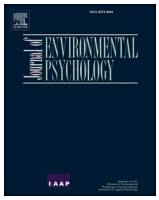
Journal of Environmental Psychology 93 (2024) 102208

Contents lists available at ScienceDirect

Journal of Environmental Psychology

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



Capturing the well-being exposome in poly-environmental scores.

M.P. van de Weijer [a,b,c,1,*], Dirk H.M. Pelt [a,b], Bart M.L. Baselmans [d], Lannie Ligthart [a,b], Floris Huider [a,b], Jouke-Jan Hottenga [a,b], René Pool [a,b], Meike Bartels [a,b]

[a] Department of Biological Psychology, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands
[b] Amsterdam Public Health Research Institute, Amsterdam University Medical Centres, Amsterdam, the Netherlands
[c] Genetic Epidemiology, Department of Psychiatry, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, the Netherlands
[d] Biomedical Technology, Faculty of Technology, Amsterdam University of Applied Sciences, Amsterdam, the Netherlands

ARTICLE INFO

Handling Editor: W. Schultz

Keywords:
Well-being
Happiness
Exposome
Environment
Gene-environment correlation

ABSTRACT

In this study we use aggregated weighted scores of environmental effects to study environmental influences on well-being and happiness. To this end, we split a sample of Netherlands Twin Register (NTR) participants into a training ($N = 4857$) and test ($N = 2077$) sample. In the training sample, we use elastic net regression to estimate effect sizes for associations between life satisfaction and two sets of environmental variables: one based on self-report socioenvironmental data, and one based on objective physical environmental data. Based on these effect sizes, we create two poly-environmental scores (PES-S and PES-O, for self-reports and objective data respectively). In the test sample, we perform association analyses between different measures of well-being and the two PESs. We find that the PES-S explains ~36% of the variance in well-being, while the PES-O does not significantly contribute to the model. Variance in other well-being measures (i.e., different life satisfaction domains, subjective happiness, quality of life, flourishing, psychological well-being, self-rated health, depressive problems, and loneliness) are explained to varying extents by the PESs, ranging from 6.36% (self-rated health) to 36.66% (loneliness). These predictive values did not change during the COVID-19 pandemic ($N = 3214$). Validating the PES-S in the UK biobank ($N = 40{,}614$), we find that the UK biobank PES-S explains about ~12% of the variance in happiness. Lastly, we examine if there is any indication for gene-environment correlation (rGE), the phenomenon where one's genetic predisposition influences exposure to the environment, by associating the PESs with polygenic scores (PGS) in a sample of Netherlands Twin Register (NTR) and UK Biobank participants. While the PES and PGS were not correlated in the NTR sample, they were correlated in the larger UK biobank sample, indicating the potential presence of rGE. We discuss several limitations pertaining to our dataset, such as a potential influence of common method bias, and reflect on how PESs might be used in future research.

1. Introduction

Many (socio) environmental exposures have been associated with human well-being. For example, meta-analyses suggest a role for social support (Jiameng & Huiyong, 2013), green space exposure (Houlden, Weich, Albuquerque, Jarvis, & Rees, 2018), and socioeconomic status (Tan, Kraus, Carpenter, & Adler, 2020), among many other factors. The totality of these environmental exposures can collectively be referred to as the well-being *exposome* (Wild, 2012), which captures all non-genetic exposures influencing variation in well-being from conception onwards (also referred to by others as the *environome* (von Stumm & d'Apice, 2022)). Approximately 60–70% of individual differences in well-being can be traced back to this exposome (Bartels, 2015; Nes & Roysamb, 2015). Complementary to the exposome is the genome, all our genetic information, which accounts for the other 30–40% of individual differences in well-being.

While there are many studies examining associations between environmental factors and well-being, they mostly follow a "pick-and-choose" approach, where potential risk factors are selected beforehand based on existing hypotheses, which can result in selective reporting (cherry picking) and overestimation of effects (publication bias, false positives). To overcome these problems, we can conduct an environment-wide association study, where many individual environmental effects are simultaneously examined in relation to well-being in a

* Corresponding author.Van der Boechorststraat 7, 1081 BT, Amsterdam, the Netherlands.
E-mail address: m.p.vandeweijer@amsterdamumc.nl (M.P. van de Weijer).
[1] Current affiliation.

https://doi.org/10.1016/j.jenvp.2023.102208
Received 17 April 2023; Received in revised form 30 November 2023; Accepted 2 December 2023
Available online 10 December 2023
0272-4944/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

hypothesis-free fashion, similar to the genome-wide association approach. By examining associations between well-being and 139 objective environmental indicators (e.g., green space, livability) in such an environment-wide association study, neighborhood variables related to socioeconomic status and safety were significantly associated with individual-level well-being (van de Weijer, Baselmans, et al., 2022).

In the genetics field it is common to perform so-called genome-wide association studies, where millions of genetic variants are associated with an outcome without a priori hypothesizing which variants might be important for the trait of interest. One of the ways in which genome-wide association findings are applied in both scientific and clinical contexts is by combining the resulting effect sizes (reflecting the strength of associations between individual genetic variants and the outcome of interest) into so-called polygenic scores (PGS) (Lewis & Vassos, 2020). These scores reflect an individual's genetic propensity or risk for a certain outcome. PGSs can be used for disease risk estimation, cross-phenotype prediction, and for answering specific research questions such as whether the effect of individual environmental exposures varies over strata of individuals with varying genetic propensities for a specific outcome (Lewis & Vassos, 2020). In previous research, well-being PGSs have been used to predict not only well-being measures (Baselmans et al., 2019), but several other phenotypes such as loneliness (Abdellaoui et al., 2018), childhood psychopathology (Akingbuwa et al., 2020), and brain morphololgy (Song et al., 2019). Similarly, we can take our investigation of the exposome one step further by summing individual environmental risk across multiple environmental factors into poly-environmental or poly-environmental scores (PESs).

One of the potential difficulties in constructing PESs is that environmental factors are correlated, and that we should take these correlations into account in some way so that we do not overestimate the effect of the environment. However, correlations between environmental factors are much complex and dynamic, complicating the construction of PESs. Existing studies using PESs have focused mainly on disease outcomes such as schizophrenia and psychosis (Jeon et al., 2022; Padmanabhan, Shah, Tandon, & Keshavan, 2017; Pries et al., 2019, 2021; Vassos et al., 2020), where the scores are used to identify at-risk individuals for these outcomes. Overall, these scores seem to be able to explain 10–20% of the variation in case-control status. The manner in which the environmental factors are combined in the calculation of these PESs varies considerably, with some simply summing estimates from systematic reviews or meta-analyses (Deckers et al., 2019; Oliver, Radua, Reichenberg, Uher, & Fusar-Poli, 2019; Padmanabhan et al., 2017) (without correcting for potential correlations), and others using training datasets where different prediction techniques are used to calculate weights that are used for weighing the estimates (He, Lakhani, Manrai, & Patel, 2019; Pries et al., 2019) in a test sample. An important distinction between these two approaches is that the latter takes correlations between predictors into account by weighting the different exposures, while the former does not.

Similar to PGSs, PESs could be used for well-being and cross-phenotype prediction. An interesting question in this regard is to what extent environmental factors predicting one specific well-being measure also predict other well-being measures. In psychology, many different definitions of well-being are employed, which can traditionally be categorized in two larger domains: subjective well-being and psychological well-being (Deci & Ryan, 2008). Subjective well-being can be defined as the presence of positive affect and absence of negative affect and is often measured by inquiring about one's general satisfaction with life or happiness (Diener, Suh, Lucas, & Smith, 1999). Psychological well-being, on the other hand, involves aspects such as self-actualization (i.e., the fulfilment of human potential) and living a meaningful life (Ryff & Singer, 2008). The extent to which subjective- and psychological well-being constructs are distinguishable or overlap is a subject of long debate (Baselmans & Bartels, 2018; van de Weijer, Baselmans, van der Deijl, & Bartels, 2018). Because of this, it is also unclear if different well-being definitions are exchangeable when examining environmental correlates. Comparing the predictive power of PESs (based on one well-being phenotype) across different well-being outcomes would allow us to get better insight into the generalization of the exposome across different well-being measures.

In addition, the PESs can be useful for stratifying individuals based on environmental (instead of genetic) 'risk or protection' (i.e., which individuals live in environments that stimulate well-being). Moreover, it allows us to broaden our understanding of the interplay between the genome and exposome with the use of research designs that combine PGSs and PESs. For example, it presents us with a new opportunity to study gene-environment correlation (rGE), the phenomenon where ones genetic predisposition influences exposure to the environment (Dick, 2011). There are three types of gene-environment correlation: 1) active, where a person's heritable traits cause them to actively select certain types of environment (e.g., people with a high predisposition for well-being select environments that stimulate their well-being), 2) evocative, where a person's heritable traits elicit a reaction in other people, which in turn influences one's environment (e.g., people with a high predisposition for well-being elicit positive responses and moods in others), and 3) passive, where genotype and environment become correlated because a child inherits both genes and a familial environment from their parents (e.g., a child receives both well-being increasing genetic variants and a well-being stimulating environment from its parents) (Jaffee & Price, 2007). Previously, researchers have studied rGE by correlating polygenic scores with individual environmental variables. For example, von Stumm and colleagues found correlations between an education polygenic score and different home, neighborhood and adversity variables (von Stumm, Kandaswamy, & Maxwell, 2023).

The objective of the current study is to explore the potential of PESs in the context of well-being. First, we calculate effect size estimates for two sets of exposome variables (one including objective environmental indicators and one including subjective environmental evaluations) in predicting life satisfaction scores (a measure of subjective well-being) in a Dutch training set using elastic net regression. Next, we use these effect sizes as weights to construct PESs in an independent test set and predict life satisfaction and several other well-being outcomes in this test set. Based on the previous finding that environmental variance for well-being increased during the pandemic (van de Weijer et al., 2022), we additionally examine if the predictive power of the PESs change from before to during the COVID-19 pandemic. Lastly, we examine if the life satisfaction PES is associated with a well-being PGS to assess potential gene-environment correlation. We compare these findings with a follow-up analysis in the UK Biobank, where we create a similar PES based on socioenvironmental variables and correlate it with the well-being PGS.

2. Methods and methods

This study was pre-registered at the open science framework (OSF) (https://osf.io/5x8kf). Additional follow-up analyses are indicated as "non-preregistered".

2.1. Sample

For the present study, we used data from two large cohort studies: 1) a sample of Netherlands Twin Register participants was used as the primary sample for constructing and testing the poly-environmental scores, and for (non-preregistered) cross-validation of the objective PES in a larger sample. 2) A sample of UKB participants was used as a (non-preregistered) cross-validation sample. For the subjective PES, another (independent) Netherlands Twin Register sample was used for cross-validation. Below, we describe the sample and measures for each cohort separately.

2.1.1. Netherlands twin register

We used a sample of participants from the Netherlands Twin Register (NTR) for our primary analyses (Ligthart et al., 2019). We included multiples (e.g., twins, triplets) and family members (e.g., parents, partners) who filled out the most recent wave of survey data collection (wave 14 collected in 2019–2022). We use two separate datasets: the first dataset consists of data collected just prior to the COVID-19 pandemic in the Netherlands (collected between June 2019 and February 2020). We randomly divide this pre-pandemic sample into a training (70%) and test (30%) sample in order to construct the scores and use them for prediction, respectively. 2) The second dataset was collected between February 2020 and May 2022, during the COVID-19 pandemic. We use this dataset to examine if the pre-pandemic PESs predict well-being to a similar extent during the pandemic.

In both samples, we remove genetic relatedness by randomly including one individual from genetically related family members, leading to final sample of $N_{max} = 6092$ individuals with satisfaction with life data (our primary well-being measure) in the pre-pandemic sample, and $N_{max} = 3214$ individuals with satisfaction with life data in the pandemic sample. Sample sizes vary across analyses and outcomes based on missingness. Sample characteristics can be found in Table 1.

In follow-up analyses, we additionally examine the predictive power of the objective PES on well-being in a larger sample from an earlier NTR survey (survey 12), collected between 2016 and 2019. This sample consisted out of 6023 individuals between 16 and 90 years old ($M = 44.34$, $SD = 14.55$). Since satisfaction with life was not collected in this survey, we used quality of life as a well-being outcome measure ($M = 7.70$, $SD = 1.06$).

Table 1
Sample descriptives pre-registered NTR analyses.

	NTR			
	full sample[a] (pre-pandemic)	training sample[b]	test sample[b]	pandemic sample
N males/females	1779/4313	1425/3432	608/1469	886/2328
M (SD) age	49.94 (12.22)	50.01 (12.19)	49.98 (12.39)	35.18 (13.32)
Age range	16–88	16–88	18–86	18–88
M (SD) life satisfaction	27.50 (5.11)	27.50 (5.14)	27.49 (5.04)	27.08 (5.48)
M (SD) family satisfaction	4.66 (.78)	4.66 (.78)	4.65 (.78)	4.74 (.88)
M (SD) financial satisfaction	4.61 (.80)	4.61 (.81)	4.60 (.79)	4.53 (.88)
M (SD) work satisfaction	4.57 (.85)	4.58 (.84)	4.55 (.87)	4.51 (.92)
M (SD) health satisfaction	4.56 (.88)	4.56 (.87)	4.56 (.90)	4.66 (.94)
M (SD) friendship satisfaction	4.71 (.75)	4.72 (.74)	4.68 (.78)	4.81 (.83)
Subjective Happiness	22.81 (4.21)	22.79 (4.22)	22.87 (4.16)	22.10 (4.78)
Quality of Life	7.84 (1.06)	7.83 (1.07)	7.86 (1.03)	7.65 (1.22)
Flourishing	46.18 (5.70)	46.18 (5.70)	46.18 (5.70)	46.02 (6.04)
Psychological well-being	20.73 (5.71)	20.76 (5.68)	20.65 (5.78)	20.47 (5.83)
Self-Rated Health	3.95 (.70)	3.95 (.70)	3.95 (.69)	4.04 (.71)
Depressive Problems	3.64 (3.64)	3.62 (3.64)	3.69 (3.65)	4.56 (4.37)
Loneliness	4.04 (1.35)	4.02 (1.34)	4.10 (1.37)	4.44 (1.51)

[a] Full sample indicates the full set of individuals with non-missing life satisfaction data.

[b] Actual sample size per analysis is lower depending on the amount of missingness per PES type.

2.1.2. UK biobank

Data from the UK Biobank (UKB) were used as a cross-validation sample to validate our elastic net poly-environmental approach (Bycroft et al., 2018) for the subjective PES. The UKB is a large cohort study with phenotypic, genetic, and biological data of UK individuals recruited between the ages of 40 and 69. The sample used in this study consisted of $N = 40{,}614$ individuals with non-missing phenotype data. Similar to the NTR analyses, we split the sample in a training (70%) and test (30%) sample and standardized all continuous predictors.

2.2. Measures

2.2.1. Netherlands twin register

2.2.1.1. Outcome measures. *Satisfaction with life* was assessed using the 5-item satisfaction with life scale (Diener, Emmons, Larsen, & Griffin, 1985). Individual items are scored on a scale from 1 to 7, with higher scores indicating higher levels of satisfaction with life. The item responses are combined into a sum-score ranging from 7 to 35 (Cronbach's $\alpha = 0.86$).

Domain-specific satisfaction items were assessed *family satisfaction, financial satisfaction, friendship satisfaction, work satisfaction*, and *health satisfaction*. For each domain, participants were asked 'in general, how satisfied are you with […]?'. The items were coded on a scale from 1 to 6, with 1 indicating extremely unhappy, and 6 indicating extremely happy.

Subjective Happiness was assessed using the 4-item subjective happiness scale (SHS) (Lyubomirsky & Lepper, 1999). The items are rated on a Likert scale ranging from 1 to 7. We recoded two reverse-coded items so that for all items, a higher score indicated higher levels of happiness. The items were combined into a sum-score ranging from 4 to 28 (Cronbach's $\alpha = 0.87$).

Quality of Life was assessed using the single-item Cantril ladder (Cantril, 1966). Participants are asked to answer the question "Where on the scale would you place your life in general?", on a scale from 0 (indicating the worst possible life) to 10 (representing the best possible life).

Flourishing was assessed using the 8-item Short Flourishing Scale (Diener et al., 2010). We combined the individual items, which are rated on a Likert scale ranging from 1 to 7, into a sum-score, where higher values indicate higher levels of flourishing (Cronbach's $\alpha = 0.90$).

Psychological well-being was assessed using the psychological well-being (PWB) subscale of the mental health continuum short form (Keyes, 2002). This subscale consists of 6 items rated on a scale from 0 to 5. Item responses were summed to create scores ranging from 0 to 30, with higher scores indicating higher levels of PWB (Cronbach's $\alpha = 0.92$).

Self-rated health was assessed with a single item "How would you rate your health in general?" (Eriksson, Undén, & Elofsson, 2001). The item was rated on a 5-point scale ranging from 'Bad' (1) to 'Excellent' (5).

Loneliness was assessed using the 3-item short scale for assessing loneliness in large epidemiological studies (Hughes, Waite, Hawkley, & Cacioppo, 2004). For each item, participants were asked to indicate how often they identified with a statement (e.g., "how often do you feel isolated from others?"), rated as 0 = almost never, 1 = sometimes, 2 = often. The items were summed to obtain a loneliness sum-score (Cronbach's $\alpha = 0.79$).

Depressive problems were assessed with the ASR DSM Depressive Problems scale (Achenbach, Bernstein, & Dumenci, 2005). The scale consists of 14 items, where each item is rated from 0 = not true, 1 = somewhat true, to 2 = very true. The items were summed to obtain a depressive problems sum-score where higher values indicate higher levels of depressive problems (Cronbach's $\alpha = 0.82$).

2.2.1.2. Predictors. We include two sets of environmental exposures:

1. Objective Environmental Exposures: we used a set of 69 objective environmental measures obtained from the Geoscience and Health Cohort Consortium (GECCO) (Timmermans et al., 2018), linked to the NTR data based on 4-numeric postal code. We previously included an overlapping set of variables in an environment-wide association study (van de Weijer, Baselmans, et al., 2022). The variables reflect aspects of the physical environment (i.e., amount of traffic in the area), culture, socioeconomic status, accessibility, education, liveability, care, and sports. The included variables reflect time-points between 2017 and 2020. A complete overview and descriptives of these variables can be found in Supplementary Table 1.
2. Subjective Environmental Exposures: A set of 21 subjective socio-environmental indicators included in the same NTR survey wave as the well-being data. These variables reflect participants' subjective evaluations on their relationships, life events, social support, leisure time activities, education, stress, and online and offline social contact. A complete overview and descriptives of these variables can be found in Supplementary Table 2.

All continuous variables were standardized to Z-scores prior to analyses.

2.2.1.3. Postal code linkage. The GECCO and NTR data were linked based on self-report postal code data from NTR participants that gave permission for data linkage. When participants register for the NTR, they are asked to provide their address. Participants are asked to contact us when they move so that their address can be updated in our administrative database. In addition, addresses are regularly updated by cross-checking with the Dutch Personal Records Database (PRD) (in Dutch: Basis Registratie Personen, BRP), and by enquiring about postal codes in several survey waves and NTR newsletters. For the current project, we used data from the pre-pandemic and pandemic wave 14 NTR data collection. In the pandemic survey, we included a question asking participants to report their current four-numeric postal code. For those individuals, we used this self-reported postal code. In the pre-pandemic survey, this item was not included. For a subset of 2623 participants of the pre-pandemic sample, addresses were cross-checked with the PRD in September 2022. For this subset, 85% of known postal codes matched with the postal codes reported in PRD. Since the pre-pandemic sample was collected 2–3 years prior to the PRD check, postal code listed at NTR at the time of data collection was used as the postal code for linkage.

2.2.1.4. Covariates. As phenotypic covariates, we include sex, age, and age^2. For the analyses that include PGSs, we additionally include genotyping platform dummies and the first 10 genomic principal components (PCs) to control for population structure. Smartpca was used to calculate the PCs, using LD-pruned 1000 Genomes imputed genetic variants genotyped on one or more platforms.

2.2.1.5. Genotype data and polygenic scores. Genotype data were available for $N = 558$ participants in the test sample. Detailed information on genotyping and quality control can be found in the supplementary methods. Polygenic scores were calculated from summary statistics from a well-being spectrum genome-wide association study (Baselmans, Jansen, et al., 2019), excluding NTR participants. Before constructing the scores, effect sizes were re-estimated, taking into account linkage disequilibrium (the correlation between genetic variants), with LDpred 0.9 (Vilhjálmsson et al., 2015) (using an infinitesimal prior and 1000 Genomes phase 3 CEU as a reference panel). The re-estimated effect sizes were used for constructing polygenic scores in PLINK (Purcell et al., 2007).

2.2.2. UKB

2.2.2.1. Outcome measure. A single item on *happiness* (UKB ID 4526) was used an outcome measure. Participants were asked to answer the question "In general how happy are you?" on a scale from 1 to 6, 1 indicating extremely happy, and 6 indicating extremely unhappy. We reverse coded the item so that higher scores indicated higher levels of happiness ($M = 4.47$, $SD = 0.69$). This item most closely resembles subjective happiness from the included NTR well-being measures.

2.2.2.2. Predictors. We selected socioenvironmental variables that most closely resembled the ones that were selected in the NTR data, resulting in a selection of $N = 15$ variables related to mobile phone use, time spend exercising, leisure time activities, educational attainment, life events, and social support (see Supplementary Table 3 for details on the included predictor variables).

2.2.2.3. Covariates. Similar to the NTR analyses, we include sex, age, and age^2 as phenotypic covariates. For the analyses that include PGSs, we include the first 10 genomic PCs to control for population structure.

2.2.2.4. Genotype data and polygenic scores. Detailed information on genotyping procedures in UKB can be found elsewhere (Bycroft et al., 2018). To account for linkage disequilibrium, the well-being spectrum summary statistics (Baselmans, Jansen, et al., 2019) (excluding UK participants) were reanalyzed with SBLUP (Robinson et al., 2017), using a reference sample of 10,000 random unrelated UKB participants. Next, the re-estimated effect sizes were used to generate PGSs in PLINK (Purcell et al., 2007). These PGSs were used to predict the PESs, using the first 10 genomic PCs, batch, age, age^2, and sex as covariates.

2.3. Statistical analyses

2.3.1. Elastic net model effect size estimation

We randomly divide our pre-pandemic sample into a training (70%) and test (30%) sample. The training sample is used to estimate effect sizes to use as weights to calculate our PES in the test set. As a primary well-being measure, we included life satisfaction. We fit two elastic net regression models in the training sample where we predict life satisfaction scores with 1) the objective environmental indicators ($N = 2152$), and 2) the subjective environmental variables ($N = 2297$). The phenotypic covariates were included in both models. While we use the same training sample for both elastic net models, the sample size varies slightly based on missingness of the predictors. Elastic net regression is a selection and shrinkage method that combines the penalties from ridge and lasso regression to optimally deal with correlated predictors and prevent overfitting. Both ridge and lasso regression use a penalty which shrinks the model coefficients. In ridge regression, the penalty is the sum of squared coefficients, while in lasso the penalty equals the sum of the coefficients. In this way, lasso sets parameters to zero while ridge only minimizes the parameters. The extent to which the ridge or lasso penalty is applied in elastic net regression is determined by a tuning parameter called alpha. The resulting penalty/shrinkage coefficient is called lambda. The best alpha and lambda are based on the lowest RMSE value from 10-fold cross-validation (the model is fit in 10 subsamples, and each subsample is used as a holdout set once with the remaining samples used as training sets) using the *caret* R package (Kuhn, 2022).

2.3.2. Poly-environmental score prediction

The non-zero coefficients from the model with the optimal tuning parameters are used for constructing the objective and subjective PESs for life satisfaction in the NTR test sample. The PESs were calculated by summing the predictor variables weighted by their respective coefficients from the training set elastic net models. The PESs were standardized so that resulting associations reflect the impact of an SD increase in the PESs on the different outcomes. We refer to the PES based on objective indicators as the PES-O, and to the PES based on the subjective indicators as the PES-S. We include the PESs (and phenotypic

covariates) as predictors in two separate, and one combined linear regression models where we predict life satisfaction scores.

2.3.3. Cross phenotype prediction

We assess cross-phenotype overlap in exposome associations by using the PESs based on life satisfaction data as predictors in models where we predict the other well-being (-related) phenotypes: subjective happiness, the specific satisfaction domains, quality of life, flourishing, psychological well-being, self-rated health, loneliness, and depressive problems. Since we run separate models for all 12 phenotypes, we use a corrected significance threshold of $\alpha = 0.05/12 = 0.004$.

2.3.4. Prediction during the corona pandemic

Part of our well-being data collection took place during the COVID-19 pandemic. Since the effect of environmental factors might be different during the pandemic, we separately analyzed this part of the sample ($N = 1898$) to examine whether the PESs predict life satisfaction and the other well-being related phenotypes similarly or differently during the pandemic. Confidence intervals, calculated using the *psychometric* R package (Olkin & Finn, 1995) were used to compare the R^2 of the prediction models before the pandemic to the R^2 of the prediction models during the pandemic.

2.3.5. Gene-environment interplay

In a sub-sample of the test set that has genotype data available ($N = 556$), we assessed potential gene-environment correlations. To this end, we include polygenic scores (PGSs) in models predicting the PESs (for a description of how these PGSs were constructed, see "NTR genotype data and polygenic scores" section above). We predict the PESs using well-being spectrum PGSs (Baselmans, Jansen, et al., 2019), including all phenotypic and genetic covariates (see covariate section above). If the PGS significantly predicts the PES ($\alpha = 0.05$), it indicates the genetic predisposition for well-being is associated with the exposure to environmental influences. This association can reflect passive, evocative or active gene-environmental correlations and indicates that environmental exposure is not (only) a random process.

2.3.6. NTR cross-validation

For the NTR cross-validation, well-being data were similarly linked to the GECCO data using 4-numeric postal code. To calculate the PES, we used the weights from the original pre-pandemic elastic net regression and used the full sample for prediction, including age, age^2 and sex as covariates.

2.3.7. UKB cross-validation

For the UKB cross-validation, we split the UKB sample in a training (70%) and test (30%) sample and standardized all continuous predictors. The training sample was used to calculate elastic net effect size estimates (using the same methods as in NTR), which were combined into PESs in the test sample. Age, age^2, and sex were included as covariates in both the training and test stage.

3. Results

3.1. Elastic net model effect size estimation and poly-environmental score prediction

The best model for the objective environmental indicators, selected through 10-fold cross validation, used an α penalty of 1, indicating that the lasso penalty function was used (i.e., regression coefficients are shrunk toward zero). The lambda parameter, which controls the weighting of the sum of both penalties, was set to 0.07, indicating that the penalty is weighted down substantially (indicating that the penalty is applied to a much lesser extent than it would have with full regularization). In the final model, 40 variables were set to zero, leading to the inclusion of 29 variables with non-zero coefficients in the PES-O (see Supplementary Table 4). When the PES-O were used to predict life satisfaction scores in the test sample ($N = 949$), it explained 0.4% of the variance ($p = .052$) in life satisfaction scores in the test data after adjusting for sex, age, and age^2.

The most optimal model for the subjective environmental indicators, selected through 10-fold cross validation, used an α penalty of .55, indicating that about equal weight was given to the ridge and lasso penalties. The λ parameter was set to 0.05, again indicating that the penalty is weighted down substantially. In the model, 4 coefficients were set to zero, leading to 17 variables with non-zero coefficients that were used for constructing the PES-S (see Supplementary Table 4 for details on these variables). The final variables included in both scores are depicted in Fig. 1. The PES-S explained 37.16% of the variance in life satisfaction scores in the test data ($N = 1155$), after adjusting for sex, age, and age^2. Combined in one model ($N = 722$), the two PESs explained 35.38% of the variation in life satisfaction. Only the PES-S ($\beta = 3.05$, $SE = 0.16$, $p < 2 \times 10^{-16}$), and not the PES-O ($\beta = 0.07$, $SE = 0.15$, $p = .65$), significantly predicted life satisfaction in the combined model. The two PESs were uncorrelated ($r = 0.07$, $p = .06$).

3.2. Cross phenotype prediction

To assess the extent to which environmental predictors overlap between life satisfaction and other well-being related phenotypes, we used the two PESs to predict 12 other phenotypes. The amount of variance explained by the PESs in the other outcomes ranged from 6.36% (self-rated health) to 36.66% (loneliness). For all outcomes, only the PES-S (and not the PES-O) was a significant predictor (see Fig. 2 and Table 2).

3.3. Prediction during the COVID pandemic

We used a sample of individuals who filled out the survey during the COVID-19 pandemic to assess if the PESs predicted life satisfaction and other well-being related phenotypes equally well during the pandemic. Using life satisfaction itself as the outcome, the variance explained by the prediction model including the two PESs did not change significantly ($R^2 = 34.33\%$, CI = 30.84–37.82%), indicating that the included environmental factors do not predict well-being to a lesser extent in these changed environmental circumstances. For all other phenotypes, the amount of variance explained by the two environmental scores combined was also similar (see Table 2). The PES-S was a significant predictor for all phenotypes during the pandemic, whereas the PES-O did not predict any of the phenotypes.

3.4. Gene-environment interplay

We computed well-being spectrum polygenic scores for the subset of participants in the test set that had genotype data available ($N = 558$). Correlations between the PGS (residualized for platform and PCs), well-being, and the PESs can be found in Table 3. The polygenic score did not significantly predict well-being when corrected for all relevant covariates ($\beta = 0.44$, $SE = 0.23$, $p = .06$). Similarly, the PGS did not significantly predict either the PES-S ($\beta = 0.11$, $SE = 0.06$, $p = .07$, $N = 335$), or the PES-O ($\beta = .-0.07$, $SE = 0.05$, $p = .22$, $N = 353$).

3.5. NTR cross-validation

We used a larger NTR sample with data collected in an earlier survey ($N = 6023$) to predict quality of life with the PES-O. The PES-O did significantly predict quality of life in this sample ($\beta = 0.05$, $SE = 0.01$, $p = .0001$), but explained only a small percentage of the variance (0.2%).

3.6. UK biobank follow-up

In the most optimal training set prediction model ($\alpha = 0.1$, $\lambda = 0.0004$), none of the variables were set to zero, leading to 15 variables

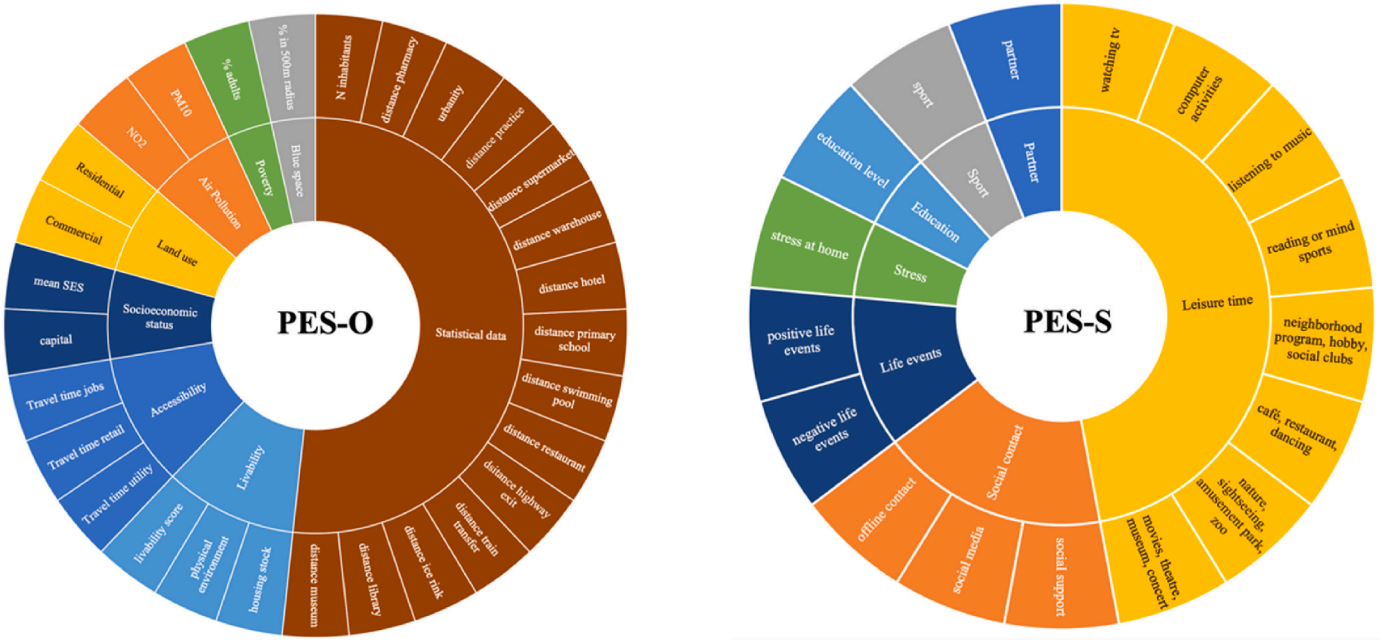


Fig. 1. Variables finally included in the objective poly-environmental score (PES-O, left) and subjective poly-environmental score (PES-S, right).

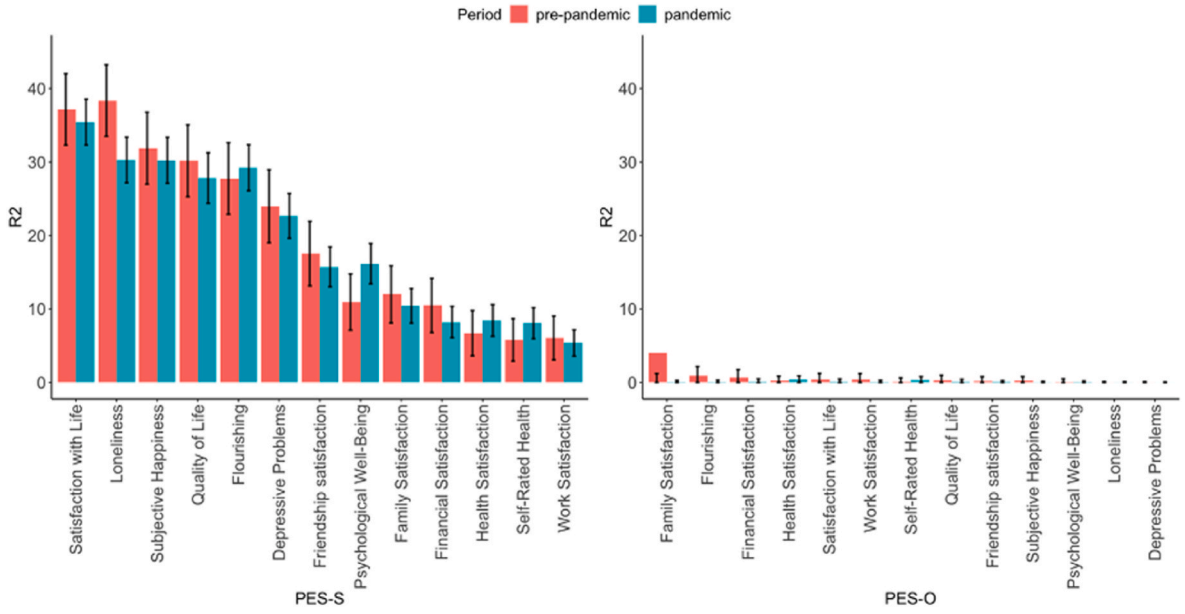


Fig. 2. Variance explained by the subjective poly-environmental score (PES-S) and objective poly-environmental score (PES-O) model on outcomes before and during the pandemic, including 95% confidence intervals.

with non-zero coefficients used for constructing the UKB PES-S in the test sample (see Supplementary Table 5 for the elastic net estimates). The PES explained 12.41% of the variance in happiness scores in the test data ($\beta = 0.24$, $SE = 0.01$, $p < 2 \times 10\text{-}^{16}$, $N = 12{,}184$), after adjusting for sex, age, and age^2.

Correlations between the PGS (residualized for platform and PCs), well-being, and the PESs can be found in Table 3. The PGS significantly predicted happiness when corrected for all relevant covariates ($\beta = 0.05$, $SE = 0.01$, $p = <2 \times 10^{-16}$, $N = 12{,}184$) but only explained a small amount of the variance (0.04%). Moreover, the PGS also significantly predicted the PES ($\beta = 0.07$, $SE = 0.01$, $p = 4.05 \times 10^{-14}$), similarly explaining .04% of the variance.

4. Discussion

This study examined the potential of combining multiple environmental correlates of well-being into well-being poly-environmental scores (PESs). To this end, we constructed two different PESs: one reflecting self-reported socio-environmental factors (the PES-S), and one reflecting objective (postal-code level) physical environmental factors (the PES-O). Moreover, we examined potential gene-environment correlation by associating well-being PESs with well-being PGSs. Lastly, we performed validation efforts in a larger NTR and UKB sample.

With respect to the predictive power of the PESs, we found a large difference between the two scores. While the score based on self-reported socioenvironmental factors explained over 35% of the

Table 2
Results from the prediction models including both the objective (PES-O) and subjective (PES-S) poly-environmental scores.

Outcome	Pre-pandemic						Pandemic					
	PES-O		PES-S				PES-O		PES-S			
	β (SE)	*p*	β (SE)	*p*	*N*	R^2	β (SE)	*p*	β (SE)	*p*	*N*	R^2
Satisfaction with life	.07 (.15)	.65	3.05 (.15)	**<2x10^{-16}**	722	35.38%	.13 (.10)	.20	3.22 (.10)	**<2x10^{-16}**	1859	34.33%
Subjective Happiness	−.08 (.12)	.54	2.38 (.13)	**<2x10^{-16}**	722	33.04%	.03 (.09)	.78	2.59 (.10)	**<2x10^{-16}**	1858	28.36%
Quality of Life	.005 (.03)	.89	.59 (.03)	**<2x10^{-16}**	720	29.83%	.04 (.02)	.09	.64 (.02)	**<2x10^{-16}**	1882	26.38%
Flourishing	.11 (.17)	.49	2.89 (.17)	**<2x10^{-16}**	716	28.90%	.09 (.11)	.44	3.25 (.12)	**<2x10^{-16}**	1857	29.75%
Psychological Well-Being	−.08 (.21)	.70	1.98 (.21)	**<2x10^{-16}**	695	10.68%	.04 (.12)	.72	2.27 (.12)	**<2x10^{-16}**	1837	15.29%
Family Satisfaction	.01 (.03)	.60	.24 (.03)	**<2x10^{-16}**	719	10.81%	−.04 (.02)	.03	.27 (.02)	**<2x10^{-16}**	1878	9.65%
Financial Satisfaction	.01 (.03)	.65	.29 (.03)	**<2x10^{-16}**	718	13.21%	.01 (.02)	.47	.27 (.02)	**<2x10^{-16}**	1878	8.50%
Work Satisfaction	.004 (.03)	.90	.22 (.03)	**1.15x10^{-11}**	699	6.63%	.03 (.02)	.12	.20 (.02)	**<2x10^{-16}**	1861	4.61%
Health Satisfaction	.01 (.03)	.82	.25 (.03)	**1.99x10^{-13}**	721	7.00%	.05 (.02)	.02	.27 (.02)	**<2x10^{-16}**	1882	8.66%
Friendship Satisfaction	−.003 (.02)	.92	.31 (.02)	**<2x10^{-16}**	720	17.65%	.01 (.02)	.56	.32 (.02)	**<2x10^{-16}**	1877	14.97%
Self-Rated Health	.01 (.03)	.74	.20 (.03)	**2.61x10^{-13}**	722	6.36%	.03 (.02)	.04	.21 (.02)	**<2x10^{-16}**	1879	8.52%
Loneliness	.06 (.04)	.14	−.79 (.04)	**<2x10^{-16}**	719	36.66%	.03 (.03)	.36	−.86 (.03)	**<2x10^{-16}**	1876	29.74%
Depressive problems	.10 (.12)	.43	−1.79 (.12)	**<2x10^{-16}**	652	24.09%	−.02 (.09)	.80	−2.14 (.09)	**<2x10^{-16}**	1795	21.15%

Note. PES-O = objective poly-environmental score, PES-S = subjective poly-environmental score, β = beta, SE = standard error, p = p-value, N = sample size, R^2 = explained variance.

Table 3
Correlations with the well-being polygenic scores (PGS).

	NTR			UKB	
	r (CI)	*p*		r (CI)	*p*
Satisfaction with life	.09 (.02–.17)	.02	Happiness	.07 (.05–.08)	4.45 × 10^{-13}
PES-S	.11 (0–.21)	.05	PES-S	.07 (.05–.09)	3.75 × 10^{-14}
PES-O	−.07 (−.17 to .04)	.20			

Note. The polygenic scores were residualized for PCs + platform. PES-O = objective poly-environmental score, PES-S = subjective poly-environmental score, NTR= Netherlands Twin Register, UKB=UK Biobank, r = correlation coefficient, CI = 95% Confidence Interval, p = p-value.

variation in well-being scores, the score based on objective, physical environmental factors explained less than 1% of the variation. It is not entirely surprising that the PES-S explained such a large part of the variation in well-being: this score contained variables that are consistently associated with well-being in previous research, such as social support (Harandi, Taghinasab, & Nayeri, 2017; Jiameng & Huiyong, 2013), feelings of stress at home (Gerhardt et al., 2021), and negative and positive life events (Ballas & Dorling, 2007; Clark & Oswald, 2002). Since environmental factors have been found to account for 60–70% of individual differences in well-being (Bartels, 2015; Nes & Roysamb, 2015), the included socioenvironmental exposures were able to explain approximately half of the environmental variation in well-being. In our own work on the relation between the social environment and well-being and adolescents, we found that genetic factors were able to explain a significant part of these associations (van de Weijer et al., 2022). An interpretation of this finding is that these associations are partly explained by a genetic predisposition for appraising one's life positively or negatively.

As a follow-up, we repeated the same analysis in a sample of UKB participants, where we included the available socioenvironmental factors that most closely resembled the ones we included in the NTR PES-S. The socioenvironmental UKB PES explained approximately 12% of the variance in well-being, which is a substantial amount but considerably less than in NTR. This difference can be traced back to differences in the amount and content of the included variables in the two PESs, where variables with large contributions to the elastic net models in NTR, such as having a partner and stress at home/work, were not available in the UKB dataset (see Supplementary Tables 4–5). This difference in explained variance illustrates a limitation of our efforts to capture the well-being exposome using the environmental scores; the extent to which the exposome is actually captured is very dependent on the variables that have been measured in the cohort in question. In this respect, our analogy of the genome (and polygenic scores) to the exposome (and poly-environmental scores) is limited. As described by Primbs and colleagues, in the context of the genome, there is a known and finite set of genes that we can measure and study. However, the same is not true for psychological phenomena and the factors that influence them (Primbs et al., 2023). Since our environmental scores depend on which environmental factors are included from an unknown set of relevant exposures, cross-context/country comparisons are difficult, especially when the same variables are not available. However, given the relatively large explained variance in the NTR dataset, we can at least conclude that we were successful in capturing a large part of the well-being exposome in a Dutch context.

It is also not entirely surprising that the PES-O explains only a small part of the variation in well-being when used individually (in both the original and larger follow-up sample) and fails to predict well-being when combined with the PES-S in one model. In previous work, similar variables on the postal code level explained only 1.45% of the variance in well-being (van de Weijer et al., 2022). Existing literature examining associations between well-being and spatial measures offers somewhat mixed results. In a British study, Ballas & Tranmer examined the extent to which variation in happiness and well-being was explained by four different levels in a multilevel design: region, district, household, and individual (Ballas & Tranmer, 2012). They found that almost all of the variation in well-being and happiness was attributable to the individual level, and some to the household level. A very small part of the variation in well-being, and none of the variation in happiness, was attributable to district/region. They conclude that, in the British context, well-being varies between people but not places. Similarly, in a Dutch study comparing the effect of subjective and objective spatial characteristics on well-being, the effect of subjective spatial characteristics on well-being was much larger than the effect of objective spatial characteristics (Ettema & Schekkerman, 2016). In contrast, in an Irish sample, the inclusion of objective spatial indicators (such as mean annual precipitation and proximity to coast) in a model where life satisfaction was predicted using socioenvironmental indicators, while controlling for socio-economic and demographic characteristics of the individuals, led to a large increase in explained variance (Brereton, Clinch, & Ferreira, 2008). Moreover, researchers have found evidence for associations between well-being and different objective environmental indicators, such as air pollution (Orru, Orru, Maasikmets, Hendrikson, & Ainsaar, 2016), urban green space (Bertram & Rehdanz, 2015), and noise levels (Benita, Bansal, & Tunçer, 2019). Nevertheless, the general consensus seems to be that subjective environmental

indicators are better suited for explaining individual differences in well-being than objective ones. While these perceptions of the environment might be stronger predictors of physical environmental ones, the latter might also include more measurement error and thus be less reliable. For example, it has been shown that reliability of neighborhood condition measures is lower in rural than urban samples (Pruitt, Jeffe, Yan, & Schootman, 2012). Nevertheless, an interesting future endeavor would be to create a PES based on subjective, instead of objective, physical environmental indicators (i.e., the perceived safety instead of the actual crime rate).

We first examined potential gene-environment correlation in the NTR sample. The well-being PGS did not significantly predict either well-being or the PESs. It is likely that because of our limited sample size we did not have sufficient power to find an effect even if present. However, when performing similar analyses in the larger UKB sample (using happiness as a well-being measure instead of satisfaction with life), the well-being PGS (even though based on the somewhat less powered summary statistics due to the exclusion of UKB participants) was associated with both well-being itself and the well-being socio-environmental PES. This finding supports the notion of gene-environment correlation for well-being, where a person's exposure to the environment depends on their genetic predisposition for well-being. Given that our sample is of an adult sample, it is unlikely that we would identify passive rGE effects. The correlation between the well-being PGS and well-being PES is thus most likely to reflect either active or evocative rGE. In case of active rGE, this would mean that people's genetic disposition for well-being results in them seeking out certain types of (social) environments. For example, those with a higher genetic predisposition for well-being might seek out environments that stimulate their well-being, such as supporting relationships. In the case of evocative rGE, people's genetic predisposition for well-being would elicit certain types of environmental reactions, which in turn influences their well-being. For example, people with a high genetic predisposition for well-being might elicit positive social relations with others, which in turn would be beneficial for well-being. Our analyses indicate there potentially is gene-environment correlation but do not allow us to provide a conclusive statement on the nature of this correlation and which rGE scenario is most likely. Importantly, previous research has found that genome-wide association studies often also capture genetic variation associated with socioeconomic status (Marees et al., 2021). This can happen due to multiple reasons, for example due to the fact that socioeconomic status is geographically clustered (Abdellaoui et al., 2019), and therefore related to many environmental factors that influence mental health. We can thus not rule out that the genome-wide association study on which the well-being polygenic score is based potentially also captures socioeconomic status-associated genetic variation, and that this causes the association with the environmental scores.

Besides examining rGE, we also used the PESs for several other purposes. One of the research questions we were interested in was how well the well-being exposome in one context predicts well-being in other contexts. For the current project, we examined if the PESs would predict well-being to a similar extent during the COVID-19 pandemic and found that this was indeed the case. Other potential interesting applications could be to compare predictions across different ages, personality types, or other personal characteristics. We additionally used the PESs to examine overlap between different well-being constructs. By comparing if the PES for one well-being phenotype is as predictive for another well-being phenotype, we are provided with new information on the overlap/distinction between these phenotypes. Our results showed that the satisfaction with life PES predicted other well-being related phenotypes to varying degrees. For example, psychological well-being was predicted to a much lesser extent than satisfaction with life, indicating that the environmental exposures associated with satisfaction with life only partly overlap with those for psychological well-being. This is in line with previous research that found only partly overlapping unique environmental effects between subjective and psychological measures of well-being (Keyes, Myers, & Kendler, 2010). An interesting future application would be to use PESs for follow-up analyses for answering research questions about phenotypes such as resilience, e.g., why do some people still thrive despite low environmental opportunities for well-being, and why do others score relatively low on well-being in "high well-being" contexts?

Our results should be interpreted in the light of multiple important limitations. First, the Netherlands and the UK are Western, Educated, Industrialized, Rich, and Democratic (WEIRD) countries, and results likely do not translate well to non-WEIRD contexts. One of the ways in which the findings might not translate well is with respect to the physical environmental exposures. The Netherlands is both a WEIRD and small country, meaning that the average distance to most amenities is relatively short. For example, the *maximum* distance between any participant 4-numeric postal code and a primary school is 7 km (Supplementary Table 1). For larger and non-WEIRD countries, distance to most amenities might be longer and less homogeneous across the country. In that case, there would be more individual differences in physical environmental measures possibly indicating a more important role in explaining individual differences in well-being. In addition, it is important to mention that while our physical environmental measures do not depend on subjective participant evaluations, there are still different ways in which one can measure these factors. For example, in this study, livability was based on 45 environment characteristics (regarding population, social cohesion, public space, safety, level of resources, and housing). However, there are many different ways in which livability can be defined and measured (Ahmed, El-Halafawy, & Amin, 2019), and the results depend on the definition used. Therefore, the results of this study might not generalize over different operationalizations of the included physical environmental measures. It would be interesting to construct well-being PESs in different cultures/contexts and compare results across these contexts. For the subjective PES, we chose to rely on self-report since we were interested in people's own evaluations of their environments. However, a limitation is that both the dependent and independent variables were obtained from the same self-report survey, meaning that the analyses might suffer from common method bias. When there is common method bias, correlations between variables can be inflated because of different types of response bias (e.g., question order bias). In our case, this would result in the PESs explaining more variance than actually is the case. With respect to the PES-O, our postal code linkage suffers from two limitations. First, for the pre-pandemic sample, we used last known postal codes for linkage. It is possible that postal code was not up-to-date for all participants, in which case the linkage would have been incorrect. Second, the GECCO data was not always available for all the years in which we collected phenotype data. For example, the pandemic dataset was collected between 2020 and 2022, but the GECCO data was only available until 2020. In this case, we had to link the phenotype data to earlier years. While it is unlikely that there were large changes in the physical environmental data in such brief periods of time, it is possible that some error was introduced there. Lastly, while we speak of prediction models, associations between the environmental exposures and well-being should not be interpreted in a causal manner. The associations between the included environmental exposures and well-being could be causal in one or the other direction or bi-directional and are not necessarily direct.

5. Conclusions

In summary, this study provides the first attempt to combine different environmental exposures into well-being poly-environmental scores. We find that a subjectively assessed socioenvironmental PES explains around half of the environmental variation in well-being in a Dutch sample, but that a PES based on objective physical environmental indicators does not predict well-being (when combined in one model with the PES-S). The socioenvironmental PES predicted well-being

during the pandemic to a similar extent, and also predicted other well-being related phenotypes, albeit to varying extents. Additionally, we find that a PGS and PES for well-being are correlated in a UKB sample, suggesting the presence of gene-environment correlation. While our WEIRD sample has limited representability, this work shows the usefulness of using PESs for studying the well-being exposome. Future research could be conducted to examine the potential of subjective physical environmental indicators, and to study how these environmental scores vary across different cultures, contexts, and ages.

Funding

This work is supported by the European Research Council Consolidator Grant (ERC-2017-COG 771057 WELL-BEING PI Bartels), the Center for Medical Systems Biology (CSMB, NWO Genomics), and Biobanking and Biomolecular Resources Research Infrastructure (BBMRI –NL, 184.021.007 and 184.033.111).

Ethics statement

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

Author statement

The study was conceptualized by M.P. van de Weijer, D.H.M Pelt and M. Bartels. Data were collected by M.P. van de Weijer and F. Huider. The phenotypic and genetic data from the Netherlands Twin Register were cleaned by L. Ligthart and J.-J. Hottenga. The polygenic scores for the Netherland Twin Register were calculated by R. Pool. Analyses were performed by M.P. van de Weijer. The article was written by M.P. van de Weijer in collaboration with all co-authors.

Declaration of competing interest

The authors have no relevant financial or non-financial interests to disclose.

Acknowledgments

This research has been conducted using the UK Biobank Resource under Application Number 40310. We would like to thank the Netherlands Twin Register and UK Biobank participants for their contribution to this study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jenvp.2023.102208.

References

Abdellaoui, A., Hugh-Jones, D., Yengo, L., Kemper, K. E., Nivard, M. G., Veul, L., et al. (2019). Genetic correlates of social stratification in Great Britain. *Nature Human Behaviour, 3*(12). https://doi.org/10.1038/s41562-019-0757-5. Article 12.

Abdellaoui, A., Nivard, M. G., Hottenga, J.-J., Fedko, I., Verweij, K. J. H., Baselmans, B. M. L., et al. (2018). Predicting loneliness with polygenic scores of social, psychological and psychiatric traits. *Genes, Brain and Behavior, 17*(6). https://doi.org/10.1111/gbb.12472

Achenbach, T. M., Bernstein, A., & Dumenci, L. (2005). DSM-oriented scales and statistically based syndromes for ages 18 to 59: Linking taxonomic paradigms to facilitate multitaxonomic approaches. *Journal of Personality Assessment, 84*(1), 49–63. https://doi.org/10.1207/s15327752jpa8401_10

Ahmed, N. O., El-Halafawy, A. M., & Amin, A. M. (2019). A critical review of urban livability. *European Journal of Sustainable Development, 8*(1). https://doi.org/10.14207/ejsd.2019.v8n1p165. Article 1.

Akingbuwa, W. A., Hammerschlag, A. R., Jami, E. S., Allegrini, A. G., Karhunen, V., Sallis, H., et al. (2020). Genetic associations between childhood psychopathology and adult depression and associated traits in 42998 individuals: A meta-analysis. *JAMA Psychiatry, 77*(7), 715–728. https://doi.org/10.1001/jamapsychiatry.2020.0527

Ballas, D., & Dorling, D. (2007). Measuring the impact of major life events upon happiness. *International Journal of Epidemiology, 36*(6), 1244–1252. https://doi.org/10.1093/ije/dym182

Ballas, D., & Tranmer, M. (2012). Happy people or happy places? A multilevel modeling approach to the analysis of happiness and well-being. *International Regional Science Review, 35*(1), 70–102. https://doi.org/10.1177/0160017611403737

Bartels, M. (2015). Genetics of wellbeing and its components satisfaction with life, happiness, and quality of life: A review and meta-analysis of heritability studies. *Behavior Genetics, 45*(2), 137–156. https://doi.org/10.1007/s10519-015-9713-y

Baselmans, B. M. L., & Bartels, M. (2018). A genetic perspective on the relationship between eudaimonic –and hedonic well-being. *Scientific Reports, 8*(1). https://doi.org/10.1038/s41598-018-32638-1

Baselmans, B. M. L., Jansen, R., Ip, H. F., van Dongen, J., Abdellaoui, A., van de Weijer, M. P., et al. (2019). Multivariate genome-wide analyses of the well-being spectrum. *Nature Genetics, 51*(3), 445–451. https://doi.org/10.1038/s41588-018-0320-8

Baselmans, B. M. L., van de Weijer, M. P., Abdellaoui, A., Vink, J. M., Hottenga, J. J., Willemsen, G., et al. (2019). A genetic investigation of the well-being spectrum. *Behavior Genetics, 49*(3), 286–297. https://doi.org/10.1007/s10519-019-09951-0

Benita, F., Bansal, G., & Tunçer, B. (2019). Public spaces and happiness: Evidence from a large-scale field experiment. *Health & Place, 56*, 9–18. https://doi.org/10.1016/j.healthplace.2019.01.014

Bertram, C., & Rehdanz, K. (2015). The role of urban green space for human well-being. *Ecological Economics, 120*, 139–152. https://doi.org/10.1016/j.ecolecon.2015.10.013

Brereton, F., Clinch, J. P., & Ferreira, S. (2008). Happiness, geography and the environment. *Ecological Economics, 65*(2), 386–396. https://doi.org/10.1016/j.ecolecon.2007.07.008

Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K., et al. (2018). The UK Biobank resource with deep phenotyping and genomic data. *Nature, 562*(7726). https://doi.org/10.1038/s41586-018-0579-z. Article 7726.

Cantril, H. (1966). *The pattern of human concerns*. Rutgers University Press.

Clark, A. E., & Oswald, A. J. (2002). A simple statistical method for measuring how life events affect happiness. *International Journal of Epidemiology, 31*(6), 1139–1144. https://doi.org/10.1093/ije/31.6.1139

Deci, E. L., & Ryan, R. M. (2008). Hedonia, eudaimonia, and well-being: An introduction. *Journal of Happiness Studies, 9*(1), 1–11. https://doi.org/10.1007/s10902-006-9018-1

Deckers, K., Nooyens, A., van Boxtel, M., Verhey, F., Verschuren, M., & Köhler, S. (2019). Gender and educational differences in the association between lifestyle and cognitive decline over 10 Years: The doetinchem cohort study. *Journal of Alzheimer's Disease, 70*(s1), S31–S41. https://doi.org/10.3233/JAD-180492

Dick, D. M. (2011). Gene-environment interaction in psychological traits and disorders. *Annual Review of Clinical Psychology, 7*, 383–409. https://doi.org/10.1146/annurev-clinpsy-032210-104518

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment, 49*(1), 71–75. https://doi.org/10.1207/s15327752jpa4901_13

Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin, 125*(2), 276–302. https://doi.org/10.1037/0033-2909.125.2.276

Diener, E., Wirtz, D., Tov, W., Kim-Prieto, C., Choi, D. won, Oishi, S., et al. (2010). New well-being measures: Short scales to assess flourishing and positive and negative feelings. *Social Indicators Research*. https://doi.org/10.1007/s11205-009-9493-y

Eriksson, I., Undén, A. L., & Elofsson, S. (2001). Self-rated health. Comparisons between three different measures. Results from a population study. *International Journal of Epidemiology, 30*(2), 326–333. https://doi.org/10.1093/ije/30.2.326

Ettema, D., & Schekkerman, M. (2016). How do spatial characteristics influence well-being and mental health? Comparing the effect of objective and subjective characteristics at different spatial scales. *Travel Behaviour and Society, 5*, 56–67. https://doi.org/10.1016/j.tbs.2015.11.001

Gerhardt, C., Semmer, N. K., Sauter, S., Walker, A., de Wijn, N., Kälin, W., et al. (2021). How are social stressors at work related to well-being and health? A systematic review and meta-analysis. *BMC Public Health, 21*(1), 890. https://doi.org/10.1186/s12889-021-10894-7

Harandi, T. F., Taghinasab, M. M., & Nayeri, T. D. (2017). The correlation of social support with mental health: A meta-analysis. *Electronic Physician, 9*(9), 5212–5222. https://doi.org/10.19082/5212

He, Y., Lakhani, C. M., Manrai, A. K., & Patel, C. J. (2019). *Poly-exposure and poly-genomic scores implicate prominent roles of non-genetic and demographic factors in four common diseases in the UK*, Article 833632. https://doi.org/10.1101/833632. bioRxiv.

Houlden, V., Weich, S., Albuquerque, J. P. de, Jarvis, S., & Rees, K. (2018). The relationship between greenspace and the mental wellbeing of adults: A systematic review. *PLoS One, 13*(9), Article e0203000. https://doi.org/10.1371/journal.pone.0203000

Hughes, M. E., Waite, L. J., Hawkley, L. C., & Cacioppo, J. T. (2004). A short scale for measuring loneliness in large surveys: Results from two population-based studies. *Research on Aging, 26*(6), 655–672. https://doi.org/10.1177/0164027504268574

Jaffee, S. R., & Price, T. S. (2007). Gene–environment correlations: A review of the evidence and implications for prevention of mental illness. *Molecular Psychiatry, 12*(5). https://doi.org/10.1038/sj.mp.4001950. Article 5.

Jeon, E.-J., Kang, S.-H., Piao, Y.-H., Kim, S.-W., Kim, J.-J., Lee, B.-J., et al. (2022). Development of the korea-polyenvironmental risk score for psychosis. *Psychiatry Investigation, 19*(3), 197–206. https://doi.org/10.30773/pi.2021.0328

Jiameng, S., & Huiyong, F.a. N. (2013). A meta-analysis of the relationship between social support and subjective well-being. *Advances in Psychological Science, 21*(8), 1357. https://doi.org/10.3724/SP.J.1042.2013.01357

Keyes, C. L. M. (2002). The mental health continuum: From languishing to flourishing in life. *Journal of Health and Social Behavior, 43*(2), 207–222.

Keyes, C. L. M., Myers, J. M., & Kendler, K. S. (2010). The structure of the genetic and environmental influences on mental well-being. *American Journal of Public Health, 100*(12), 2379–2384. https://doi.org/10.2105/AJPH.2010.193615

Kuhn, M. (2022). *caret: Classification and regression training* (R package version 6.0-92) [Computer software] https://CRAN.R-project.org/package=caret.

Lewis, C. M., & Vassos, E. (2020). Polygenic risk scores: From research tools to clinical instruments. *Genome Medicine, 12*(1), 44. https://doi.org/10.1186/s13073-020-00742-5

Ligthart, L., van Beijsterveldt, C. E. M., Kevenaar, S. T., de Zeeuw, E., van Bergen, E., Bruins, S., et al. (2019). The Netherlands twin register: Longitudinal research based on twin and twin-family designs. *Twin Research and Human Genetics, 22*(6), 623–636. https://doi.org/10.1017/thg.2019.93

Lyubomirsky, S., & Lepper, H. S. (1999). A measure of subjective happiness: Preliminary reliability and construct validation. *Social Indicators Research, 46*(2), 137–155. https://doi.org/10.1023/A:1006824100041

Marees, A. T., Smit, D. J. A., Abdellaoui, A., Nivard, M. G., van den Brink, W., Denys, D., et al. (2021). Genetic correlates of socio-economic status influence the pattern of shared heritability across mental health traits. *Nature Human Behaviour, 5*(8). https://doi.org/10.1038/s41562-021-01053-4. Article 8.

Nes, R. B., & Roysamb, E. (2015). The heritability of subjective well-being: Review and meta-analysis. In *The genetics of psychological well-being: The role of heritability and genetics in positive psychology* (pp. 75–79). Oxford University Press.

Oliver, D., Radua, J., Reichenberg, A., Uher, R., & Fusar-Poli, P. (2019). Psychosis polyrisk score (PPS) for the detection of individuals at-risk and the prediction of their outcomes. *Frontiers in Psychiatry, 10*. https://www.frontiersin.org/articles/10.3389/fpsyt.2019.00174.

Olkin, I., & Finn, J. D. (1995). Correlations redux. *Psychological Bulletin, 118*, 155–164. https://doi.org/10.1037/0033-2909.118.1.155

Orru, K., Orru, H., Maasikmets, M., Hendrikson, R., & Ainsaar, M. (2016). Well-being and environmental quality: Does pollution affect life satisfaction? *Quality of Life Research, 25*(3), 699–705. https://doi.org/10.1007/s11136-015-1104-6

Padmanabhan, J. L., Shah, J. L., Tandon, N., & Keshavan, M. S. (2017). The "polyenviromic risk score": Aggregating environmental risk factors predicts conversion to psychosis in familial high-risk subjects. *Schizophrenia Research, 181*, 17–22. https://doi.org/10.1016/j.schres.2016.10.014

Pries, L.-K., Erzin, G., Rutte, B. P. F., van Os, J., & Guloksuz, S. (2021). Estimating aggregate environmental risk score in psychiatry: The exposome score for schizophrenia. *Frontiers in Psychiatry, 12*. https://www.frontiersin.org/articles/10.3389/fpsyt.2021.671334/full.

Pries, L.-K., Lage-Castellanos, A., Delespaul, P., Kenis, G., Luykx, J. J., Lin, B. D., et al. (2019). Estimating exposome score for schizophrenia using predictive modeling approach in two independent samples: The results from the EUGEI study. *Schizophrenia Bulletin, 45*(5), 960–965. https://doi.org/10.1093/schbul/sbz054

Primbs, M. A., Pennington, C. R., Lakens, D., Silan, M. A. A., Lieck, D. S. N., Forscher, P. S., et al. (2023). Are Small Effects the Indispensable Foundation for a Cumulative Psychological Science? A Reply to Götz et al. (2022. *Perspectives on Psychological Science, 18*(2), 508–512. https://doi.org/10.1177/17456916221100420

Pruitt, S. L., Jeffe, D. B., Yan, Y., & Schootman, M. (2012). Reliability of perceived neighbourhood conditions and the effects of measurement error on self-rated health across urban and rural neighbourhoods. *Journal of Epidemiology & Community Health, 66*(4), 342–351. https://doi.org/10.1136/jech.2009.103325

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., et al. (2007). Plink: A tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics, 81*(3), 559–575. https://doi.org/10.1086/519795

Robinson, M. R., Kleinman, A., Graff, M., Vinkhuyzen, A. A. E., Couper, D., Miller, M. B., et al. (2017). Genetic evidence of assortative mating in humans. *Nature Human Behaviour, 1*(1). https://doi.org/10.1038/s41562-016-0016. Article 1.

Ryff, C. D., & Singer, B. H. (2008). Know thyself and become what you are: A eudaimonic approach to psychological well-being. *Journal of Happiness Studies, 9*(1), 13–39. https://doi.org/10.1007/s10902-006-9019-0

Song, L., Meng, J., Liu, Q., Huo, T., Zhu, X., Li, Y., et al. (2019). Polygenic score of subjective well-being is associated with the brain morphology in superior temporal gyrus and insula. *Neuroscience, 414*, 210–218. https://doi.org/10.1016/j.neuroscience.2019.05.055

von Stumm, S., & d'Apice, K. (2022). From genome-wide to environment-wide: Capturing the environome. *Perspectives on Psychological Science, 17*(1), 30–40.

von Stumm, S., Kandaswamy, R., & Maxwell, J. (2023). Gene-environment interplay in early life cognitive development. *Intelligence, 98*, Article 101748. https://doi.org/10.1016/j.intell.2023.101748

Tan, J. J. X., Kraus, M. W., Carpenter, N. C., & Adler, N. E. (2020). The association between objective and subjective socioeconomic status and subjective well-being: A meta-analytic review. *Psychological Bulletin, 146*(11), 970–1020. https://doi.org/10.1037/bul0000258

Timmermans, E. J., Lakerveld, J., Beulens, J. W. J., Boomsma, D. I., Kramer, S. E., Oosterman, M., et al. (2018). Cohort profile: The geoscience and health cohort Consortium (GECCO) in The Netherlands. *BMJ Open, 8*(6), Article e021597. https://doi.org/10.1136/bmjopen-2018-021597

Vassos, E., Sham, P., Kempton, M., Trotta, A., Stilo, S. A., Gayer-Anderson, C., et al. (2020). The Maudsley environmental risk score for psychosis. *Psychological Medicine, 50*(13), 2213–2220. https://doi.org/10.1017/S0033291719002319

Vilhjálmsson, B. J., Yang, J., Finucane, H. K., Gusev, A., Lindström, S., Ripke, S., et al. (2015). Modeling linkage disequilibrium increases accuracy of polygenic risk scores. *The American Journal of Human Genetics, 97*(4), 576–592. https://doi.org/10.1016/j.ajhg.2015.09.001

van de Weijer, M. P., Baselmans, B. M. L., Hottenga, J.-J., Dolan, C. V., Willemsen, G., & Bartels, M. (2022). Expanding the environmental scope: An environment-wide association study for mental well-being. *Journal of Exposure Science and Environmental Epidemiology, 32*(2), 195–204. https://doi.org/10.1038/s41370-021-00346-0

van de Weijer, M. P., Baselmans, B., van der Deijl, W., & Bartels, M. (2018). A growing sense of well-being: A literature review on the complex framework well-being. *PsyArXiv*. https://doi.org/10.31234/osf.io/3rmx9

van de Weijer, M. P., Pelt, D. H. M., de Vries, L. P., Huider, F., van der Zee, M. D., Helmer, Q., et al. (2022). Genetic and environmental influences on quality of life: The COVID-19 pandemic as a natural experiment. *Genes, Brain and Behavior, 21*(8), Article e12796. https://doi.org/10.1111/gbb.12796

van de Weijer, M. P., Pelt, D. H. M., van Beijsterveldt, C. E. M., Willemsen, G., & Bartels, M. (2022). Genetic factors explain a significant part of associations between adolescent well-being and the social environment. *European Child & Adolescent Psychiatry, 31*(10), 1611–1622. https://doi.org/10.1007/s00787-021-01798-3

Wild, C. P. (2012). The exposome: From concept to utility. *International Journal of Epidemiology, 41*(1), 24–32. https://doi.org/10.1093/ije/dyr236