# Association between real-world experiential diversity and positive affect relates to hippocampal-striatal functional connectivity

**Aaron S. Heller**[1,*], **Tracey C. Shi**[2,4], **C.E. Chiemeka Ezie**[1,3], **Travis R. Reneau**[1], **Lara M. Baez**[1], **Conor J. Gibbons**[1], **Catherine A. Hartley**[4,5,*]

[1.]University of Miami, Department of Psychology

[2.]Columbia University, College of Physicians and Surgeons

[3.]New York University, School of Medicine

[4.]New York University, Department of Psychology

[5.]New York University Center for Neural Science and Langone Health Neuroscience Institute

## Abstract

Experiential diversity promotes well-being in animal models. Here, using geolocation tracking, experience sampling, and neuroimaging, we found that daily variability in physical location was associated with increased positive affect in humans. This effect was stronger for individuals who exhibited greater functional coupling of the hippocampus and striatum. These results link diversity in real-world daily experiences to fluctuations in positive affect and identify a hippocampal-striatal circuit associated with this bidirectional relation.

Experiential diversity confers substantial benefits for cognitive and affective function[1]. Animals able to roam freely within "enriched" environments that offer diverse and variable experiences exhibit greater affective well-being relative to those reared in standard laboratory environments[1]. In rodents, exposure to enriched environments facilitates play[2], social affiliative behavior[3], optimistic evaluative biases[4], and stress resilience[5]. Each of these behavioral phenotypes in rodent models is characteristic of positive affective states in humans[6], underscoring the benefits of experiential diversity for psychological well-being.

The behavioral effects of environmental enrichment are paralleled by pronounced changes in brain structure and function across a number of regions[7]. In particular, plastic synaptic changes in the hippocampus and striatum of animals reared in complex environments[7] highlight these brain regions as potential mediators of the beneficial consequences of experiential diversity.

Increased exposure to novelty may be a central mechanism through which enriched environments promote positive affect. While the increased physical activity and social interaction afforded by enriched environments facilitates positive behavioral outcomes[7], environmental novelty alone is sufficient to produce many benefits of enrichment[8]. Behavioral and neural data suggest that exposure to novelty has subjectively rewarding and reinforcing properties[9,10]. For example, novel stimuli can effectively reinforce action[10], elicit activity in reward-sensitive dopaminergic neurons[9], and promote a preference for the contexts in which they are encountered[11]. The same hippocampal-striatal circuit implicated in the effects of environmental enrichment is also centrally involved in novelty detection as well as the encoding of reward and positive affective states[9,10,12]. Collectively, these empirical findings corroborate a longstanding theoretical proposal that novel experiences may function as intrinsic rewards and generate positive affect[13]. However to date, no study has examined whether variation in real-world experiential diversity is associated with positive affective states in humans.

Because humans live in a complex and dynamic environment, daily living ostensibly affords limitless opportunity for experiential diversity. However, exposure to novel experiences, and their rewarding effects, depends on the degree to which individuals exploit this opportunity by exploring their environment. Roaming entropy (RE), which indexes the variability in an individual's physical location over the course of a day, is a quantitative location-derived metric of experiential diversity[14]. In rodents, increased levels of RE reproduce the key neural and cognitive effects of environmental enrichment[14]. In the present study, we examined whether daily RE might serve as an ecologically valid measure of experiential diversity in humans, and accordingly, whether day-to-day variation in RE might relate to fluctuations in positive affect.

We conducted geolocation tracking on a sample of 132 human participants (90 female, aged 18–31, mean = 23, SD = 3.4) recruited in three cohorts in New York and Miami. Participants' locations were tracked continuously for 3–4 months (mean 104.6 days, SD = 22.1). During geolocation tracking, participants were prompted to provide assessments of positive (PA) and negative (NA) affect via a questionnaire sent to their mobile phones. Affect assessments were delivered at random times every other day. Ten participants responding to fewer than 20 ecological momentary assessments (EMA) were excluded (Supplementary Fig. 1), leaving 122 (84 female, mean age = 22.75, SD = 3.36) participants.

Each day with adequate geolocation data (Supplementary Figures 2 & 3) was downsampled to one-minute temporal resolution and converted into a measure of daily RE, previously used to quantify rodent exploratory behavior[14]:

$$RE = \sum_{j=1}^{n} (p_{ij} \times (log_2 p_{ij}))/\log(n)$$

In this equation, $p_{ij}$ is the within-day historical probability that location $j$ was visited by participant $i$ (i.e., the proportion of the day spent in each unique location) and n is the total number of unique locations in the environment (at four decimal degrees of GPS resolution). Thus, RE is higher for days in which one visits a greater number of locations and exhibits greater uniformity in the distribution of her time across visited locations. A minimum RE value would be achieved by spending all day in a single location, whereas maximum entropy would, theoretically, be achieved by spending an equal proportion of the day in each unique location in an environment (Figure 1).

We examined whether RE was associated with day-to-day variation in self-reported PA and NA (Supplementary Figs. 4 & 5), with the following potentially confounding factors included in the model: distance travelled for that day, city (Miami or New York), day of the week, time of day the EMA survey was completed, and temperature and precipitation at modal location for that day. Across participants, self-reported positive affect was higher on days when RE was greater (B = 0.79, t(3834) = 2.464, p = 0.014; Supplementary Table 1). The RE metric corresponds closely to the number of unique locations visited in a day (mean within-subjects correlation: 0.85; Supplementary Note 1 & Supplementary Figure 6) and both RE and this simpler metric were significantly predictive of PA. Of the other variables in the model, only day of the week was significantly associated with PA (Supplementary Table 1), with PA lower early to mid-week relative to the weekend (Extended Data Fig. 1). There was no relation between RE and negative affect (B = −0.19, t(3835) = −0.66, p = 0.509; Supplementary Table 2).

This finding suggests that increased diversity in one's daily spatial environment is associated with greater positive affect. We next sought to test whether this relation might reflect increased exposure to novel and diverse experiences on higher RE days. Our limited geolocation tracking period does not allow us to determine whether an individual has never previously visited a specific location and it is truly novel. However, we derived a putative measure of location novelty by calculating the number of unique locations an individual visited each day that had not previously been visited during the tracking period. Number of novel locations correlated positively with RE (B = 0.0005, t(9697) = 141.03, p < .001; mean within-subjects correlation: 0.79), and was also significantly associated with PA (B = 1.60, t(3473) = 4.676, p < .001). Moreover, when both RE and number of novel locations were included in the model, the latter remained associated with PA (B = 2.44, t(3472) = 4.40, p < .001), whereas RE was no longer significantly associated with PA (B = −1.06, t(3472) = −1.94, p = 0.053). These analyses suggest that increases in RE may indeed be associated with heightened exposure to novelty, and that daily novelty exposure is associated with positive affect. However, given our limited temporal window of observation, this result should be corroborated in a study in which absolute location novelty can be accurately determined.

We next tested whether RE was associated with diversity in specific dimensions of the experiences that an individual has in a given day. We extracted US Census Bureau data capturing sociodemographic features of each location visited by a participant (e.g., population density, median age, total businesses, educational attainment, unemployment rate, race, gender; Supplementary Appendix 1). We used principal component analysis (PCA) to reduce the dimensionality of the data (Supplementary Table 3). We discretized the values for the first three components into high and low values and recoded each location as one of eight possible locations in this $2 \times 2 \times 2$ (PC1 (low/high) x PC2 (low/high) x PC3 (low/high)) sociodemographic feature space. Consistent with our hypothesis that RE might be associated with greater diversity in one's daily experiences, RE was significantly correlated with entropy over this sociodemographic feature space (B = 5.55, t(9697) = 78.56, p < .001). As large spatial extents of both New York and Miami contain a limited range of values in this sociodemographic space (Figure 2), this correlation is not a trivial consequence of each visited GPS location having a unique location in the PCA-derived sociodemographic feature space. Greater entropy over sociodemographic environmental features was also associated with increased PA (B = 1.14, t(3834) = 3.48, p = .0005; Supplementary Table 3). Moreover, when both RE and sociodemographic feature entropy were included in the model, the latter remained significantly associated with PA (B = 0.94, t(3833) = 2.21, p = 0.027), whereas RE was no longer associated with PA (B = 0.35, t(3833) = 0.870, p = 0.384). Collectively, these analyses suggest that the RE measure of location variability is associated with exposure to diverse sociodemographic features of the environment, and that increases in such experiential diversity, above and beyond location variability, are associated with day-to-day variation in positive affect.

While it is possible that the novel and diverse experiences associated with high roaming entropy days facilitate positive affect, another possibility is that individuals are more motivated to explore their environments during days in which they feel better. To examine the potential directionality of our observed effect, we conducted analyses of the temporal relationship between roaming entropy and positive affect. We first tested whether the previous day's RE influenced PA on a given day. When including both the previous ($RE_{t-1}$) and current ($RE_t$) day's RE in a model predicting the current day's PA ($PA_t$), only $RE_t$ significantly predicted $PA_t$ (B = 0.87, t(3367) = 2.498, p = 0.0125), whereas $RE_{t-1}$ did not (B = 0.31, t(3367) = 0.957, p = 0.339). We also examined whether positive affect on a given day ($PA_t$) predicted roaming entropy the following day ($RE_{t+1}$). There was no significant relationship between $PA_t$ and $RE_{t+1}$: B = 0.0008, t(3396) = 1.029, p = 0.3037, in a model which also included $RE_t$.

Next, we repeated these temporal analyses with the measures of daily novelty exposure and sociodemographic experiential diversity, each of which was more strongly related with PA than RE itself. Both the concurrent (B = 1.66, t(3031) = 4.476, p < .001) and the prior day's (B = 0.73 t(3031) = 2.121, p = 0.034) number of novel locations and the concurrent (B = 1.12, t(3367) = 3.219, p = 0.0013) and prior day's (B = 0.61, t(3367) = 2.085, p = 0.037) sociodemographic entropy significantly predicted a given day's PA. We also observed a significant effect in the opposite temporal direction, such that a given day's positive affect significantly predicted both the following day's number of novel locations (B = 0.003, t(3079) = 3.477, p = 0.0005) and entropy over sociodemographic features

of the environment (B = 0.002, t(3396) = 3.066, p = 0.002 ), with the concurrent day's value for each variable included in the models. These analyses suggest a bidirectional and temporally extended interaction in which novelty and experiential diversity are associated with greater subsequent positive affect, which, in turn, is associated with more novel and diverse experiences on the subsequent day. This is consistent with theoretical proposals that behaviors that promote positive emotions can engender self-reinforcing "upward spirals"[15], fostering subsequent repetition of those same behaviors. Nonetheless, future studies that directly manipulate positive affect, or exposure to novel and diverse experiences, will be critical in order to provide causal evidence for such bidirectional effects.

To better understand the neural mechanisms associated with the relation between roaming entropy and positive affect, we performed a resting-state fMRI scan in approximately half of the participants (n = 58) at the end of the geolocation tracking period. We used the random-effects estimates from the multilevel model for each individual as an index of the strength of coupling between RE and PA. We defined an anatomical seed region in the ventral striatum, an area widely associated with the subjective experience of positive affect as well as the processing of reward and novelty[12]. We conducted a whole-brain analysis, corrected for multiple comparisons, to identify regions of the brain in which resting-state functional connectivity with the ventral striatum was associated with a stronger relation between RE and PA (controlling for mean PA, mean RE, mean framewise displacement, and cohort; Figure 3). This analysis identified a cluster within the right posterior hippocampus ([x = 34, y = −24, z = −8], B = 0.08, t(52) = 4.51, p < 0.001; Supplementary Table 5, Supplementary Fig. 7). This finding suggests that functional integration of the striatum and the hippocampus, regions centrally implicated in the processing of reward and environmental novelty, is associated with the strength of the relation between experiential diversity and positive affect.

Our findings demonstrate that objectively measured increases in the diversity of one's physical environment are associated with increases in positive affect, and link connectivity within a hippocampal-striatal circuit to the strength of this effect. The hippocampus is centrally implicated in both the representation of spatial location as well as the detection of novelty[9,10]. Exposure to novel spatial environments elicits hippocampal activity that drives dopamine release in the nucleus accumbens via glutamatergic projections[9,10]. These projections play a necessary role in the formation of associations between spatial locations and primary rewards[16] and are also proposed to underpin the rewarding properties of novelty[9,10]. Consistent with this proposal, individuals with greater hippocampal-ventral striatal tract strength also exhibit higher levels of trait novelty-seeking[17]. Collectively, this literature suggests that hippocampus-nucleus accumbens functional connectivity might influence the degree to which novelty and diversity in one's spatial environment are experienced as subjectively rewarding, which in turn may promote subsequent exploratory or novelty-seeking behavior. While these data do not afford a direct test of this hypothesis, future studies might examine whether changes in the magnitude of hippocampal-striatal functional coupling mediate the strength of the relation between roaming entropy and positive affective states.

By harnessing data that objectively quantifies defining features of locations, we found that exposure to diverse social and demographic features of visited locations was also associated with positive affect. This same approach could be leveraged to examine whether experiential diversity across other environmental feature dimensions (e.g., developed versus natural environments[18]) relates to fluctuations in affective state. While our analysis focused on location-related diversity in daily experience, diverse experiences that occur independent of any change in physical location (e.g., social interactions, professional or leisure activities) have also been related to positive affect[19]. Future studies seeking to link real-world daily activities to emotional well-being might assay such behaviors, as well as examine whether the relation between experiential diversity and positive affect depends on the degree to which one's movements or activities are volitional or subject to external constraints[20].

Evidence from rodent models indicates that diversity of daily experience alters the function of neural circuits and increases behavioral indices of affective well-being. Here we demonstrated that objectively measured increases in the daily diversity of one's physical location were associated with corresponding increases in positive affect, an effect that was related to greater functional coupling of a hippocampal-striatal circuit. Collectively, these findings translate a broad literature on the beneficial consequences of environmental enrichment across species, demonstrating a relation between real-world exposure to novel and diverse experiences and increases in positive affect in humans, and linking this bidirectional association to the engagement of hippocampal-striatal neural circuitry.

## Methods:

### Participants

132 young adult participants (90 female, mean age: 23, sd = 3.4, range=18–31) were recruited in three cohorts from the New York City and Miami areas via fliers and posted advertisements. Participants were required to have a smartphone that could receive text messages and that met system requirements to run the study application. A subset of participants recruited in New York City (the first cohort) were not screened for magnetic resonance imaging (MRI) eligibility and only completed the geolocation and affect sampling assessments (n = 31). An additional group of MRI-eligible (right-handed, no MRI contraindications) participants were recruited both in New York City (second New York City cohort; n = 60) and Miami (third cohort; n = 41). All participants provided informed consent and were paid for their participation in the study. There was no randomization of subjects to experimental conditions, and thus, data collection and analysis did not require blinding. This study was approved by the Weill Cornell Medicine and University of Miami Institutional Review Boards.

### Procedure

During an initial session in the lab prior to the start of geolocation tracking, participants completed a series of baseline questionnaires including measures of depression (Beck Depression Inventory [BDI])[21] and trait anxiety (State-Trait Anxiety Instrument [STAI-T])[22]. A geolocation application (Moves; ProtoGeo, Helsinki, Finland) was installed on each participant's phone to passively acquire geolocation data whenever the phone's

accelerometer sensed motion. Geolocation data was synced to a participant-specific account on the application's online servers. Following the initial in-lab session, participants' geolocations were tracked continuously for 3–4 months (mean 104.6 days, sd = 22.1) via the Moves application.

During this period, participants were prompted to provide assessments of their mood via a questionnaire sent by SMS to their mobile phones (MRI-eligible cohort) or via an app (first cohort only) installed on their phones. Participants rated their current mood in response to the following 10 adjectives taken from the Positive and Negative Affect Scale (PANAS-X)[23]: Positive Affect (PA): happy, excited, strong, relaxed, attentive; Negative Affect (NA): nervous, jittery, upset, irritable, sluggish. Participants were also asked to rate their previous night's sleep on a scale of "slept poorly" to "slept well." SMS assessments of mood and sleep were delivered at random times every other day during daytime hours. Due to technological limitations, the first cohort used a mood recording application (T2 Mood Tracker; DHA Connected Health, Joint Base Lewis-McChord, WA, USA) instead of SMS assessment, and were queried to report their mood every other day at the same time of day. PA and NA were operationalized as the mean of the adjectives for that assessment. Ten participants who responded to fewer than 20 SMS EMA assessments during the 3–4 month tracking period were excluded from analyses, leaving 122 participants (Supplementary Fig. 1; mean age = 22.75 years, SD = 3.36, 84 females). Of the included participants, the number of completed affect assessments ranged from 22 to 78 (mean=42.8; SD=9.4), minimum (22) and maximum (78). PA and NA exhibited an inverse relationship across subjects (B = −0.60, t(5080) = −28.39, p < .001), with a mean correlation of −0.54.

Following the mobile tracking period, participants completed a second session in the lab. During this session, participants again completed the BDI and STAI-T. The depression and trait anxiety measures obtained prior to (mean (SD) BDI: 6.99 (6.38); STAI: 40.86 (11.25)) and following the 3–4 month sampling period (mean (SD) BDI: 5.78 (5.31); STAI: 38.83 (11.06)) did not differ (BDI: B = 0.46, t(110) = 0.50, p = 0.62; STAI: B = −1.11, t(110) = −0.48, p = 0.63, in a model that included cohort number to control for potential effects of seasonality), and mean RE (see formula below) was not correlated with self-reported depression (r = −0.02) or trait-anxiety (r = 0.01).

At both sites, the subset of participants in the MRI-eligible cohort who were willing to undergo an fMRI scan completed a third session that involved scanning no later than 4 weeks after the second lab session. During this session, a 15-minute resting-state scan (2.0s TR, Flip Angle = 75, TE: 29 ms, 3.4 × 3.4 × 3.4 mm, 38 slices) as well as T1- (1 mm$^3$) and T2-weighted anatomical scans were collected on each participant. 38 participants were scanned at Weill Cornell on a 3T Siemens Trio and 25 participants were scanned at the University of Miami on a 3T GE750. Two of the scanned participants were excluded from analysis for failure to respond to a sufficient number of SMS EMA assessments (as described above), and three participants were excluded from analyses due to excessive motion (see fMRI data preprocessing below), leaving a total of 58 participants included in our reported fMRI analyses.

## Geolocation data processing

Prior to performing any analyses, we sought to determine the amount of geolocation data required on each day to extract a stable RE estimate. Missing data is unavoidable when collecting geolocation data from cellular phones (i.e., due to the device being shut off). Thus, we performed several analyses to identify an acceptable threshold for the amount of missing data that would enable reliable estimation of within-day roaming entropy (RE), while excluding the fewest number of days so that the day-to-day dynamics of roaming entropy could be examined.

The first analysis identified the robustness of RE estimates (see equation below) to various levels of missing data by randomly removing a percentage of each day's data from days for which data was at least 99% complete. RE was recalculated from this subsample of data and compared to the full day's RE. A threshold of 90% complete geolocation data within a day closely approximates 99% complete data (Supplementary Fig. 2).

The second analysis examined how many days would be excluded as a function of the threshold of data required for inclusion. At the within-day thresholds of 30%, 20%, 10%, 1%, 0% allowable missing data the mean percent of days removed were: 16.1%, 18.9%, 23.2%, 34.4%, and 77.2%, respectively. Selecting a threshold of <= 10% missing data resulted in an average of 80.2 useable days per subject (Supplementary Fig. 3).

As a result of these analyses, we identified a threshold of a maximum of 10% of missing data within-day as being sufficient to retain a sufficient number of days, while ensuring that calculation of a single day's RE would be accurate.

Geolocation data were preprocessed such that latitude and longitude coordinates were rounded to the fourth decimal (a spatial resolution of approximately 4–11 meters between each unique location). For each day, we converted continuous geolocation data, downsampled to one-minute temporal resolution, into a measure of daily roaming entropy (RE) that has previously been used in studies quantifying rodent exploratory behavior[14,24], interpolating over timepoints when data were missing:

$$RE = \sum_{j=1}^{n} (p_{ij} \times (log_2 p_{ij}))/\log(n)$$

In this equation, $p_{ij}$ is the within-day historical probability that location $j$ was visited by participant $i$, quantified as the proportion of the day spent in each unique location (number of minutes in that location divided by the 1440 minutes in a day) and n is the total number of unique locations in the environment (calculated to be 648,000,000 unique locations on earth at four decimal degrees of GPS resolution).

Several variables that might impact RE or affect were also acquired and used as covariates in analyses. These included: time and day of week that mood surveys were completed, amount of precipitation and mean temperature for the modal location of the participant for that day, total distance traveled for the day, and cohort.

To extract weather information, weather data was collected using the API at http://weatherunderground.com. API requests for mean temperature and total precipitation data were made using participants' modal longitude/latitude for each day.

To calculate total distance traveled per day, the sum of distances between successive points visited in a given day was calculated. Distance was calculated using the rdist.earth function in the R package 'fields,' which uses the haversine formula,

$$d = 2r \arcsin\left(\sqrt{\sin^2\left(\frac{\phi_2 - \phi_1}{2}\right) + \cos(\phi_1)\cos(\phi_2)\sin^2\left(\frac{\lambda_2 - \lambda_1}{2}\right)}\right)$$

where $\Phi$ represents latitude, $\lambda$ longitude, r the radius of the Earth, and d the distance between the two points.

## Location data analysis

Multilevel models were run in R (3.5.0) using the nlme package[25] to examine whether RE predicted daily positive or negative affect. Subject and RE were treated as random factors. T-statistics and p-values were extracted from these models using lmerTest. We elected to run a model that included RE, as well as a number of additional factors that we hypothesized might relate to PA, including the time of day that mood assessment was completed, amount of precipitation for modal location of that day, mean temperature for modal location of that day, total distance traveled for that day, day of week (treated as a factor), and cohort (treated as a factor). Results reported in the manuscript include this full set of covariates that we hypothesized, a priori, might account for variance in PA. However, we also performed model fit comparisons using the 'dredge' function as part of the "MUMIn" library in R which ranks models via the Akaike Information Criterion (AIC)[26]. The best-fitting model only included RE, cohort and day of the week. We repeated all analyses using this best fitting model and the relation between RE and PA held under this reduced model (B = 0.63, t(3993) = 2.07, p = 0.038). Random-effect coefficients from the full model, reflecting the strength of the RE-PA association, were extracted for each subject, and used as individual difference metrics to predict rsFMRI connectivity.

## Determining whether analytic assumptions drive the observed RE – PA relationship

In our geolocation analysis, we interpolated over gaps in location data caused by phones being off or loss of GPS signal (e.g., due to underground subway commutes). To test whether interpolation might artificially increase RE and account for effects, we recalculated RE assuming that the participant remained in the last observed location throughout the temporal window of the data gap, instead of interpolating, and reran our analyses testing for relationships between daily RE and PA. An identical linear mixed effects model also demonstrated a significant association between non-interpolated RE and PA (B = 0.78, t(3834) = 2.457, p = 0.014), suggesting that this result is robust to assumptions about behavior when the phone is off.

**Extraction of number of novel locations visited**—For analyses examining whether the RE-PA effect could be accounted for by an estimate of the degree to which a participant

encountered "novel" locations on that day, we extracted the number of novel locations visited in a day (defined as number of unique locations not previously visited during the observation period). An important caveat regarding this metric is that we are not able to determine in any absolute sense whether an individual has never previously visited a location, and thus whether the location is truly novel for a participant. Instead, we can only make the more restricted inference of whether the participant has previously visited this location during the 3–4 month tracking period. In order to avoid inflation of this novelty estimate early in the observation period and allow for adequate observation of visited locations prior to the designation of novel locations, we held out data from the first 10 days before estimating novelty.

**Estimation of entropy over sociodemographic features of the environment—**
We suggest that a greater degree of daily 'experiential diversity' on high RE days might contribute to the association between RE and PA. To test this hypothesis, we attempted to estimate the degree of diversity in the social and demographic features of the locations visited by participants on a given day.

The first step in this analysis was to map each GPS coordinate that appeared in our dataset to its corresponding Federal Information Processing Standard (FIPS) code. A FIPS code uniquely identifies locations at the spatial resolution of a "Block Group", the smallest geographical unit for which the U.S. Census Bureau publishes data. The areas encompassed by Block Groups differ in size as they are determined by the approximate population size of roughly 600 to 3,000 people each.

Block Groups are used by US government agencies and other geographic information systems (GIS) to link geographic and demographic information to specific physical locations. We downloaded 53 variables capturing demographic and socioeconomic features of a given block group (e.g., population density, median age, total businesses, educational attainment, unemployment population, race, gender - see Supplementary Appendix 1 for a complete list of data sources) from the U.S. Census Bureau.

We performed a Principal Component Analysis (PCA) to reduce the dimensionality of the 53 sociodemographic variables. PCA derives a smaller set of artificial variables ("components") that explain a maximum of variance in all of the variables. All variables were mean-centered and scaled by standard deviation. The "principal" function from the R package psych version 1.8.12[27] was utilized to perform a Varimax rotated PCA on the 53 sociodemographic variables. Varimax rotation was utilized to maximize orthogonality among the components. Examination of the scree plot and eigenvalues showed that the 53 Census variables optimally clustered into 12 components that explained 80% of the variance in the data. Component weights indicate how strongly a sociodemographic variable is related to its component, with 1 reflecting the strongest positive associations and −1 indicating a strong inverse relationship (component weights for all 53 variables can be found in Supplementary Table 3). Individual component values for each block group were calculated using the regression method.

In order to better differentiate our measure of location in GPS coordinate space from a measure of location in sociodemographic feature space, we sought to define a sociodemographic feature space at a relatively lower level of spatial resolution. We took the first 3 components in our PCA analysis that explained the largest amount of variance in the sociodemographic data and discretized this 3-dimensional space into "high" and "low" values for each component, based on a median split. This results in each block group being assigned one of eight possible positions in this 3-dimensional sociodemographic feature space (PC1 (2 levels) x PC2 (2 levels) x PC3 (2 levels)).

We converted each subject's minute-by-minute GPS coordinate timeseries into a timeseries consisting of minute-by-minute values in this 3-dimensional sociodemographic feature space for the corresponding Block Group. We then calculated daily entropy over this feature space. This entropy measure reflects day-to-day variability in exposure to sociodemographic features of the environment, with higher entropy achieved on days when an individual is exposed to greater sociodemographic diversity.

## fMRI data preprocessing

Resting state data were preprocessed using AFNI (version AFNI_17.2.17). Resting state functional MRI (rsFMRI) data were despiked, slice-time corrected, coregistered to the subject's T1 image, spatially registered using linear and nonlinear registration (3dQwarp) to the MNI152 template, and smoothed with a $6mm^3$ isotropic kernel. Normalized T1 data were segmented into grey matter, white matter, and the ventricles. We removed variance associated with motion, the derivative of motion, and the timeseries of BOLD activity in white matter and ventricles[28,29]. TRs with motion above 2mm were censored and the resulting data were bandpass filtered at 0.01 – 0.1 Hz. Participants with >3mm head motion in any direction were excluded from analyses (n = 3). The resulting residuals from the remaining 58 participants were used in the connectivity analyses detailed below.

Because of ongoing debate within the resting state fMRI literature on preprocessing best practices – particularly work indicating that the effects of motion on resting state BOLD are non-linear and spatially variable[30–32] – we performed a second preprocessing approach similar to the one used by multiple groups[32]. Using this approach, we first identified frames that contained framewise displacement > 0.2mm. Those volumes were marked as outliers, however we also marked one volume before and two volumes after any volumes with FD > 0.2. In addition, uncensored segments of data lasting fewer than five contiguous volumes were flagged as outliers. It has been suggested that this approach more completely removes volumes in which movement occurred as well as the potential aftereffects of movement on the BOLD signal.

We then used the same nuisance regressors (six demeaned motion parameters, their derivative, their square, the square of the derivative, the averaged ventricular signal, its derivative and square, as well as the averaged white matter signal, its derivative and square), with a band pass of 0.01 to 0.10 Hz. With the residual fMRI timeseries, we then implemented the Lomb-Scargle Periodogram censoring interpolation method[31] to perform signal recovery and interpolation. We extracted VS-hippocampal connectivity from these

preprocessed data and tested whether we obtained the same result using this alternative preprocessing approach (Supplementary Note 2).

We conducted additional analyses to determine whether the seed region in the ventral striatum and the hippocampus cluster fell within traditional resting state networks[33] (Supplementary Note 3).

### fMRI data analysis

For each participant, the mean timeseries from voxels in the nucleus accumbens (NAcc), anatomically defined using the Harvard-Oxford Atlas[34], was extracted from the preprocessed rsFMRI data. This timeseries was correlated with the rest of the brain, and an r-z transformation was performed to complete parametric statistical tests. The random effect coefficients from the multilevel model described above (along with average framewise displacement, scanner site (Miami or New York City), mean RE, and mean PA as control variables) were then included as predictors in a robust regression model to predict NAcc-whole brain connectivity. These random effects maps were corrected for multiple comparisons using FSL's FLAME with a threshold of $z > 2.7$, corrected at $p < .05$.

## Statistics

Multilevel models were run in R (3.5.0) using the nlme package[25] to examine whether our geolocation-derived measures (roaming entropy, number of novel locations, entropy over sociodemographic [US Census derived] space, number of unique locations) predicted daily positive or negative affect. Data were assumed to be normally distributed, but this was not formally tested. To view the distribution, scatterplots illustrating participant-level associations between roaming entropy (z-scored) and positive and negative affect are presented in Supplementary Figs. 4 & 5. Imaging analyses used voxelwise multiple regression using FSL. No statistical methods were used to predetermine sample sizes, but our sample sizes were similar to those reported in previous publications[18,35].

## Extended Data
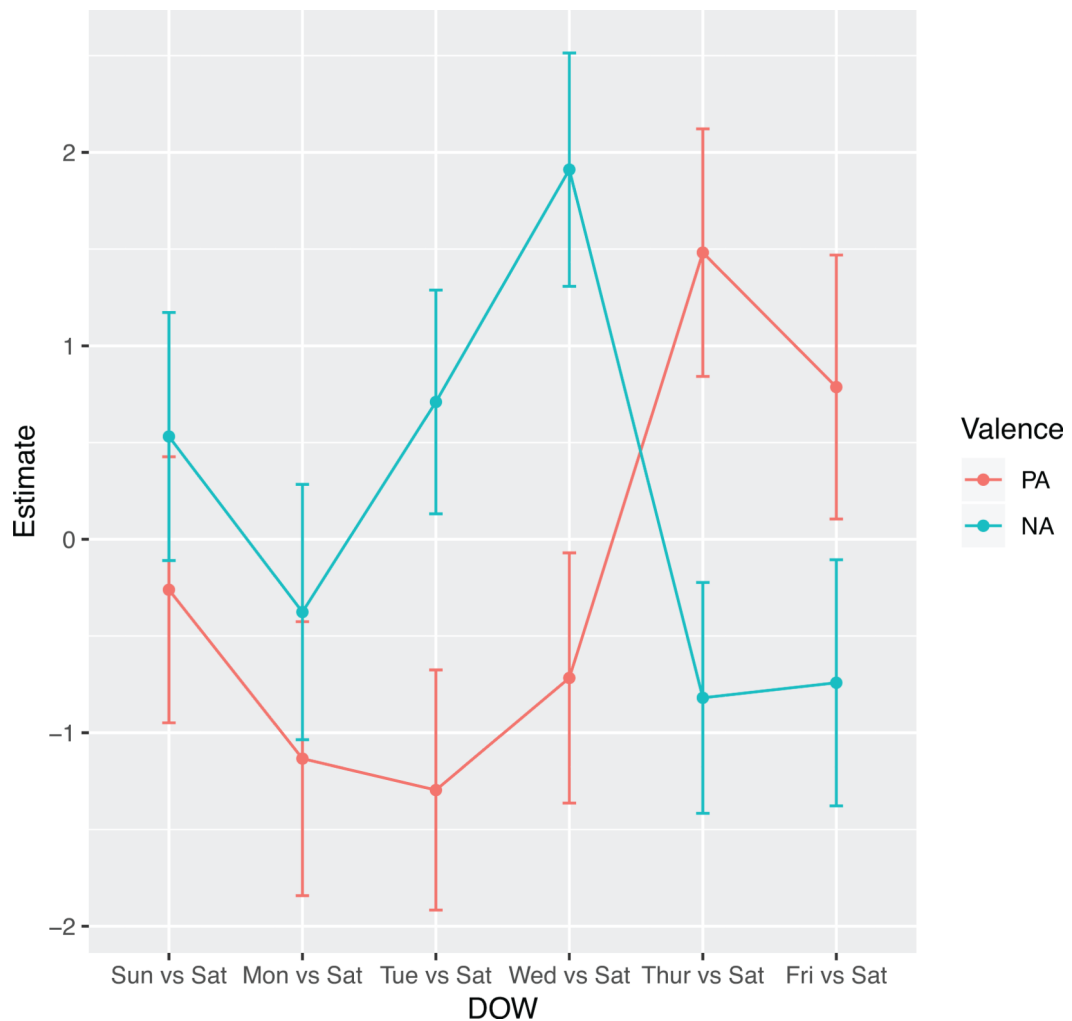


**Extended Data Figure 1.**

Positive and negative affect by day of week. This figure depicts the parameter estimates of the effect of day of week (DOW) on positive and negative affect from linear mixed-effects models that included RE, time of day EMA was collected, distance travelled, mean temperature and precipitation (from the modal location of that day for that participant), and cohort as predictors (n = 122 independent human participants). Of the variables in models, only day of the week was also significantly associated with PA, with PA lower early to mid-week relative to the weekend (F(3834) = 2.8, p = 0.011). Both temperature (B = 0.062, t(3835) = 2.585, p = 0.01) and day of the week (F(3835) = 3.0, p = 0.006) were significantly associated with negative affect, which was greater on higher temperature days and in the early to mid-week relative to the weekend. Error bars represent standard error of parameter estimate.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements

## Data availability

The raw geolocation datasets generated and/or analyzed during the current study are not publicly available due to the inherently identifiable nature of geolocation data, but are available from the corresponding authors upon reasonable request. The processed data necessary to reproduce the central findings in the manuscript are available at http://github.com/manateelab/NatNeuro.

## References

1. Tost H, Champagne FA & Meyer-Lindenberg A Nature Neuroscience 18, 1421–1431 (2015). [PubMed: 26404717]

2. Marashi V, Barnekow A, Ossendorf E & Sachser N Hormones and Behavior 43, 281–292 (2003). [PubMed: 12694638]

3. Laviola G et al. European Journal of Neuroscience 20, 1655–1664 (2004). [PubMed: 15355333]

4. Brydges NM, Leach M, Nicol K, Wright R & Bateson M Animal Behaviour 81, 169–175 (2011).

5. Fox NA, Henderson HA, Marshall PJ, Nichols KE & Ghera MM Annu Rev Psychol 56, 235–262 (2005). [PubMed: 15709935]

6. Fredrickson BL Philosophical Transactions of the Royal Society B: Biological Sciences 359, 1367–1377 (2004).

7. van Praag H, Kempermann G & Gage FH Nat. Rev. Neurosci. 1, 191–198 (2000). [PubMed: 11257907]

8. Walsh RN & Cummins RA Psychological Bulletin 82, 986 (1975). [PubMed: 1103204]

9. Lisman JE & Grace AA Neuron 46, 703–713 (2005). [PubMed: 15924857]

10. Düzel E, Bunzeck N, Guitart-Masip M & Düzel S Neuroscience & Biobehavioral Reviews 34, 660–669 (2010). [PubMed: 19715723]

11. Bevins RA & Bardo MT Behavioural Brain Research 99, 53–60 (1999). [PubMed: 10512572]

12. Haber SN & Knutson B Neuropsychopharmacology 35, 4–26 (2010). [PubMed: 19812543]

13. Berlyne DE Perception & Psychophysics 8, 279–286 (1970).

14. Freund J et al. Science 340, 756–759 (2013). [PubMed: 23661762]

15. Garland EL et al. Clinical Psychology Review 30, 849–864 (2010). [PubMed: 20363063]

16. Sjulson L, Peyrache A, Cumpelik A, Cassataro D & Buzsáki G Neuron (2018).doi:10.1016/j.neuron.2018.04.015

17. Cohen MX, Schoene-Bake J-C, Elger CE & Weber B Nature Neuroscience 12, 32–34 (2009). [PubMed: 19029888]

18. Kühn S et al. Scientific Reports 7, (2017).

19. Schooler C Intelligence 8, 259–281 (1984).

20. Schwanen T, Kwan M-P & Ren F Geoforum 39, 2109–2121 (2008).

## Methods only references:

21. Beck AT & Steer RA (Psychological Corp. ; Harcourt Brace Jovanovich: San Antonio, Tex.; New York, 1987).

22. Spielberger CD (1983).at <http://ubir.buffalo.edu/xmlui/handle/10477/1873>

23. Watson D & Clark LA (1999).

24. Freund J et al. Neuroscience 309, 140–152 (2015). [PubMed: 25987202]

25. Pinheiro J, Bates D, DebRoy S & Sarkar D (2014).

26. Akaike H IEEE Transactions on Automatic Control 19, 716–723 (1974).

27. Revelle W URL http://cran.r-project.org/web/packages/psych/index.html (2018).

28. Satterthwaite TD et al. NeuroImage 60, 623–632 (2012). [PubMed: 22233733]

29. Satterthwaite TD et al. NeuroImage 64, 240–256 (2013). [PubMed: 22926292]

30. Power JD, Barnes KA, Snyder AZ, Schlaggar BL & Petersen SE Neuroimage 59, 2142–2154 (2012). [PubMed: 22019881]

31. Power JD et al. Neuroimage 84, (2014).

32. Kebets V et al. bioRxiv 637827 (2019).doi:10.1101/637827

33. Buckner RL, Andrews-Hanna JR & Schacter DL Annals of the New York Academy of Sciences 1124, 1–38 (2008). [PubMed: 18400922]

34. Desikan RS et al. Neuroimage 31, 968–980 (2006). [PubMed: 16530430]

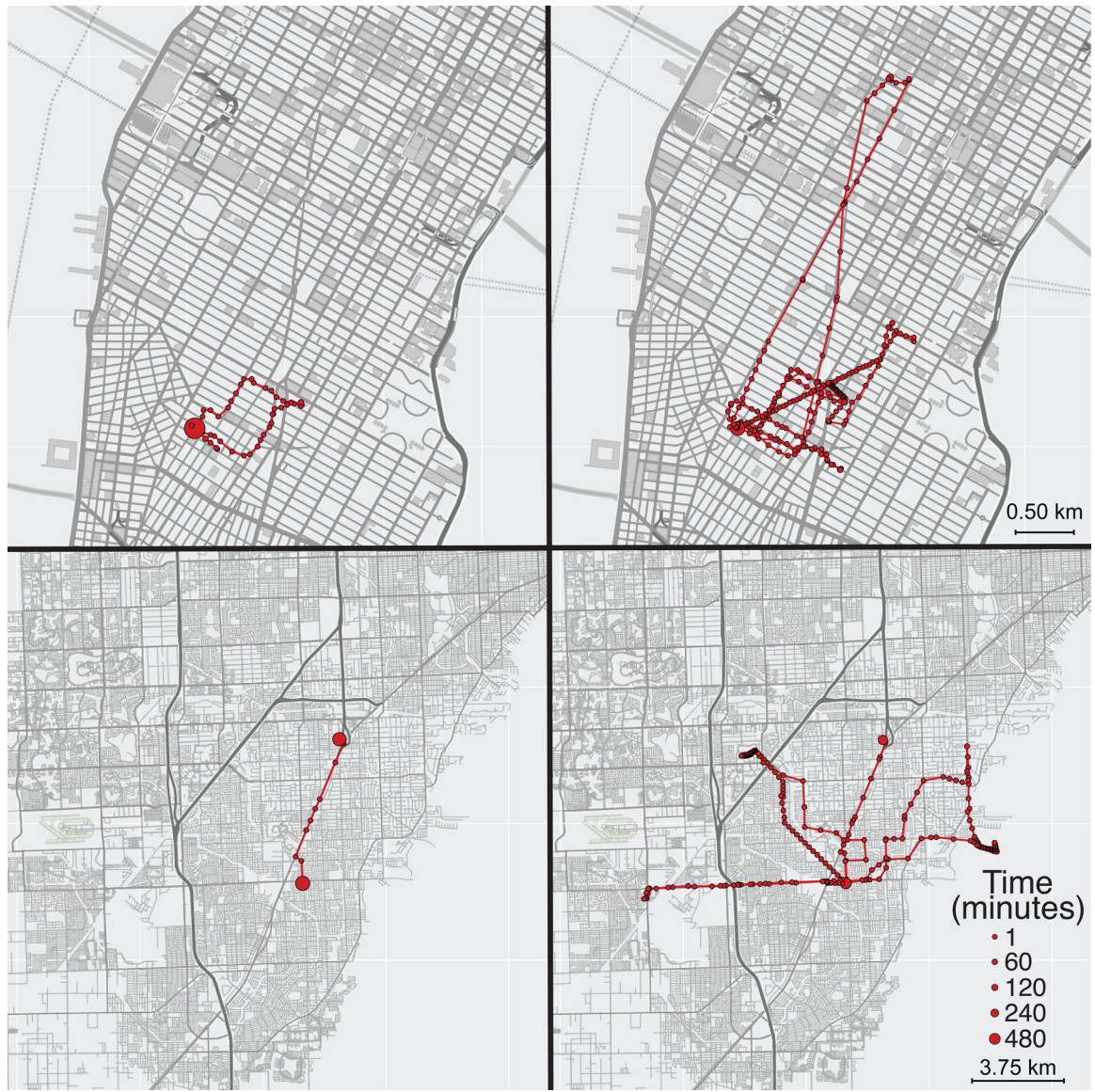35. Tost H et al. Nat Neurosci 22, 1389–1393 (2019). [PubMed: 31358990]

**Figure 1. Geolocation tracking in New York and Miami.**
Representative low (left column) and high (right column) roaming entropy days from one participant in New York (top row) and one in Miami (bottom row). Map tiles by © Stamen Design, under CC BY 3.0. Map data by OpenStreetMap.org contributors.
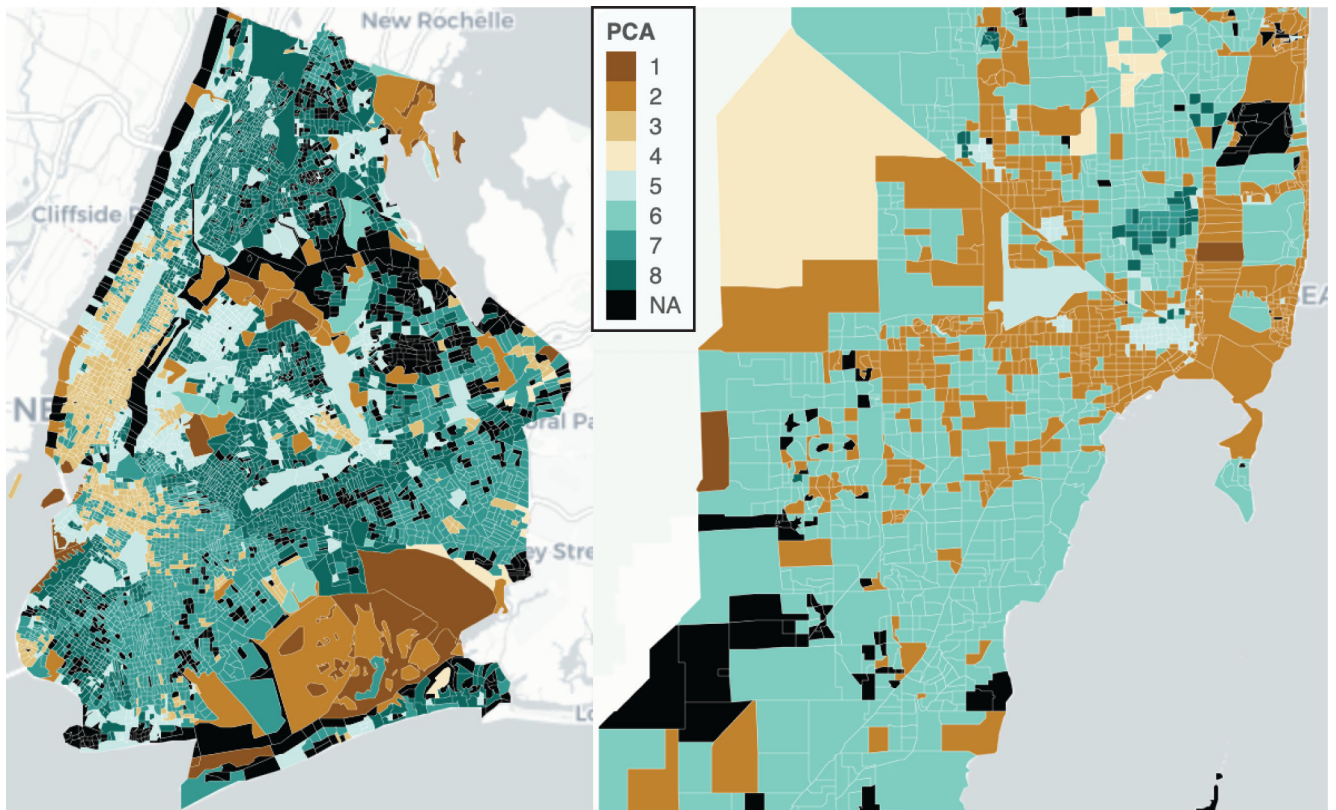
**Figure 2. Physical location transformed into sociodemographic feature space.**
Maps depicting the sociodemographic feature space of New York City and Miami at the
spatial resolution of a block group (the smallest geographical unit for which the US Census
publishes data). Color of locations represents the positions in this sociodemographic feature
space derived from our discretized PCA. Daily entropy over this sociodemographic feature
space was associated with positive affect (n = 122 independent human subjects). Map tiles
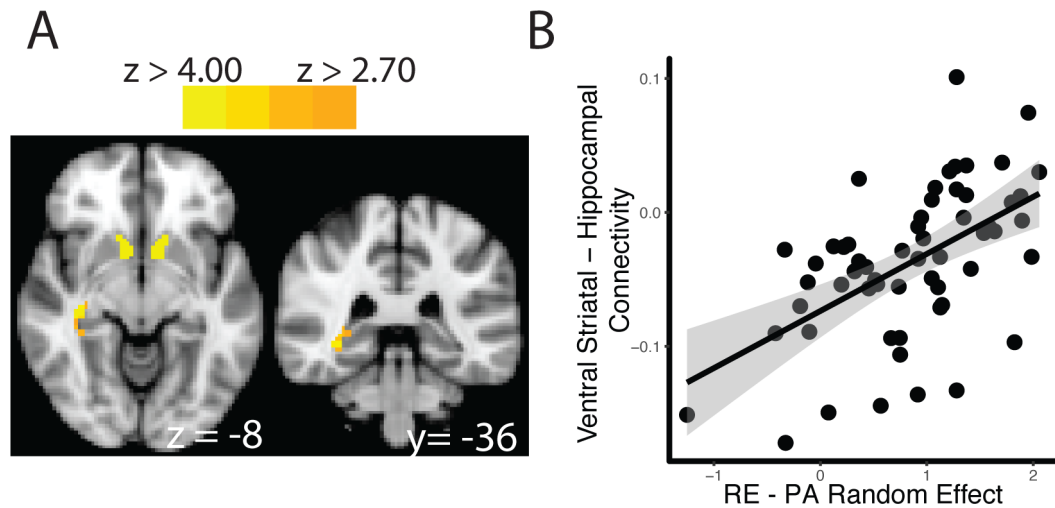by CARTO under CC BY 4.0. Map data by OpenStreetMap contributors.

**Figure 3. Entropy-affect association is linked to hippocampal-ventral striatal connectivity.**
Random-effects, reflecting individual differences in the degree to which roaming entropy is associated with positive affect, correlate positively with resting-state functional connectivity between a ventral striatal seed and the posterior hippocampus (A & B). Shaded band represents the 95% confidence interval on the best fitting regression line (n = 58 participants).