolescents may pay more attention to the quality of parks, as suggested by Feng and Astell-Burt's (2017) study. Low-quality, urban parks may therefore be associated with poorer mental health. This explanation, again, is speculative, and more research investigating the role of quality in the link between greenspace and mental health is needed.

Study limitations

Before we draw final conclusions, we must note several limitations. First, our study was limited to the London region. London is a large urban area in the southeast of England, so findings cannot be applied to rural areas or, indeed, other urban areas in England.

Furthermore, London may not be representative of urban areas across the world (or even Europe). Therefore, studies in other areas are needed to test whether findings are generalisable across geographies and cultures. Second, the focus on London (rather than the whole of the UK) resulted in a great decrease in sample size. The relatively small sample size makes it more difficult to detect small effect sizes. Third, a related limitation is the large number of tests performed in our study. Due to multiple outcomes, exposures, and age groups, we had to run multiple tests, which increases the probability of a type 1 error. However, due to the generally small effect and sample sizes, and to avoid a type 2 error, we did not correct for multiple tests. We argue that this is a fair approach, especially as we took care to interpret patterns of associations (rather than individual associations). Fourth, our study was prone to exposure misclassification for several reasons: 1) we only had data on the residential neighbourhood but not on other relevant environments (such as the school); 2) the neighbourhood was defined as a circular buffer around a postcode (or as a LSOA); 3) greenspace data were from 2016 (green land cover) and 2020 (green land use), whereas UKHLS data were from 2009-2018; and 4) we did not have data on actual use of greenspaces. Taken together, we have to assume that our exposure variables could only approximate adolescents' true exposure to greenspace. Future studies would benefit from more accurate objective measures of exposure (further addressing issues related to exposure misclassification) but also subjective (self-reported) measures of exposure (complementing objective measures and providing insight into the role of perceived frequency of exposure and use/usage of greenspace). Fifth, although we distinguished between different types of greenspace, we did not include information on specific characteristics (e.g. vegetation, features, or facilities) or quality (e.g. safety, accessibility, or cleanliness). Arguably, these are important factors that may affect whether adolescents visit green spaces and, if yes, what they get out of these visits. Future studies would benefit from including information on a range of characteristics of green spaces. Further, a mixed-methods approach, collecting both quantitative and qualitative data, could be useful. Qualitative data could provide important insight into why adolescents use, and benefit from, certain types of spaces (or not use, or 'dis-benefit' from, other types of spaces). Sixth, there may have been residual confounding, as we could not account for some potentially relevant confounders, such as access to a private garden (i.e. access to proximal greenspace). Finally, we can assume that the association between neighbourhood greenspace and adolescent mental health and well-being is complex and influenced by a range of factors at several levels (e.g. individual, family, and neighbourhood). We were not able to capture this complexity fully, and future studies would benefit from investigating other factors that may play a role in the association (e.g. individual sex, family socio-economic background, and neighbourhood safety). Indeed, as we have observed a potential role of age in our study, it would certainly be interesting to investigate the role of other (individual) factors too. This would allow for a more comprehensive understanding of the association and, in turn, would have more specific implications for real-world applications (e.g. policy, and urban planning and design).

Conclusion

Our study is an important contribution to the literature, as it moves away from generic measures of greenspace quantity to more nuanced measures of different types of greenspace, thereby adding to an evolving stream in the literature. A better understanding of the effects of different types of greenspace could have important implications for policymaking, and urban planning and design (e.g. what types of spaces to provide, i.e. maintain, extend, and/or build). In summary, however, our results do not allow for clear conclusions about what types of greenspaces may be most beneficial for adolescent mental health and well-being. Patterns suggest that parks and gardens and outdoor sports facilities may be more beneficial than natural and semi-natural urban greenspaces or green land cover in general. They also suggest that older adolescents (13 to 15 years) may benefit more from green spaces than younger adolescents (10 to 12 years). However, individual associations and also patterns of associations were not consistent across analyses. Therefore, and in light of the study limitations, results must be viewed with caution and future studies are needed to confirm the patterns observed in this study. Future studies may not only investigate different types of greenspace and greenery but also explore the role of specific characteristics and the quality of green spaces.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Leverhulme Doctoral Training Programme for the Ecological Study of the Brain, under grant number DS-2017-026 (awarded to Marie A. E. Mueller); and the Economic and Social Research Council, under grant number S/N007921/1 (awarded to Eirini Flouri).

Notes on contributors

Marie A. E. Mueller is a scientist based at the Department of Epidemiology and Applied Clinical Research at the UCL

Division of Psychiatry. ***Emily Midouhas*** and ***Eirini Flouri*** are scientists based at the Department of Psychology and Human Development at the UCL Institute of Education. They are interested in physical, social, and biological factors influencing the development (e.g., cognition, behaviour, and mental health) of both typically-developing children and children with special educational needs. MAEM is a research fellow in psychiatric epidemiology, EM is an associate professor, and EF is a professor in developmental psychology.

ORCID

Marie A. E. Mueller http://orcid.org/0000-0001-9616-1201

References

Air Quality Data – London Datastore. data.london.gov.uk. (2022). Retrieved 2 September 2022, from https://data.london.gov.uk/air-quality/.

Amoly, E., *et al.*, 2015. Green and blue spaces and behavioral development in Barcelona schoolchildren: the BREATHE project. *Environmental health perspectives*, 122 (12), 1351–1358. doi:10.1289/ehp.1408215.

Balseviciene, B., *et al.*, 2014. Impact of residential greenness on preschool children's emotional and behavioral problems. *International journal of environmental research and public health*, 11 (7), 6757–6770. doi:10.3390/ijerph110706757.

Bannink, R., Pearce, A., and Hope, S., 2016. Family income and young adolescents' perceived social position: associations with self-esteem and life satisfaction in the UK Millennium Cohort Study. *Archives of disease in childhood*, 101 (10), 917–921. doi:10.1136/archdischild-2015-309651.

Bezold, C.P., *et al.*, 2018. The association between natural environments and depressive symptoms in adolescents living in the United States. *Journal of adolescent health*, 62 (4), 488–495. doi:10.1016/j.jadohealth.2017.10.008.

Bloemsma, L.D., *et al.*, 2018. Green space visits among adolescents: frequency and predictors in the PIAMA birth cohort study. *Environmental health perspectives*, 126 (4), 1–9. doi:10.1289/EHP2429.

Bratman, G.N., Olvera-Alvarez, H.A., and Gross, J.J., 2021. The affective benefits of nature exposure. *Social and personality psychology compass*, 15 (8). doi:10.1111/spc3.12630.

Carstairs, V., Morris, R., and Carstairs, V., 1989. Deprivation, mortality and resource allocation. *Journal of public health (United Kingdom)*, 11 (4), 364–372. doi:10.1093/oxfordjournals.pubmed.a042469.

Christensen, M.K., *et al.*, 2020. The cost of mental disorders: a systematic review. *Epidemiology and psychiatric sciences*, 29. doi:10.1017/S204579602000075X.

Christie, D. and Viner, R., 2005. Adolescent development. *The BMJ*, 330 (7486), 301–304. doi:10.1136/bmj.330.7486.301.

de Girolamo, G., *et al.*, 2012. Age of onset of mental disorders and use of mental health services: needs, opportunities and obstacles. *Epidemiology and psychiatric sciences*, 21 (1), 47–57. doi:10.1017/S2045796011000746.

Dzhambov, A.M., *et al.*, 2018. Residential greenspace might modify the effect of road traffic noise exposure on general mental health in students. *Urban forestry & urban greening*, 34, 233–239. doi:10.1016/j.ufug.2018.06.022.

Feda, D.M., *et al.*, 2015. Neighbourhood parks and reduction in stress among adolescents: results from Buffalo, New York. *Indoor and Built Environment*, 24 (5), 631–639. doi:10.1177/1420326X14535791.

Feng, X. and Astell-Burt, T., 2017. Residential green space quantity and quality and child well-being: a longitudinal study. *American journal of preventive medicine*, 53 (5), 616–624. doi:10.1016/j.amepre.2017.06.035.

Flouri, E., Midouhas, E., and Joshi, H., 2014. The role of urban neighbourhood green space in children's emotional and behavioural resilience. *Journal of environmental psychology*, 40, 179–186. doi:10.1016/j.jenvp.2014.06.007.

Ford, T. and Parker, C., 2016. Emotional and behavioural difficulties and mental (ill)health. *Emotional and behavioural difficulties*, 21 (1), 1–7. doi:10.1080/13632752.2016.1139300.

Goodman, R., 1997. The strengths and difficulties questionnaire: a research note. *Journal of child psychology and psychiatry, and allied disciplines*, 38 (5), 581–586. doi:10.1111/j.1469-7610.1997.tb01545.x.

Goodman, R., Meltzer, H., and Bailey, V., 1998. The strengths and difficulties questionnaire: a pilot study on the validity of the self-report version. *European child & adolescent psychiatry*, 7 (3), 125–130. doi:10.1007/s007870050057.

Greenspace Information for Greater London CIC - GIGL. gigl.org.uk. (2022). Retrieved 2 September 2022, from https://www.gigl.org.uk/.

Houlden, V., *et al.*, 2021. Does nature make us happier? A spatial error model of greenspace types and mental wellbeing. *Environment and planning B: Urban analytics and city science*, 48 (4), 655–670. doi:10.1177/2399808319887395.

Institute for Social and Economic Research, 2020. *Understanding society: waves 1-10, 2009-2019 and Harmonised BHPS: waves 1-18, 1991-2009, user guide.* University of Essex: Colchester. 29 October 2020.

Kelly, Y., *et al.*, 2016. BMI development and early adolescent psychosocial well-being: uK Millennium Cohort Study. *Pediatrics*, 138 (6). doi:10.1542/peds.2016-0967.

Kelly, Y., *et al.*, 2018. Social media use and adolescent mental health: findings from the UK Millennium Cohort Study. *EClinicalMedicine*, 6, 59–68. doi:10.1016/j.eclinm.2018.12.005.

Kessler, R.C., *et al.*, 2005. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the national comorbidity survey replication. *Archives of general psychiatry*, 62 (6), 593. doi:10.1001/archpsyc.62.6.593.

Li, M., *et al.*, 2020. Sound absorption by tree bark. *Applied acoustics*, 165, 165. doi:10.1016/j.apacoust.2020.107328.

London Green and Blue Cover – London Datastore. data.london.gov.uk. (2018). Retrieved 2 September 2022, from https://data.london.gov.uk/dataset/green-and-blue-cover.

Madzia, J., *et al.*, 2019. Residential greenspace association with childhood behavioral outcomes. *The journal of pediatrics*, 207, 233–240. doi:10.1016/j.jpeds.2018.10.061.

Maes, M.J.A., *et al.*, 2021. Benefit of woodland and other natural environments for adolescents' cognition and mental health. *Nature sustainability*, 4 (10), 851–858. doi:10.1038/s41893-021-00751-1.

Markevych, I., *et al.*, 2014. Access to urban green spaces and behavioural problems in children: results from the GINIplus and LISAplus studies. *Environment international*, 71, 29–35. doi:10.1016/j.envint.2014.06.002.

Markevych, I., *et al.*, 2017. Exploring pathways linking greenspace to health: theoretical and methodological guidance. *Environmental research*, 158, 301–317. doi:10.1016/j.envres.2017.06.028.

Mavoa, S., *et al.*, 2019. Natural neighbourhood environments and the emotional health of urban New Zealand adolescents. *Landscape and urban planning*, 191, 191. doi:10.1016/j.landurbplan.2019.103638.

Mills, J.G., *et al.*, 2017. Urban habitat restoration provides a human health benefit through microbiome rewilding: the microbiome rewilding hypothesis. *Restoration ecology*, 25 (6), 866–872. doi:10.1111/rec.12610.

Mueller, M.A.E. and Flouri, E., 2020. Neighbourhood greenspace and children's trajectories of self-regulation: findings from the UK Millennium Cohort Study. *Journal of environmental psychology*, 71, 71. doi:10.1016/j.jenvp.2020.101472.

Mueller, M.A.E. and Flouri, E., 2021. Urban adolescence: the role of neighbourhood greenspace in mental well-being. *Frontiers in psychology*, 12, 12. doi:10.3389/fpsyg.2021.712065.

Mueller, M.A.E., Flouri, E., and Kokosi, T., 2019. The role of the physical environment in adolescent mental health. *Health & place*, 58, 58. doi:10.1016/j.healthplace.2019.102153.

Nguyen, T., *et al.*, 2015. Relationship between types of urban forest and PM2.5 capture at three growth stages of leaves. *Journal of environmental sciences (China)*, 27 (C), 33–41. doi:10.1016/j.jes.2014.04.019.

NHS Digital, 2018. *Mental health of children and young people in England, 2017*. PAS. NHS Digital. https://files.digital.nhs.uk/A0/273EE3/MHCYP%202017%20Trends%20Characteristics.pdf

Passmore, H.-A. and Holder, M.D., 2017. Noticing nature: individual and social benefits of a two-week intervention. *The journal of positive psychology*, 12 (6), 537–546. doi:10.1080/17439760.2016.1221126.

Pebesma, E., 2018. Simple features for R: standardized support for spatial vector data. *The R journal*, 10 (1), 439–446. doi:10.32614/RJ-2018-009.

Planning Policy Guidance 17: Planning for open space, sport and recreation. webarchive.nationalarchives.gov.uk. (2012). Retrieved 2 September 2022, from https://webarchive.nationalarchives.gov.uk/ukgwa/20120919132719/http://www.communities.gov.uk/documents/planningandbuilding/pdf/ppg17.pdf.

Raghunathan, T., *et al.*, 2001. A multivariate technique for multiply imputing missing values using a sequence of regression models. *Survey methodology*, 27 (1), 85–96.

Richardson, E.A., *et al.*, 2017. The role of public and private natural space in children's social, emotional and behavioural development in Scotland: a longitudinal study. *Environmental research*, 158, 729–736. doi:10.1016/j.envres.2017.07.038.

Rosenberg, M., 1965. *Society and the adolescent self-image*. Princeton University Press. https://www.jstor.org/stable/j.ctt183pjjh

Rubin, D.B., 1987. *Multiple imputation for nonresponse in surveys*. New York: John Wiley.

Shaw, B., *et al.* (2015). Children's independent mobility: an international comparison and recommendations for action. In Policy Studies Institute. Retrieved from http://www.psi.org.uk/docs/7350_PSI_Report_CIM_final.pdf

Stevenson, M.P., Schilhab, T., and Bentsen, P., 2018. Attention restoration theory II: a systematic review to clarify attention processes affected by exposure to natural environments. *Journal of toxicology and environmental health - part B: critical reviews*, 21 (4), 227–268. doi:10.1080/10937404.2018.1505571.

Suhrcke, M., Pillas, D., and Selai, C. 2007. Economic aspects of mental health in children and adolescents. *Geneva: WHO: Social Cohesion for Mental Wellbeing among adolescents*.

Taylor, A.F., Kuo, F.E., and Sullivan, W.C., 2002. Views of nature and self-discipline: evidence from inner city children. *Journal of environmental psychology*, 22 (1–2), 49–63. doi:10.1006/jevp.2001.0241.

Teucher, A. and Russell, K. (2022). Rmapshaper: client for 'mapshaper' for 'Geospatial' operations. R package version 0.4.6. https://CRAN.R-project.org/package=rmapshaper.

Thygesen, M., *et al.*, 2020. The association between residential green space in childhood and development of attention deficit hyperactivity disorder: a population-based cohort study. *Environmental health perspectives*, 128 (12), 1–9. doi:10.1289/EHP6729.

University of Essex, Institute for Social and Economic Research. (2020). Understanding society: waves 1-9, 2009-2018 and harmonised BHPS: waves 1-18, 1991-2009. [data collection]. 12th Edition. UK Data Service. SN: 6614, 10.5255/UKDA-SN-6614-13.

University of Essex, Institute for Social and Economic Research, 2021. Understanding society: waves 1-10, 2009-2019: special licence access. *Census 2011 lower layer super output areas. [data collection]*. 10th. UK Data Service. SN: 7248. 10.5255/UKDA-SN-7248-10.

University of Essex, Institute for Social and Economic Research. (2022). Understanding Society: Waves 1-11, 2009-2020 and Harmonised BHPS: Waves 1-18, 1991-2009: Secure Access. [data collection]. 13th Edition. UK Data Service. SN: 6676, 10.5255/UKDA-SN-6676-13.

Wang, P., *et al.*, 2019. Green space and serious psychological distress among adults and teens: a population-based study in California. *Health & place*, 56, 184–190. doi:10.1016/j.healthplace.2019.02.002.

Wheeler, B., 2019. *Carstairs index 2011 for lower-layer super output areas*. [Data Collection]. Colchester, Essex: UK Data Archive.10.5255/UKDA-SN-851497.

Wickham, H., *et al.*, 2019. Welcome to the tidyverse. *Journal of open source software*, 4 (43), 1686. doi:10.21105/joss.01686.

Younan, D., *et al.*, 2016. Environmental determinants of aggression in adolescents: role of urban neighborhood greenspace. *Journal of the American academy of child and adolescent psychiatry*, 55 (7), 591–601. doi:10.1016/j.jaac.2016.05.002.