Hunters, busybodies, and the knowledge network building associated with curiosity

David M. Lydon-Staley[a], **Dale Zhou**[b], **Ann Sizemore Blevins**[a], **Perry Zurn**[c], **and Danielle S. Bassett**[a,d,e,f,g,*]

[a]Department of Bioengineering, School of Engineering & Applied Science, University of Pennsylvania, Philadelphia, PA 19104 USA
[b]Department of Neuroscience, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104 USA
[c]Department of Philosophy, American University, Washington, DC 20016 USA
[d]Department of Electrical & Systems Engineering, School of Engineering & Applied Science, University of Pennsylvania, Philadelphia, PA 19104 USA
[e]Department of Neurology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104 USA
[f]Department of Physics & Astronomy, College of Arts & Sciences, University of Pennsylvania, Philadelphia, PA 19104 USA
[g]Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104 USA
[*]dsb@seas.upenn.edu

ABSTRACT

The information gained when practicing curiosity promotes well-being over extended timescales. The open-ended and internally driven nature of curiosity, however, makes characterizing the diverse styles of information seeking that accompany it a daunting endeavor. A recently developed historicophilosophical taxonomy of curious practice distinguishes between the collection of disparate, loosely connected pieces of information and the seeking of related, tightly connected pieces of information. With this taxonomy, we use a novel knowledge network building framework of curiosity to capture styles of curious information seeking in 149 participants as they explore Wikipedia for over 5 hours spanning 21 days. We create knowledge networks in which nodes consist of distinct concepts (unique Wikipedia pages) and edges represent the similarity between the content of Wikipedia pages. We quantify the tightness of each participants' knowledge networks using graph theoretical indices and use a generative model of network growth to explore mechanisms underlying the observed information seeking. We find that participants create knowledge networks with small-world and modular structure. Deprivation sensitivity, the tendency to seek information that eliminates knowledge gaps, is associated with the creation of relatively tight networks and a relatively greater tendency to return to previously-visited concepts. We further show that there is substantial within-person variability in knowledge network building over time and that building looser networks than usual is linked with higher than usual sensation seeking. With this framework in hand, future research can quantify the information collected during curious practice and examine its association with well-being.

Introduction

Curiosity is characterized by intrinsically motivated information seeking[1–3]. The information sought while practicing curiosity often has no immediate, tangible benefit[4–6]. Acting on one's curiosity can lead to suboptimal behavior when the goal is to maximize reward, including accepting monetary costs and risking exposure to electric shocks to receive non-instrumental information[7–10]. Despite a lack of immediate benefits, the tendency to frequently experience and practice curiosity is robustly associated with positive well-being[11–13]; curiosity facilitates focused engagement with novel and challenging stimuli and, in the process, the accrual of information and other resources that, although not of immediate benefit, may have utility when encountering future challenges[14–16]. And irrespective of its immediate or potential utility, the practice of curiosity may well be valuable in itself[17].

Characterizing how individuals seek information when internally driven is fundamental to understanding how the practice of curiosity leads to the shoring up of information and resources that impact well-being. Historicophilosophical studies tracing the use of the word curiosity (including its Greek, Latin, French, German, and English terms) have identified a few styles of curious information seeking that span multiple millennia, cultures, and languages[18]. The styles include the busybody and the hunter. The information seeking of the busybody is marked by a preference for sampling diverse concepts, characterized by "distraction" and "never-dwelling anywhere"[19] (p.161). The busybody will "frisk about, and rove about, at random, wherever

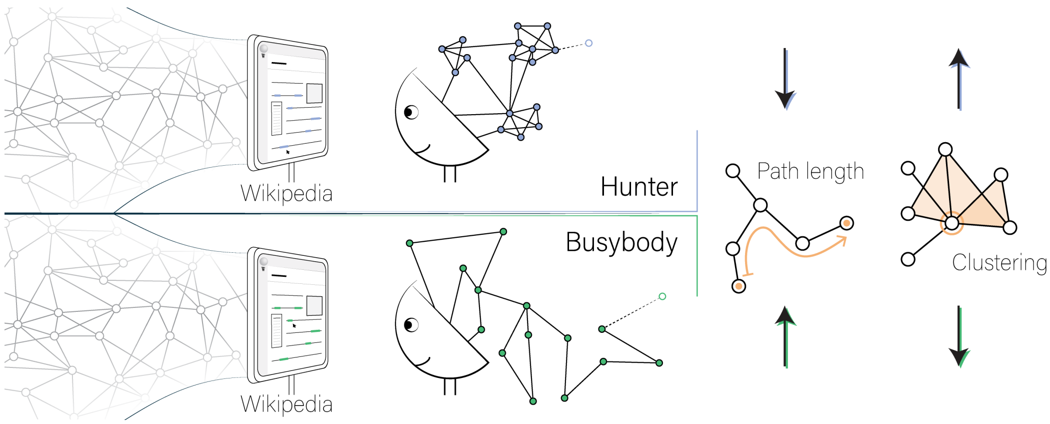

Figure 1. Hunter and busybody styles of information seeking. Participants explored Wikipedia for 15 minutes every day for 21 days. We represent participants' information seeking as knowledge networks[24]. Nodes represent the unique Wikipedia pages visited, and edges represent the similarity between the text content of each page. We use a historicophilosophical taxonomy of curious information seeking[18] to examine between-person differences in the resulting networks. The busybody samples diverse concepts and creates loose knowledge networks of sparsely connected concepts. In contrast, the hunter creates tight knowledge networks characterized by sampling related concepts. We operationalize notions of network tightness using graph theoretical indices. Intuitively, the characteristic path length assesses the average distance between all pairs of nodes in a network. When path length is short, the network is easily traversed and representative of the hunter's tight networks. The clustering coefficient indicates the extent to which a node's neighbors are connected. A high average clustering coefficient indicates a tight network of closely connected concepts, which is the kind we expect of the hunter.

they please"[20] (sec. 34). The information seeking of the hunter, by contrast, is characterized by sampling closely connected concepts. The hunter does not "turn aside and follow every scent"[21] (p.520e) in the manner of the busybody. The hunter instead "wishes [they] had a few hundred helpers and good, well-trained hounds that [they] could drive into the history of the human soul to round up [their] game"[22] (p.59) in a targeted search for information. Both styles are considered curious practice, but there are individual differences in the extent to which each type of curious practice is expressed in behavior[23]. Tendencies to exhibit one form of curiosity practice over another will lead to the accumulation of different types of information and resources over time. The busybody's store of information will be more diverse relative to that of the hunter, but the hunter's information store will contain greater depth on fewer subjects.

The open-ended, internally driven nature of curiosity makes characterizing the diverse information seeking styles of humans a daunting endeavor. Existing approaches include the examination of saccadic exploration of visual scenes and responses to trivia questions designed to evoke curiosity[25,26]. Carefully designed experimental paradigms are beginning to shed light on curiosity, but they have been met with calls to consider more complex forms of information seeking that occur over extended timescales[2]. We claim that styles of information seeking identified through historicophilosophical methodologies can be readily accommodated within a recently developed knowledge network building framework[24]. From this perspective, network nodes represent distinct concepts and network edges represent the manner in which the concepts are related. While practicing curiosity, an individual traverses edges on knowledge networks as they move from one concept to the next. Some of the edges they traverse may have large weights, indicating that the two concepts joined by the edge are very similar, and some edges may have very small weights indicating that the two concepts are virtually unrelated. Casting curiosity as a knowledge network building practice reflects the interconnectedness of informational units[27] and allows an application of the mathematical language of graph theory[28,29] to precisely quantify complex manifestations of curious behavior. The easily distracted busybody will be marked by the creation of loose knowledge networks of sparsely connected, seemingly unrelated concepts. In the parlance of graph theory, their networks will have small edge weights, low clustering, and high characteristic path length. The more targeted sampling of the hunter, in contrast, will result in tight knowledge networks consisting of closely-connected concepts and their networks will have large edge weights, high clustering, and low characteristic path length (Fig. 1).

We can measure the associations between knowledge network structure and existing instruments designed to measure curiosity. Increasingly, curiosity is measured as a multifaceted trait, often containing a facet capturing the tendency to seek information due to feelings of deprivation[30,31]. The extent to which one's information seeking is motivated to overcome the feeling of being deprived of knowledge has emerged as a robust individual difference[32,33]. Individuals high in deprivation sensitivity possess a strong drive to know, and they tend to seek information that eliminates gaps in their knowledge[30,34]. Thus, we hypothesize that individuals high in deprivation sensitivity will create tighter knowledge networks as they encounter new information, recognize gaps in their knowledge, and eliminate these knowledge gaps by searching for closely related concepts[35]. In contrast, we expect individuals low in deprivation sensitivity to create looser knowledge networks.

Here, we operationalize curiosity as a knowledge network building practice. We monitor the internally directed information seeking of 149 individuals on Wikipedia, an online encyclopedia, over the course of 21 days. The choice of Wikipedia reflects its use as a knowledge network in previous research (e.g.,[36]) and findings that intrinsic learning forms a major motivation for use of Wikipedia in everyday life[37,38]. We treat each Wikipedia page as a distinct concept or node in a knowledge network, and we use natural language processing tools to determine the weight of network edges reflecting similarities in semantic content between any two pages. Once the network is constructed, we use graph theory to quantify general notions of tight and loose knowledge networks to realize busybody and hunter styles of information seeking. We complement this analytic investigation with the development of a generative model of information search to uncover potential mechanisms underlying knowledge network growth. We also examine associations between the structure of knowledge networks and existing measures of trait deprivation sensitivity, with the hypothesis that individuals high in deprivation sensitivity will create relatively tight, hunter-like knowledge networks and conversely that individuals low in deprivation sensitivity will create relatively loose, busybody-like knowledge networks. We further examine the extent to which the tightness of knowledge networks changes over the course of the study. Curiosity exhibits within-person fluctuations over relatively short time-scales (e.g., from day to day[11,39]) and hunter and busy-body styles of knowledge network building are thought to be expressed to different degrees across time *within-person*[24]. We hypothesize that periods during which loose knowledge networks are created, reflecting the pursuit of novel, diverse, and varied information, will be periods in which individuals experience heightened sensation seeking tendencies that promote the "seeking of varied, novel, complex, and intense sensations and experiences" (p.26[40];[41]).

We demonstrate the feasibility of quantifying complex styles of internally directed information seeking within a network science framework and find that the observed knowledge networks have a small-world structure, are highly modular, and show substantial within-person variation. We discover that participants high in deprivation sensitivity create tight knowledge networks with high edge weights, high clustering, and short characteristic path lengths, and that they have a greater tendency to return to concepts they have previously visited, relative to participants low in deprivation sensitivity. We find that knowledge network building shows substantial within-person variability and that building looser networks than usual is associated with higher than usual sensation seeking.

Results

We aim to operationalize busybody and hunter styles of curious practice in a specific instance of knowledge network building. Using the internally-directed Wikipedia browsing of 149 participants for 15 minutes each day across 21 days, we treat each Wikipedia page visited as a network node, and we define the weight of network edges as the cosine similarity of the term-frequency inverse document frequency of the text contained within each page (Fig. 2A-B). Thus, a high edge weight indicates similarity in terms of the text contained in the two nodes connected by the edge. We interrogate the structure of each network using graph theoretical indices, and we apply a generative model of network growth to provide insight into mechanisms underlying network building. Participants completed a self-reported survey of trait curiosity[30] prior to beginning the information seeking task, allowing us to examine associations between aspects of network structure and trait deprivation sensitivity. In all models, we include four other facets of curiosity (joyous exploration, social curiosity, thrill seeking, and stress tolerance) to examine the extent to which the hypothesized associations are specific to deprivation sensitivity. Participants also reported on their sensation seeking tendencies each day across the 21 day period immediately prior to the Wikipedia browsing task, allowing us to examine how within-person changes in sensation seeking are associated with changes in knowledge network building. Details of knowledge network construction and all measures are provided in the Materials and Methods section and Supplementary Materials. Descriptive statistics of key variables can be found in Table S1.

Deprivation sensitivity is positively associated with the average edge weight of knowledge networks

Participants completed an average of 17.90 (SD = 3.21) days of Wikipedia browsing. The median number of edges in participants' knowledge networks is 168 (IQR = 143), where each edge indicates a transition from one Wikipedia page to another. Participants visited a median of 135 (IQR = 99) unique nodes. The average weight of all edges in each participant's network is 0.18 (SD = 0.04). Intuitively, network building that is more hunter-like is reflected in the right side of the distribution of average edge weights in the sample (relatively high average edge weights) and busybody-like network building is reflected in the left side of the distribution (relatively lower average edge weights); see Fig. 2C.

We used a multilevel model to assess the relation between information seeking behavior and trait curiosity. We find that deprivation sensitivity is positively associated with average edge weight (b=0.004, p=0.01; Cohen's d = 0.44[42], Table S2), indicating that participants high in deprivation sensitivity are more hunter-like in their knowledge network building relative to participants low in deprivation sensitivity, who in contrast are more busybody-like in their information seeking (Fig. 2D). Interestingly, two other facets of curiosity show significant associations with average edge weight; joyous exploration is

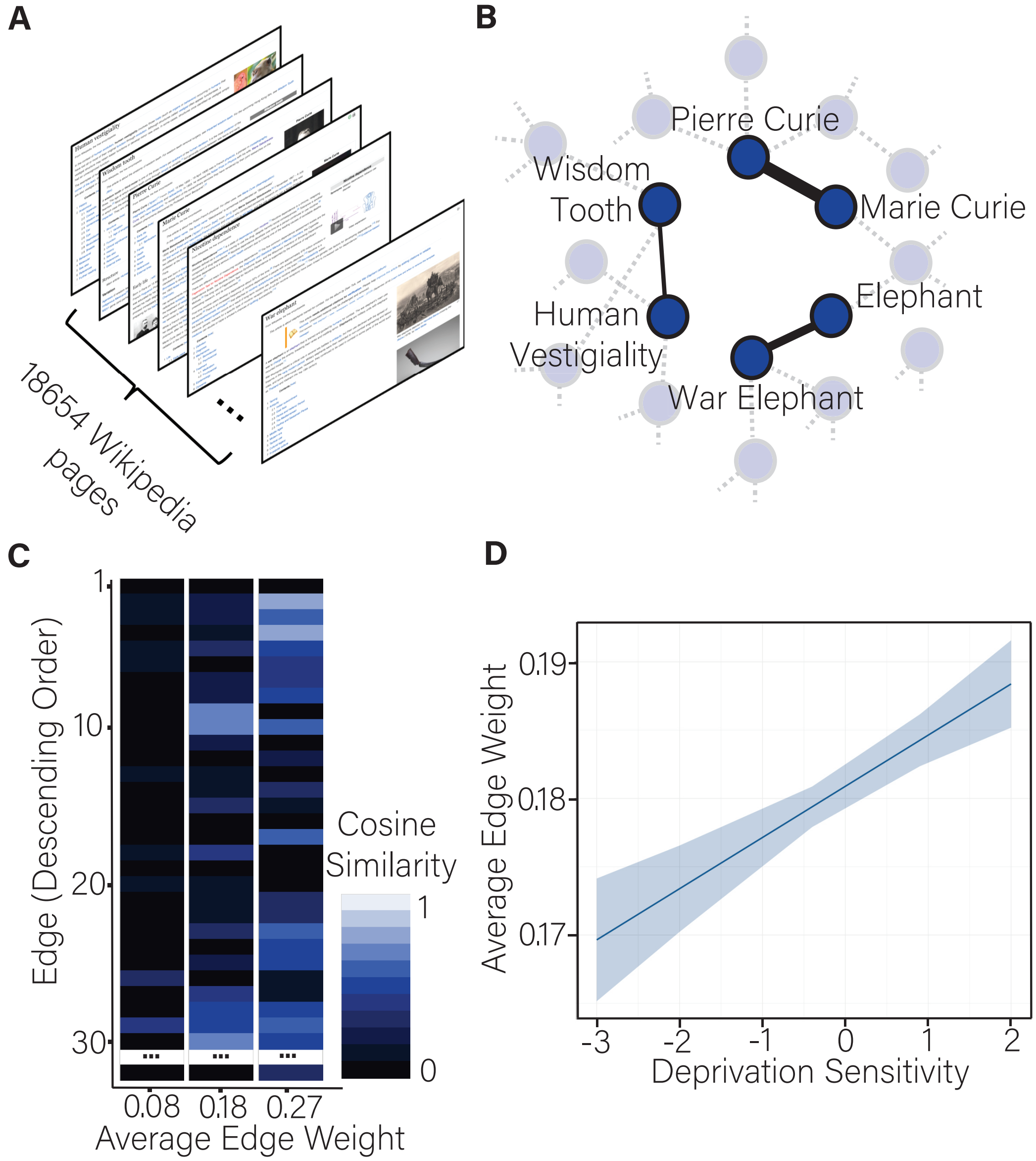

Figure 2. Knowledge network construction and the association between deprivation sensitivity and edge weight. (*A*) The sample visited 18654 Wikipedia pages. (*B*) Network nodes represent all the unique pages visited by all participants in the sample. Weighted network edges represent the cosine similarity (bounded between 0 and 1) between all possible pairs of vectors of term-frequency inverse document frequencies associated with the text of each page. Edges with higher weights indicate relatively greater semantic similarity between nodes. For example, the edge between "Marie Curie" and "Pierre Curie" has a cosine similarity value of 0.8, and the edge between "Wisdom Tooth" and "Human Vestigiality" has a value of 0.2. (*C*) The partial time series of edges traversed by an individual who tended to visit loosely connected concepts (*left*), an individual who tended to visit strongly connected concepts (*right*), and an individual whose network had the average edge weight for the sample (*middle*). In a section of their edge weight time series, the participant on the left with lower than average edge weight sought out "Physical chemistry", "Me Too movement", "The Partridge Family", "Harborne Primary School", "HIP 79431", and "Tom Bigelow", which collectively appear to be a rather diverse set of concepts. In contrast, the participant with relatively high average edge weight visited "History of the Jews in Germany", "Hep-Hep riots", "Zionism", "Nathan Birnbaum", and "Theodor Herzl", which comprise a closely connected set of concepts in Jewish history. (*D*) Multilevel model results show that participants high in deprivation sensitivity had higher average edge weights, indicating that they tended to visit similar concepts as they traversed Wikipedia. The ribbon around the model estimated association represents the standard error.

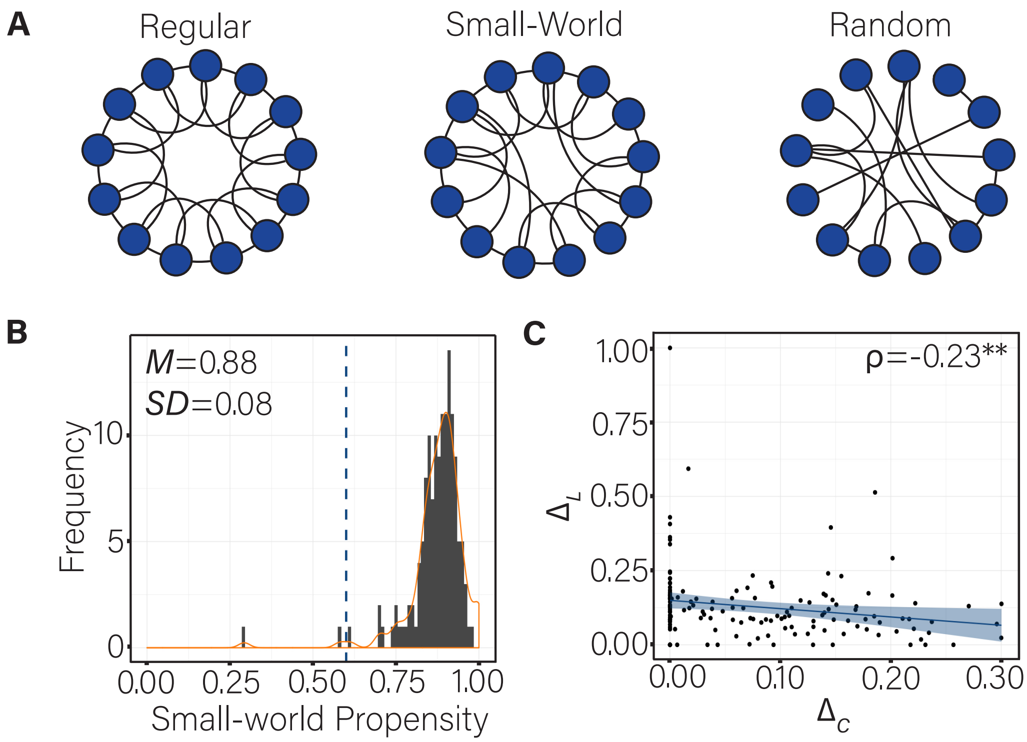

Figure 3. Small-world propensity of knowledge networks. (*A*) Small-world networks (*middle*) show high clustering like regular lattices (*left*) and short characteristic path lengths like random graphs[43] (*right*). (*B*) The distribution of small-world propensity values of the knowledge networks in the sample. All but two participants' networks were above the pragmatic but heuristic small-world propensity cut-off value of 0.60 indicated by the dashed blue line[44]. Deviations from small-world organization can reflect deviations in clustering coefficient, in characteristic path length, or in both metrics. (*C*) The fractional deviation of the observed networks from the null model shown for the clustering coefficient ($\triangle_C$) and characteristic path length ($\triangle_L$). The blue line indicates a line of best fit. The ribbon around the line of best fit represents the standard error. Results of a Spearman correlation between both variables is shown in the top right. *Notes*: $^{**}p<0.01$; *M*=mean; *SD*=standard deviation.

negatively associated with average edge weight (b=-0.005, p=0.01, Cohen's d = 0.44), while stress tolerance is positively associated with average edge weight (b=0.005, p<0.001, Cohen's d = 0.61). Neither social curiosity (b=-0.001, p=0.32, Cohen's d = 0.17) nor thrill seeking (b=-0.002, p=0.11, Cohen's d = 0.27) are associated with average edge weight. Notably, these associations remain significant after controlling for the number of edges traversed (b=0.00003 , p=0.04, Cohen's d = 0.36) and the proportion of edges traversable via a hyperlink (b=0.18, p<0.001, Cohen's d = 2.59). Note that participants could use the search function to move from page to page and were not restricted to the use of hyperlinks. Moreover, the associations also remain significant when accounting for the positive association between the cosine similarity value of each edge and the presence of a hyperlink between each pair of nodes (b=0.19, p<0.001, Cohen's d = 1.07). The robustness of the results is further underscored by the fact that the associations remain significant after excluding a participant with an average edge weight value that was six standard deviations above the mean (Table S3).

Deprivation sensitivity is positively associated with knowledge network clustering and negatively associated with characteristic path length

Next, we created participant-specific networks consisting of all Wikipedia pages that a participant visited and all possible edges between those nodes, even if those specific edges were not traversed during the information seeking task. We calculated the average clustering coefficient of each participant's knowledge network. The clustering coefficient provides an indication of the extent to which a node's neighbors are connected. In the form that we employ, the weighted clustering coefficient gives the average intensity (geometric mean) of all triangles associated with each node[45]. We took the mean clustering coefficient of each node in participants' knowledge networks to quantify general notions of tight and loose knowledge networks, with high average clustering coefficients indicative of networks consisting of closely connected concepts and low average clustering coefficients indicative of networks consisting of sparsely connected concepts. The average clustering coefficient in these

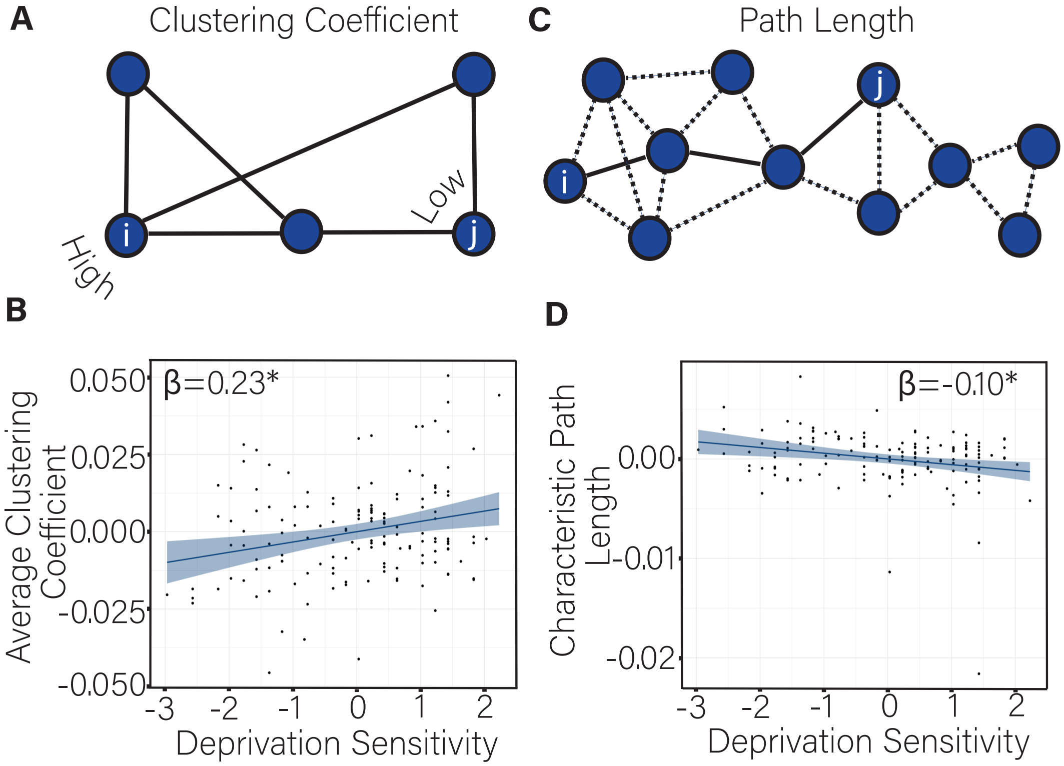

Figure 4. Deprivation sensitivity and the clustering and path length of knowledge networks. (*A*) We characterized the extent to which a node's neighbors are connected by calculating the clustering coefficient on participant-specific networks in which each node is a unique Wikipedia page visited by the participant and edges exist between all possible node pairs and are weighted by a cosine similarity value. Here we show a network schematic in which node i has a high clustering coefficient while node j has a low clustering coefficient; the neighbors of node i are more likely to be neighbors of one another than the neighbors of node j. (*B*) A partial residual plot shows that deprivation sensitivity is positively associated with the average clustering coefficient. The ribbon around the line of best fit represents the standard error. (*C*) We also quantified the characteristic path length of each participant's network. The shortest path between node i and node j is displayed as a continuous line. The characteristic path length can be thought of as the average distance along the shortest paths for all possible pairs of nodes in the network. (*D*) A partial residual plot shows that deprivation sensitivity was negatively associated with the characteristic path length. The ribbon around the line of best fit represents the standard error. In panels (*A*) and (*C*) we show binary networks to provide intuition, but in all analyses we used weighted path length and clustering coefficient to maintain sensitivity to individual differences in network geometry. *Note*: $*p<0.05$; β=standardized regression coefficient.

networks is 0.09 (SD=0.02). To conceptually link this metric to curious practices, we note that network building that is more hunter-like is reflected in the right side of the distribution of average clustering coefficients (relatively high clustering) and more busybody-like network building is reflected in the left side of the distribution (relatively low clustering).

In a complementary assessment of the extent of similarity between concepts explored by participants, we computed the characteristic path length of each participant's network. Intuitively, the characteristic path length assesses the average distance between all pairs of nodes in a network. When the characteristic path length is short, the network is easily traversed[46]. The mean characteristic path length in these networks is 0.99 (SD=0.03). To conceptually link this metric to curious practices, we note that network building that is more hunter-like is reflected in the left side of the distribution of the characteristic path length (relatively short path lengths) and more busybody-like network building is reflected in the right side of the distribution (relatively long path lengths).

We find a negative Pearson correlation between the average clustering coefficient and the characteristic path length (r=-0.50, p<0.001), such that people high in clustering tend to have short path lengths. This moderate tracking of clustering with path length in the sample motivated an analysis of the propensity of knowledge networks to form small-world structures, which show high clustering and short characteristic path length relative to random graphs (Fig. 3A-B). Formally, the small-world propensity[44] assesses the degree to which the observed clustering coefficient and characteristic path length align with that expected in lattice and random benchmark networks, respectively. Intuitively and in the context of these data, the metric can probe the balance between exploring related concepts and efficiently spanning the chosen space of concepts. The average small-world propensity value (ϕ) was 0.88 (*SD*=0.08). For all but two participants, the ϕ values of the networks fall above the commonly used threshold of 0.6, indicating that the majority of networks are small-world in nature. Thus, participants' knowledge network building exhibits a priority for exploiting similar concepts, reflected in higher clustering, while efficiently spanning the chosen space of concepts, reflected in shorter path length, than would be expected in random and lattice graphs. To unpack this finding further, we examine the contributions to ϕ of deviations in clustering ($\triangle_C$) and path length ($\triangle_L$) from the respective benchmark models (Fig. 3C). We observe an even distribution of deviations in clustering coefficient (*mean* = 0.08, *SD*=0.08, *skew*=0.75), and a more heavy tailed distribution of deviations in characteristic path length (*mean*=0.13, *SD*=0.12, *skew*=3.32), suggesting that deviations in clustering play a greater role relative to path length in the observed individual differences in the small-world propensity of knowledge networks. Deviations in clustering and deviations in path length are negatively correlated (ρ=-0.23, p=0.004), indicating a balance between exploiting neighboring concepts in knowledge networks and efficiently traversing the concept space as a whole, which is a trade-off that we also observe in the raw clustering and path length values. In sum, these findings suggest that participants balance local explorations with efficient traversal of the concept space.

We used multiple regression analysis to test the extent to which deprivation sensitivity is involved in this balance of exploiting information in neighboring concepts while also ensuring that the chosen concept space can be efficiently traversed. We regressed the average clustering coefficient on deprivation sensitivity while controlling for the other four facets of curiosity (joyous exploration, social curiosity, thrill seeking, and stress tolerance) as well as network density and size, which are known confounds in network studies[47] (Fig. 4A). We removed an outlier value of the clustering coefficient (0.28, 9.5 standard deviations above the mean) before performing the analysis. The predictors explain 20 percent of the variance in the clustering coefficient (R^2=0.20, $F(7,140)$=5.08, $p < 0.001$; Table S4). Deprivation sensitivity is positively associated with the average clustering coefficient (b=0.003, p=0.01, β=0.23, Fig. 4B), suggesting that participants high in deprivation sensitivity examine closely related concepts during information seeking to a greater extent than participants low in deprivation sensitivity. Social curiosity is negatively associated with the average clustering coefficient (b=-0.003, p=0.03, β=-0.19). No other facets of curiosity are associated with the average clustering coefficient of the subject-specific knowledge network. Notably, the associations are significant when controlling for network density (b=0.07, p=0.08, β=0.15) and network size (i.e., number of nodes; b=-0.0001, p=0.001, β=-0.27).

We next regressed the characteristic path length on deprivation sensitivity while controlling for the other four facets of curiosity as well as both network density and size (Fig. 4C). We removed an outlier value of the characteristic path length (0.69, 10 standard deviations below the mean) before performing the analysis. The predictors explain 82 percent of the variance in the characteristic path length, R^2=0.82, $F(7,140)$=90.64, $p < 0.001$ (Table S5). Deprivation sensitivity is negatively associated with the characteristic path length (b=-0.001, p=0.02, β=-0.10; Fig. 4D) such that participants high in deprivation sensitivity, while exploiting local information, also have networks that are easily traversable from one end to the next. No other facets of curiosity are associated with path length. Notably, the associations are significant when controlling for network density (b=-0.15, p<0.001, β=-0.82) and network size (b=0.00002, p<0.001, β=0.19).

Knowledge networks exhibit modular structure

Average clustering coefficient and path length provide two complementary measures of tightness and looseness of knowledge networks. To further examine the organization of knowledge networks[48], we performed community detection. This approach allowed us to assign each node in each participant's knowledge network to a community within which concepts explored were similar to one another while dissimilar to the concepts explored in nodes outside of the community (Fig. 5A). The sample average of the network modularity quality index, Q[49], is 0.59 (*SD*=0.12). Comparing the Q of observed networks against the Q in randomized versions of the observed networks that preserved degree distributions[50], a paired samples t-test shows higher Q of the observed relative to the null networks $t(148)$=41.601, p<0.001, Cohen's D=3.41, indicating that the observed knowledge networks show non-trivial community structure (Fig. 5B). Regression analysis (Table S6) reveals that higher clustering is associated with higher Q (b=2.52, p=<0.001, β=0.39) and that longer path length is associated with higher Q (b=6.70, p=0.003, β=0.39) while controlling for density (b=-1.82, p<0.001, β=-0.57) and network size (b=-0.0004, p<0.001, β=-0.30). We examine the correlation between Q and deprivation sensitivity (Fig. 5C), providing some evidence that participants high in deprivation sensitivity create more modular networks (r=0.17, p=0.04), although this association did not survive when controlling for network density (Table S7). Thus, although deprivation sensitivity's positive association with average clustering coefficient may promote more modular structure given the positive association between clustering and Q, this tendency towards greater modularity is likely balanced by deprivation sensitivity's negative association with characteristic path length.

To further probe the structure of participants' knowledge networks, we examine the number of communities identified by community detection. Knowledge networks contain an average of 13.82 (*SD*=5.46) communities. Linear regression indicates that deprivation sensitivity is positively associated with the number of communities (b=0.10, p=0.03, β=0.18) such that the knowledge networks of participants high in deprivation sensitivity contain more communities relative to participants low in deprivation sensitivity (Fig. 5D). This association is significant after controlling for a number of network confounds, including network density (b=-3.09, p<0.001, β=-0.37), network size (b=0.002, p=0.007, β=0.20), and Q (b=1.10, p=0.03, β=0.19). Joyous exploration is negatively associated with the number of communities (b=-0.15, p=0.01, β=-0.23) and stress tolerance is positively associated with the number of communities (b=0.09, p=0.04, β=0.17). No other facets of curiosity are associated with the number of communities.

Principles of knowledge network growth and associations with curiosity

Our examination of average edge weights, clustering coefficient, characteristic path length, and modular structure are descriptions of the networks created by participants. In a next step, we moved beyond descriptions of network structure by using a generative model to explore potential network mechanisms underlying the observed patterns of information seeking. Our generative model represents the network growth mechanisms that a simulated agent uses to construct networks with different structures. By fitting an agent's growth mechanisms to the empirical sequence in which participants traversed edges on Wikipedia, we characterized how participants' differing information seeking patterns arise from formal growth rules. Tight networks could emerge from a greater tendency to revisit similar concepts, a lesser propensity to make large conceptual leaps when moving from page to page, or a combination of both. These possibilities guided our choice of network growth model. We formalized these possibilities for underlying principles that led to differences in the tightness of knowledge networks using two growth rules. The first growth rule is *reinforcement* and entails a participant strengthening the weights of traversed edges (Fig. 6A). When an edge is strengthened it becomes more likely that the participant will revisit the nodes connected by the reinforced edge. Higher values of reinforcement indicate greater strengthening of traversed edges. The second growth rule is *regularity*. Regularity indicates the willingness of the agent to take short versus long topological steps. Higher values of regularity indicate a relatively greater preference for taking shorter topological steps (Fig. 6B). The effect of reinforcement and regularity on knowledge network growth of simulated participants is illustrated in Supplemental Movie 1 and described in Fig. 6.

To determine the roles of reinforcement and regularity on observed knowledge network growth, we fit the generative model to each participant's network separately. Mean reinforcement is 39.55 (*SD*=6.69). Intuitively, hunter-like network building is indicated by higher values of reinforcement, suggesting that participants return to previously visited concepts in order to fill information gaps. Indeed, values of reinforcement are positively correlated with both average node degree (r=0.17, p=0.04) and the number of traversed loops (r=0.18, p=0.03), which are indices that reflect a practice of revisiting network nodes, in the observed directed knowledge networks. Busybody-like network building, in contrast, is indicated by lower values of reinforcement, and lesser tendency to return to previously visited concepts. We used multiple regression analysis to test if deprivation sensitivity is associated with reinforcement while controlling for the other four facets of curiosity as well as network density and network size. The predictors as a group do not explain a significant percent of the variance in reinforcement, R^2=0.09, F(7,141)=1.97, p = 0.06 (Table S9). Deprivation sensitivity is positively associated with reinforcement (b=1.36, p=0.01, β=0.24; Fig. 6C). This association indicates that participants with high values of deprivation sensitivity have a greater tendency to return to previously visited concepts during knowledge network building. Joyous exploration is negatively associated with reinforcement (b=-1.47, p=0.04, β=-0.23). No other facets of curiosity are associated with reinforcement. Notably, the associations are significant when controlling for network density (b=5.13, p=0.46, β=0.06) and network size (b=0.01, p=0.05, β=0.17).

In addition to reinforcement, the regularity term of the generative model constitutes the preference for taking shorter versus longer topological steps during information seeking. Intuitively, network building that is more busybody-like is indicated by smaller regularity values and the tendency to take relatively long topological steps along the knowledge network. In contrast, hunter-like network building is indicated by larger regularity values and the tendency to take shorter topological steps, potentially in an effort to sample closely related concepts. We used multiple regression analysis to test if deprivation sensitivity is associated with regularity while controlling for the other four facets of curiosity and network strength. The predictors explain a significant percent of the variance in the regularity (R^2=0.10, F(7,141)=2.26, p = 0.03; Table S10). No facets of curiosity, including deprivation sensitivity, are associated with regularity (all p-values greater than 0.05). Although not significantly associated with facets of curiosity, the mean regularity is 2.11 (*SD*=0.15), approaching 2 and therefore suggesting that the information seeking character of the sample is consistent with Lévy-like dynamics[51–53]. A Lévy flight is a specialized random walk expressed as fractal movement patterns, occurring when the distribution of distances traversed with discrete movements falls in a power-law distribution with an exponent of 2, as observed in the current data (Fig. 6B).

Variability in hunter and busybody styles

In a final step, we partitioned the time series of Wikipedia browsing data into thirds to create early, middle, and late information seeking knowledge networks in order to examine the extent to which participants exhibit variability in their styles of information

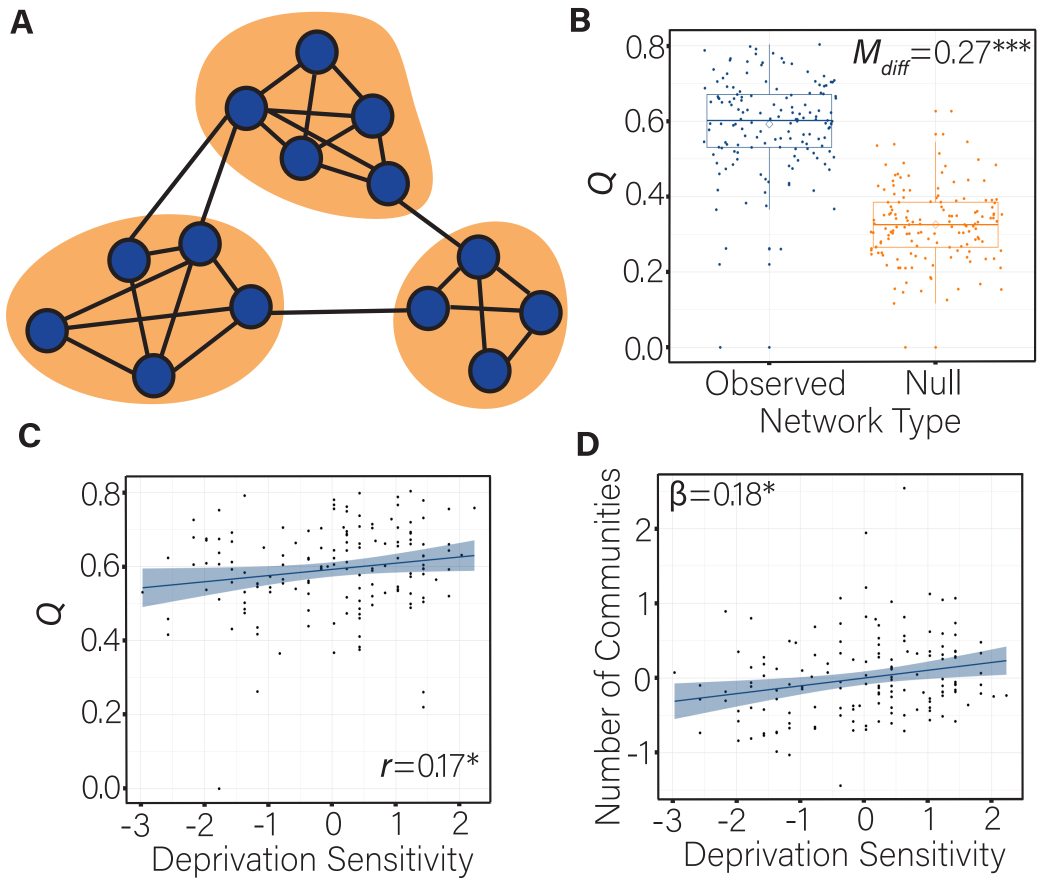

Figure 5. Mesoscale structure of knowledge networks. (*A*) Networks can be decomposed into communities, the nodes of which exhibit dense connections with each other and sparse connections with other network nodes. Three communities are indicated in orange. The modularity quality, Q, of all knowledge networks (Observed) and of degree distribution preserving randomized networks (Null) is shown in panel (*B*). The boxplots represent the median, the interquartile range, and the largest and smallest values within 1.5 times the interquartile range above and below the 75th and 25th percentiles. A paired samples t-test indicates that the observed networks have higher Q relative to randomized networks. The positive correlation between deprivation sensitivity and Q is shown in panel *C*. The ribbon around the line of best fit represents the standard error. A partial residual plot from a regression analysis (Table S8) indicates that deprivation sensitivity is positively associated with the number of communities within knowledge networks (*D*). The ribbon around the line of best fit represents the standard error. *Note*: *$p<0.05$; ***$p<0.001$; *SD*=standard deviation; *M*=mean; M_{diff}=mean difference between Q of observed and null networks; β=standardized regression coefficient.

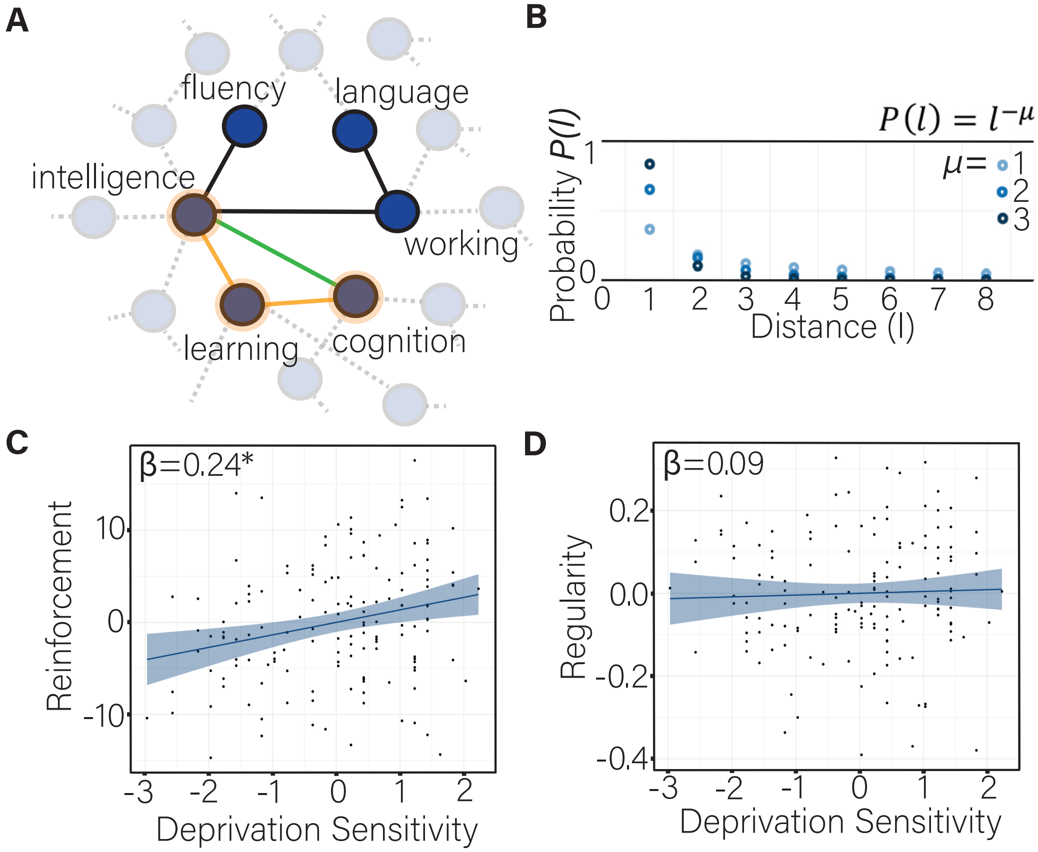

Figure 6. **Generative model and associations with deprivation sensitivity.** Our generative model of knowledge network growth consists of two growth rules. Reinforcement is the value that an agent places on similar and previously sought information while traversing the network. Regularity reflects the preference for taking short versus long topological steps while exploring the network. During a random walk (*A*), an agent draws length l=2 from a Pareto distribution with μ=1 (*B*), meaning that the agent will traverse two edges. Starting at the node i "intelligence" at time t=0, the probability of visiting neighbors $P^t(i,j)$ is given by $\frac{w^t_{ij}}{\sum_l w^t_{il}}$ where w is the edge weight. At time t=1, the agent walks to nodes j "learning" and k "cognition" (orange edges, *A*). The edge between initial and target nodes E(i,k), in green, gets strengthened by the reinforcement value δw. For participants with higher values of reinforcement, the edge is reinforced to a relatively high extent, leading to a greater likelihood of returning to previously visited concepts and resulting in tight networks characteristic of the hunter. High regularity values are associated with a preference to take shorter topological leaps when walking on the knowledge network. This tendency is shown in panel (*B*) where three values of regularity (μ) are shown along with their associated probabilities of making jumps of distance l. Participants with high regularity values (μ=3, dark blue, *B*) have a higher probability of taking steps of distance 1 relative to participants with low values (μ=1 or 2, lighter blues, *B*) and a lower probability of taking jumps of distances greater than 1. High regularity values, then, would result in tight networks akin to the hunter. See Movie S1 for dynamic illustrations. Partial residual plots from regression analyses indicate that deprivation sensitivity is positively associated with reinforcement (*C*). The ribbon around the line of best fit represents the standard error. We observe no association between deprivation sensitivity and regularity (*D*). The ribbon around the line of best fit represents the standard error. *Note*: ; β=standardized regression coefficient.

seeking across time. Intraclass correlations indicate that 35% of the variance in the average edge weight, 26% of the variance in average clustering coefficient, 64% of the variance in characteristic path length, 30% of variance in Q, and 17% of the variance in number of communities is due to between-person variance. Thus, a substantial amount of the variance in network metrics across early, middle, and late information seeking stages is due to within-person fluctuations.

We hypothesized that fluctuations in participants' sensation seeking tendencies – their preferences for novel and exciting experiences – would be associated with the tightness of their knowledge networks, such that periods of high sensation-seeking tendencies would be periods during which looser knowledge networks were created. Repeated measures correlations provide evidence for this hypothesis. The repeated measures correlation between sensation seeking and average edge weight of knowledge networks is significant and negative (r=-0.16, p=0.004; Fig. 7B), indicating that periods of higher sensation seeking are periods in which networks with lower average edge weights are constructed. Periods of higher than usual sensation seeking are also periods in which participants create knowledge networks of lower than usual clustering (r=-0.14, p=0.01, Fig. 7C), and longer than usual characteristic path length (r=0.19, p<0.001, Fig. 7D). No significant repeated measures correlation emerged between sensation seeking and Q (r=0.04, p=0.45), but periods of higher than usual sensation seeking are also periods in which fewer communities are observed in participants' knowledge networks (r=-0.16, p=0.004).

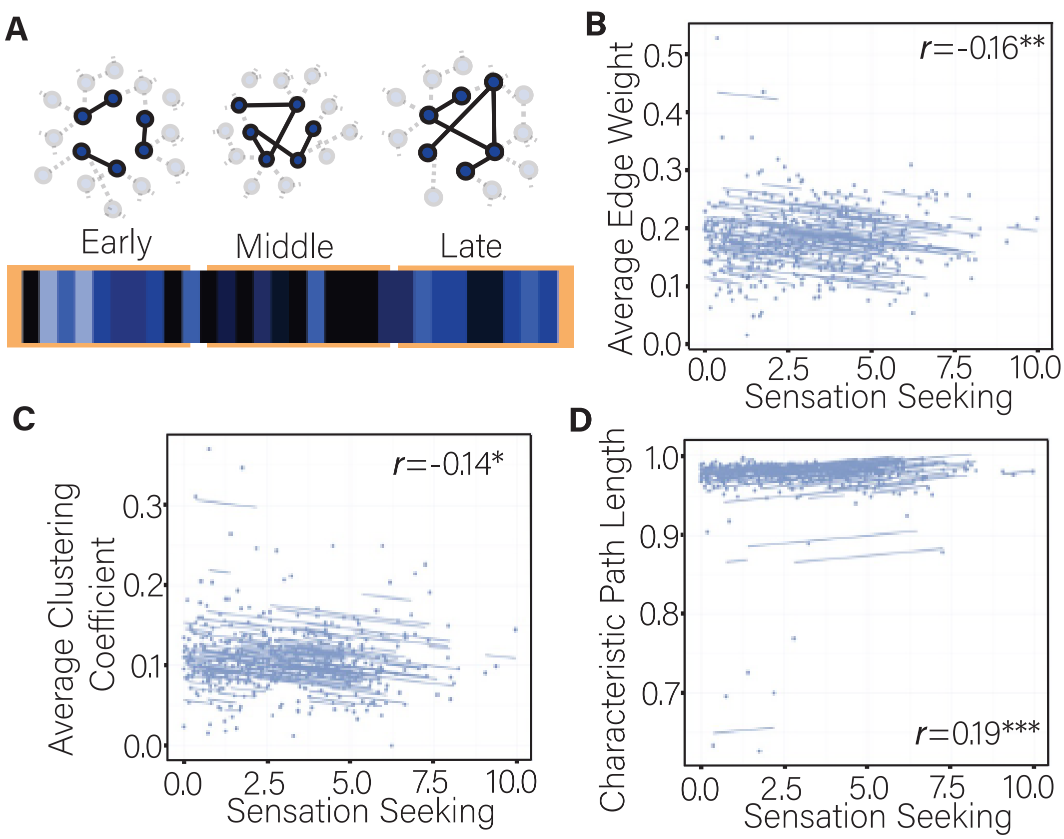

Figure 7. Within-person variability in hunter and busybody styles. We partitioned each participant's time series of edges traversed into early, middle, and late periods to examine within-person fluctuations in the expression of hunter and busybody styles of information seeking (*A*). Repeated measures correlations (estimates in top right corners) indicate that periods of higher than usual sensation seeking as assessed via daily diary are periods during which knowledge networks with lower than usual average edge weights (*B*), lower than usual average clustering coefficients (*C*), and longer than usual characteristic path lengths (*D*) are created. Each dot represents one of three observations for a participant and lines represent the repeated measures correlation fit for each participant (panels *B-D*). *Notes*: ***p<0.001; **p<0.01; *p<0.05.

Robustness and additional analyses

Additional analyses confirm that the results for the association between deprivation sensitivity and average edge weight (Table S11), deprivation sensitivity and clustering coefficient (Table S12), deprivation sensitivity and characteristic path length (Table S13), deprivation sensitivity and number of modules (Table S14), and deprivation sensitivity and reinforcement (Table S15) are robust to the removal of non-significant covariates. Including age and gender in the regression analyses revealed no significant associations between age and gender and average edge weight (Table S16), average clustering coefficient (Table S17), characteristic path length (table S18), number of communities (Table S19), and reinforcement (Table S20). The inclusion of age and gender did not change the observed associations between deprivation sensitivity and any graph metrics.

Discussion

Across species, the propensity to seek out non-instrumental information is notable[4–6]. Curiosity is characterized by this intrinsically motivated information seeking and is strongly associated with well-being due to the many informational and social resources reaped by consistently acting on one's curiosity over time[11–13]. The open-ended and internally driven nature of curiosity makes it notoriously difficult to quantify the information resources that are collected during curious practice, which in turn are theorized to promote well-being on extended timescales. Here, we sought to overcome this challenge by integrating historicophilosophical styles of curious information seeking[18,23] with a knowledge network building approach to curiosity[24] to characterize and quantify the internally driven and idiosyncratic information seeking of individuals under minimal external constraints.

Hunter and busybody styles of curiosity and associations with deprivation sensitivity

By intensively monitoring the information seeking of participants browsing Wikipedia for over five hours throughout the course of 21 days, we constructed networks consisting of the unique Wikipedia pages visited by participants and the semantic similarity between the content of those pages. Transforming the 18654 pages visited by participants into networks, we were able to represent complex information seeking in a manner that could be readily quantified. Knowledge networks exhibited small-world and modular structure. Individual differences in average edge weight, clustering coefficient, and characteristic path length captured general notions of tight and loose knowledge networks, providing an intuitive mapping for hunter and busybody styles of curious practice.

In addition to quantifying qualitative notions of loose and tight knowledge networks that sit at the heart of historicophilosophical styles of curious information search, we examined the potential role of deprivation sensitivity in styles of knowledge network building by incorporating a recently developed self-report, multidimensional curiosity scale[30]. In line with the theoretical notion that individuals high in deprivation sensitivity have a drive to eliminate the unknown as they encounter new information and recognize gaps in their knowledge[32,33,35], we observed that deprivation sensitivity was consistently associated with the three graph theoretical indices used to quantify the tightness and looseness of participants' knowledge networks. More specifically, greater deprivation sensitivity was associated with higher average edge weights, higher clustering coefficients, and shorter characteristic path lengths. Our findings support theoretical propositions regarding deprivation sensitivity and provide new insight into its expression during open-ended, internally directed information seeking.

In examining the association between deprivation sensitivity and knowledge network architecture, we controlled for other facets of curiosity in order to examine aspects of knowledge network building that were specific to deprivation sensitivity. Although treated as a covariate, joyous exploration, a facet of curiosity associated with pure enjoyment of novel stimuli[54], was negatively associated with both average edge weight and reinforcement. This pattern of relations suggests that participants characterized by high motivation to seek new knowledge are more likely to visit relatively dissimilar concepts as they traverse Wikipedia compared to those low in joyous exploration. Few associations between facets of curiosity beyond deprivation sensitivity or joyous exploration and knowledge network indices were observed. Contexts beyond Wikipedia will be better suited to examine how other facets of curiosity are expressed in networks during information seeking in the contexts of uncertainty (stress tolerance), social information (social curiosity), and perceptually intense (thrill seeking) information.

Generative model of curiosity and information seeking

As well as describing the resulting knowledge networks in the language of graph theory, we further formalized the study of hunter and busybody styles of information search by specifying a generative model of network growth. The model consisted of two growth rules with an intuitive mapping to hunter and busybody styles involving the tendency to revisit previously traversed edges (i.e., reinforcement) and individual differences in the propensity to travel across different topological distances at each edge between concepts (i.e., regularity). Reinforcement was associated with deprivation sensitivity. The association underlines the importance of revisiting information in explaining the tendency for participants high in deprivation sensitivity to create tight networks. Reinforcement was negatively associated with joyous exploration, suggesting that individual differences in the enjoyment of novel stimuli is associated with a reduced tendency to return to previously visited concepts. Regularity was not associated with deprivation sensitivity (or any other facet of curiosity), suggesting that the mechanism underlying the relative tightness of the knowledge networks of participants high in deprivation sensitivity is more likely due to the revisiting of similar concepts and less likely due to individual differences in the tendency to take short versus long topological leaps.

An interesting finding from the generative model is that the sample, on average, exhibited a regularity value of 2.11, consistent with a particular type of random walk termed a Lévy flight. A Lévy flight is a specialized random walk expressed as fractal movement patterns, occurring when the distribution of distances traversed with discrete movements falls in a power-law distribution with an exponent of 2, as observed in the current data (Fig. 6B). Fractal movement patterns make Lévy flights particularly apt for efficiently searching for resources embedded in complex environments with hierarchical, lattice, patchy, or heterogeneous organizations[55–57]. Lévy flights have been observed in the movement trajectories of diverse systems, including cells, animals, and humans[58–60]. Although remaining an active area of interdisciplinary research and debate[61–64], observations of Lévy kinesthetics in nature have motivated proposals that evolution selected for cognitive processes that result in efficient Lévy flight exploration[65–68]. Evolutionary adaptations leading to Lévy flight foraging in physical environments may have also been co-opted for the exploration of abstract conceptual spaces[69,70]. Our findings suggest, then, that humans tend to display a specific type of information seeking behavior typically observed during the optimally efficient search for scarce, randomly distributed, and subjectively rewarding information during knowledge network building on Wikipedia. This finding motivates the interpretation of knowledge network exploration during internally directed information seeking under minimal constraints as searching through a conceptual space for subjectively rewarding concepts with an optimally efficient strategy. Future work could directly test this interpretation by administering a similar information seeking task that regularly polls individuals for the subjective value of the sought pieces of information.

Time-varying nature of knowledge network building

We considered styles of knowledge network building and deprivation sensitivity as both traits and states. This approach is in line with recent work indicating substantial within-person variability in curiosity[11]. We find that all indices of knowledge network tightness exhibit substantial within-person variability across time. Thus, while our individual differences analyses indicate variability across persons in the expression of hunter and busybody information seeking styles, these tendencies fluctuate within persons across short (i.e., over the course of days) periods of time. Making use of daily reports of sensation seeking, we find that periods during which looser than usual knowledge networks are created are also periods during which sensation seeking tendencies are higher than usual. This mapping between fluctuations in knowledge network style and sensation seeking tendencies is intuitive given the association between sensation seeking and drives for novel experiences. An important future direction will be to determine if sensation seeking tendencies influence knowledge network building by changing the desire for differing types of information during curious practice. Alternatively, findings may reflect more diverse information seeking spurred on by the more diverse array of activities undertaken prior to Wikipedia exploration during periods of high sensation seeking.

Benefits of the historicophilosophical approach

Our specific interdisciplinary approach has the benefit of broadening and deepening the now classical psychological perspectives on curiosity. Our analysis develops and redirects a long tradition of distinguishing specific and diversive curiosity[71]. Specific curiosity refers to an aroused state experienced when confronted with ambiguous stimuli, leading to specific exploration to obtain depth of knowledge[72]. Diversive curiosity, by contrast, refers to the need to seek new experiences to obtain a breadth of knowledge[73]. These dimensions of curiosity continue to be probed[74]. The strength of this literature lies not only in its careful attention to states vs. traits of curiosity, but to the objects that induce curiosity (e.g., novel perceptual or epistemic stimuli) and the internal impetuses that prompt curiosity (e.g., interest, boredom, conflict, complexity, ambiguity, anxiety, etc.). We build on this literature in two ways. First, by pressing back, across philosophical thought, we can attend to a rich, under-utilized history of curiosity not merely as a state or a trait, but as a panoply of personas and practices. Tracking these transhistorical archetypes across eons of wisdom literatures, we are equipped to appreciate and to test inherited taxonomies of curious practices[75]. Second, by pressing forward, through network science, we can attend more directly to curiosity as an act of connecting, rather than merely acquiring, new pieces of information. Not limiting ourselves to understanding how knowledge is amassed, we utilize this framework to explore the elegant architectures of knowledge network building itself.

Future outlook

The collection of data under few restrictions provides an important match to the internally directed information seeking that lies at the core of contemporary definitions of curiosity[1–3]. We situate hunter and busybody styles of information seeking on a dimension ranging from loose to tight networks. Implicit in this formalism is the notion that individuals practice both forms of information seeking but that each form can be expressed to a differing degree, and that the relative expression of diverse forms of curious practice is an important individual difference[23]. Note that both styles are considered curious practice and that this conceptualization aligns with multidimensional conceptions of curiosity that emphasize not the existence of curious versus incurious people, but individual differences in the way that curiosity is expressed[30]. As such, this analysis may contribute to a broader appreciation of the diverse range of curious practices manifested across the spectrum of neurotypical and neuroatypical learners[76]. A next step will be to examine how tendencies to practice these different styles is reflected in the types of resources (in this case narrow versus wide store of information) that individuals collect over time, which are theorized to impact well-being[11,13]. In future applications of this work, it is conceivable that educators can develop exercises that specifically target either hunter or busybody styles of information seeking, whether because one style is better suited to a given task or because a student will benefit from developing both styles as complementary empowerment skills.

Conclusions

We make use of a novel knowledge network building framework of curiosity to capture and quantify styles of information seeking put forward in a recently developed taxonomy of curious information seeking. We show that individuals' highly idiosyncratic, internally directed information seeking can be represented as knowledge networks and that general notions of tight and loose knowledge networks can be operationalized using graph theoretical indices and growth mechanisms to provide insight into the organizing principles of curiosity-driven exploration. We provide support for a role for deprivation sensitivity in motivating distinct styles of information seeking by finding evidence that individuals high in deprivation sensitivity create tight knowledge networks and exhibit a tendency to return to previously visited concepts.

Methods

We used data from the Knowledge Networks Over Time (KNOT) study, a study designed to provide insight into behavior across a range of domains of functioning, including curiosity[11,41]. All data and code used in the manuscript are available upon request from the corresponding author. Greater detail on the design, data preparation, and analysis can be found in the supplement.

Participants

Our participant sample comprised 149 individuals (121 female, 26 male, 2 other gender) recruited through poster, Facebook, Craigslist, and university research site advertisements in Philadelphia and the surrounding university community, who completed a task that is the focus of the current manuscript, from a full sample of 167 participants on which we have previously reported[11,41]. Participants were aged between 18.21 and 65.24 years (M = 25.05, SD = 6.99), and identified as African American/Black (6.71%), Asian (25.50%), Hispanic/Latino (5.37%), Multiracial (5.37%), other (5.37%), White (49.66%), and missing information (2.01%). Data collection began in October 2017 and ended in July 2018. All research was conducted in accordance with the institutional review board (IRB) at our host university. The IRB board at the university declared the study exempt due to the minimal risk the study posed to participants.

Procedure

Interested participants were sent a baseline survey through Qualtrics containing demographic questionnaires and the curiosity measure. Participants engaged in a laboratory session at which they completed additional questionnaires on Qualtrics, received training in the daily assessment protocol, and were guided through the installation of tracking software (Timing) necessary for a Wikipedia browsing task. Only the participant and one researcher were present during the laboratory visit. Following the laboratory visit, a 21-day diary assessment protocol was initiated. The 21-day diary assessment consisted of two components. The first was a daily diary, delivered using Qualtrics, consisting of survey questionnaires that took approximately 5 minutes to complete. The second came immediately after the daily diary component and was a 15 minute Wikipedia browsing task. Links to the daily assessments were emailed to participants at 6:30 PM each evening and participants completed them outside of the laboratory on their personal computers. The researcher was not blind to study hypotheses during data collection.

Measures

We used participants' reports of demographic information and trait curiosity from the baseline surveys, their ratings of sensation seeking during the 21-day diary, and their daily Wikipedia browsing.

Each evening following the daily diary, participants were prompted to open a browser and to navigate to Wikipedia.org. Participants were instructed to spend 15 minutes in self-directed information seeking on Wikipedia and to explore whatever topics interested them. The choice of Wikipedia, an open content online encyclopedia, reflects its use as a knowledge network in previous research (e.g.,[36]). Following the 15 minutes of open browsing, participants exported and uploaded their browsing history.

Curiosity was measured using the Five Dimensional Curiosity Scale (5D)[30]. The 5D captures multiple dimensions of curiosity that include deprivation sensitivity, joyous exploration, stress tolerance, social curiosity, and thrill seeking. Participants rate the extent to which five items within each subscale accurately describes them on a 0 ("Does not describe me at all") to 6 ("Completely describes me") scale. Reliability (Cronbach's α) of subscales in the current sample were satisfactory.

We measured day's sensation-seeking using 2 items adapted from the Fun-Seeking subscale of the BIS/BAS scales[77] and the Excitement-Seeking subscale of the Revised Neuroticism, Extraversion, and Openness Personality Inventory[78]. Participants were instructed to rate how accurately the statement reflected how they behaved today on a scale from 0 ("None of the time") to 10 ("All of the time") in increments of 0.1.

Data Preparation

To construct knowledge networks for each participant, we created for each individual a list of nodes (unique Wikipedia pages visited) and an edge list that indicated the similarity between each node. To create edge weights, we computed term frequency-inverse document frequency (*tf-idf*) values for the text within each Wikipedia page visited during the study for all participants (n=18654) and calculated the cosine similarity between all pairs of nodes. Higher cosine similarity values indicate greater similarity of the text within each Wikipedia page. Nodes connected by a hyperlink had a larger cosine similarity value (*Mean* = 0.27, *SD*=0.21) than nodes that were not connected by a hyperlink (*Mean* = 0.07, *SD*=0.13), and this difference was statistically significant ($t(26976)=97.85$, $p<0.001$). We chose to define edges by their cosine similarity rather than by the binary hyperlink indicator because the cosine similarity ranges from 0 to 1, providing a more fine-grained measure of concept similarity than is available through binary, hyperlink-defined edges.

Data Analysis

Given the novelty of the information seeking task used in the present study, we undertook thorough descriptive analyses of the structure of participants' knowledge networks. We then used model-based approaches to uncover the mechanisms underlying knowledge network growth. Throughout, we examined associations between knowledge network structure and deprivation sensitivity. See supplementary materials for definitions of network statistics and descriptions of analysis approaches.

References

1. Gottlieb, J., Oudeyer, P. Y., Lopes, M. & Baranes, A. Information-seeking, curiosity, and attention: Computational and neural mechanisms. *Trends Cogn. Sci.* **17**, 585–593, DOI: 10.1016/j.tics.2013.09.001 (2013). NIHMS150003.

2. Gottlieb, J. & Oudeyer, P. Y. Towards a neuroscience of active sampling and curiosity, DOI: 10.1038/s41583-018-0078-0 (2018).

3. Kidd, C. & Hayden, B. Y. The Psychology and Neuroscience of Curiosity. *Neuron* **88**, 449–460, DOI: 10.1016/j.neuron.2015.09.010 (2015). 15334406.

4. Blanchard, T. C., Hayden, B. Y. & Bromberg-Martin, E. S. Orbitofrontal cortex uses distinct codes for different choice attributes in decisions motivated by curiosity. *Neuron* **85**, 602–614, DOI: 10.1016/j.neuron.2014.12.050 (2015).

5. Brydevall, M., Bennett, D., Murawski, C. & Bode, S. The neural encoding of information prediction errors during non-instrumental information seeking. *Sci. Reports* **8**, 1–11, DOI: 10.1038/s41598-018-24566-x (2018).

6. Daddaoua, N., Lopes, M. & Gottlieb, J. Intrinsically motivated oculomotor exploration guided by uncertainty reduction and conditioned reinforcement in non-human primates. *Sci. Reports* **6**, 20202, DOI: 10.1038/srep20202 (2016).

7. Bennett, D., Bode, S., Brydevall, M., Warren, H. & Murawski, C. Intrinsic Valuation of Information in Decision Making under Uncertainty. *PLoS Comput. Biol.* **12**, 1–21, DOI: 10.1371/journal.pcbi.1005020 (2016).

8. Hsee, C. K. & Ruan, B. The Pandora Effect. *Psychol. Sci.* **27**, 659–666, DOI: 10.1177/0956797616631733 (2016).

9. Lau, J. K. L., Ozono, H., Kuratomi, K., Komiya, A. & Murayama, K. Hunger for Knowledge: How the Irresistible Lure of Curiosity is Generated in the Brain. *bioRxiv* 473975, DOI: 10.1101/473975 (2018).

10. Rodriguez Cabrero, J. A., Zhu, J. Q. & Ludvig, E. A. Costly curiosity: People pay a price to resolve an uncertain gamble early. *Behav. Process.* **160**, 20–25, DOI: 10.1016/j.beproc.2018.12.015 (2019).

11. Lydon-Staley, D. M., Zurn, P. & Bassett, D. S. Inconsistent curiosity: Augmentation and blunting of curiosity in daily life and implications for well-being. (2018).

12. Park, S., Kim, M.-S. & Chun, M. M. Concurrent working memory load can facilitate selective attention: evidence for specialized load. *J. experimental psychology. Hum. perception performance* **33**, 1062–75, DOI: 10.1037/0096-1523.33.5.1062 (2007).

13. Peterson, C., Ruch, W., Beermann, U., Park, N. & Seligman, M. E. Strengths of character, orientations to happiness, and life satisfaction. *J. Posit. Psychol.* **2**, 149–156, DOI: 10.1080/17439760701228938 (2007).

14. Fredrickson, B. L. *Positive Emotions Broaden and Build*, vol. 47 (Copyright © 2013, Elsevier Inc. All rights reserved., 2013), 1 edn.

15. Fredrickson, B. L. 05 The Role of Positive Emotions in Positive Psychology. *Am. Psychol. Assoc.* **56**, 218–226, DOI: 10.1O37//0OO3-O66X.56.3.218 (2001). PMC3122271.

16. Kashdan, T. B., Rose, P. & Fincham, F. D. Curiosity and exploration: Facilitating positive subjective experiences and personal growth opportunities. *J. Pers. Assess.* **82**, 291–305, DOI: 10.1207/s15327752jpa8203_05 (2004).

17. Nowotny, H. *Insatiable curiosity: innovation in a fragile future* (MIT Press, 2010).

18. Zurn, P. Busybody, hunter, dancer: three historical models of curiosity. In Papastefanou, M. (ed.) *Toward new philosophical explorations of the epistemic desire to know: Just curious about curiosity*, 27–49 (Cambridge Scholars Press, Cambridge, 2019).

19. Heidegger, M. Being and time, trans. joan stambaugh. *State Univ. New York Press. Albany* (1996).

20. Yonge, C. D. *et al.* *The works of Philo: Complete and unabridged* (Hendrickson Pub, 1993).

21. Helmbold, W. C. *et al.* *Plutarch's Moralia*, vol. 1 (Harvard University Press, 1960).

22. Good, B. Evil, trans. walter kaufmann. *Basic Writings Nietzsche* 1878–1879 (1966).

23. Zurn, P. & Bassett, D. S. On Curiosity: A Fundamental Aspect of Personality, a Practice of Network Growth. *Pers. Neurosci.* **1**, e13, DOI: 10.1017/pen.2018.3 (2018).

24. Bassett, D. S. A network science of the practice of curiosity. In Zurn, P. & Shankar, A. (eds.) *Curiosity studies: Toward a new ecology of knowledge* (University of Minnesota Press, Minneapolis, MN, forthcoming).

25. Baranes, A., Oudeyer, P. Y. & Gottlieb, J. Eye movements reveal epistemic curiosity in human observers. *Vis. Res.* **117**, 81–90, DOI: 10.1016/j.visres.2015.10.009 (2015).

26. Risko, E. F., Anderson, N. C., Lanthier, S. & Kingstone, A. Curious eyes: Individual differences in personality predict eye movement behavior in scene-viewing. *Cognition* **122**, 86–90, DOI: 10.1016/j.cognition.2011.08.014 (2012).

27. West, R. & Leskovec, J. Human wayfinding in information networks. 619, DOI: 10.1145/2187836.2187920 (2012).

28. Gross, J. L. & Yellen, J. *Handbook of graph theory* (CRC press, 2004).

29. Newman, M. *Networks* (Oxford university press, 2018).

30. Kashdan, T. B. *et al.* The five-dimensional curiosity scale: Capturing the bandwidth of curiosity and identifying four unique subgroups of curious people. *J. Res. Pers.* **73**, 130–149, DOI: 10.1016/j.jrp.2017.11.011 (2018).

31. Litman, J. A. Interest and deprivation factors of epistemic curiosity. *Pers. Individ. Differ.* **44**, 1585–1595, DOI: 10.1016/j.paid.2008.01.014 (2008).

32. Litman, J. A. & Mussel, P. Validity of the Interest-and Deprivation-Type Epistemic Curiosity Model in Germany. *J. Individ. Differ.* **34**, 59–68, DOI: 10.1027/1614-0001/a000100 (2013).

33. Litman, J. A. Interest and deprivation factors of epistemic curiosity. *Pers. Individ. Differ.* **44**, 1585–1595, DOI: 10.1016/j.paid.2008.01.014 (2008).

34. Golman, R. & Loewenstein, G. Information gaps: A theory of preferences regarding the presence and absence of information. (2016).

35. Loewenstein, G. The psychology of curiosity: A review and reinterpretation., DOI: 10.1037/0033-2909.116.1.75 (1994).

36. Scholtes, I. When is a Network a Network? Multi-Order Graphical Model Selection in Pathways and Temporal Networks. DOI: 10.1145/3097983.3098145 (2017). 1702.05499.

37. Lemmerich, F., Sáez-Trumper, D., West, R. & Zia, L. Why the world reads wikipedia: Beyond english speakers. In *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining*, 618–626 (ACM, 2019).

38. Singer, P. *et al.* Why we read wikipedia. In *Proceedings of the 26th International Conference on World Wide Web*, 1591–1600 (International World Wide Web Conferences Steering Committee, 2017).

39. Kashdan, T. B. & Steger, M. F. Curiosity and pathways to well-being and meaning in life: Traits, states, and everyday behaviors. *Motiv. Emot.* **31**, 159–173, DOI: 10.1007/s11031-007-9068-7 (2007).

40. Zuckerman, M. *Behavioral expressions and biosocial bases of sensation seeking* (Cambridge university press, 1994).

41. Lydon-Staley, D. & Bassett, D. Within-person variability in sensation-seeking during daily life: Positive associations with alcohol use and self-defined risky behaviors. (2019).

42. Kleiman, E. *EMAtools: Data Management Tools for Real-Time Monitoring/Ecological Momentary Assessment Data* (2017). R package version 0.1.3.

43. Watts, D. J. & Strogatz, S. H. Collective dynamics of 'small-world'networks. *nature* **393**, 440 (1998).

44. Muldoon, S. F., Bridgeford, E. W. & Bassett, D. S. Small-world propensity and weighted brain networks. *Sci. reports* **6**, 22057 (2016).

45. Onnela, J. P., Saramäki, J., Kertész, J. & Kaski, K. Intensity and coherence of motifs in weighted complex networks. *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.* **71**, 1–4, DOI: 10.1103/PhysRevE.71.065103 (2005). 0408629.

46. Latora, V. & Marchiori, M. Efficient behavior of small-world networks. *Phys. review letters* **87**, 198701 (2001).

47. Van Wijk, B. C., Stam, C. J. & Daffertshofer, A. Comparing brain networks of different size and connectivity density using graph theory. *PloS one* **5**, e13701 (2010).

48. Betzel, R. F. & Bassett, D. S. Multi-scale brain networks. *Neuroimage* **160**, 73–83 (2017).

49. Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. *J. statistical mechanics: theory experiment* **2008**, P10008 (2008).

50. Maslov, S. & Sneppen, K. Specificity and stability in topology of protein networks. *Science* **296**, 910–913 (2002).

51. Viswanathan, G. M. *et al.* Optimizing the success of random searches. *nature* **401**, 911 (1999).

52. Viswanathan, G. *et al.* Lévy flights in random searches. *Phys. A: Stat. Mech. its Appl.* **282**, 1–12 (2000).

53. Sims, D. W. *et al.* Scaling laws of marine predator search behaviour. *Nature* **451**, 1098 (2008).

54. Kashdan, T. B. *et al.* The Five-Dimensional Curiosity Scale: Capturing the bandwidth of curiosity and identifying four unique subgroups of curious people. *J. Res. Pers.* DOI: 10.1016/j.jrp.2017.11.011 (2017).

55. Santos, M., Viswanathan, G., Raposo, E. & da Luz, M. Optimization of random searches on regular lattices. *Phys. Rev. E* **72**, 046143 (2005).

56. Wosniack, M., Santos, M., Raposo, E., Viswanathan, G. & Da Luz, M. Robustness of optimal random searches in fragmented environments. *Phys. Rev. E* **91**, 052119 (2015).

57. Raposo, E. P. *et al.* How landscape heterogeneity frames optimal diffusivity in searching processes. *PLoS computational biology* **7**, e1002233 (2011).

58. Harris, T. H. *et al.* Generalized lévy walks and the role of chemokines in migration of effector cd8+ t cells. *Nature* **486**, 545 (2012).

59. Rhee, I. *et al.* On the levy-walk nature of human mobility. *IEEE/ACM transactions on networking (TON)* **19**, 630–643 (2011).

60. Rhodes, T. & Turvey, M. T. Human memory retrieval as lévy foraging. *Phys. A: Stat. Mech. its Appl.* **385**, 255–260 (2007).

61. Pyke, G. H. Understanding movements of organisms: it's time to abandon the lévy foraging hypothesis. *Methods Ecol. Evol.* **6**, 1–16 (2015).

62. Reynolds, A. Liberating lévy walk research from the shackles of optimal foraging. *Phys. life reviews* **14**, 59–83 (2015).

63. Da Luz, M., Raposo, E. & Viswanathan, G. And yet it optimizes. comment on" liberating lévy walk research from the shackles of optimal foraging" by am reynolds. *Phys. life reviews* **14**, 94–98 (2015).

64. Gallotti, R., Bazzani, A., Rambaldi, S. & Barthelemy, M. A stochastic model of randomly accelerated walkers for human mobility. *Nat. communications* **7**, 12600 (2016).

65. Bartumeus, F. Lévy processes in animal movement: an evolutionary hypothesis. *Fractals* **15**, 151–162 (2007).

66. Humphries, N. E. *et al.* Environmental context explains lévy and brownian movement patterns of marine predators. *Nature* **465**, 1066 (2010).

67. Viswanathan, G. M., Da Luz, M. G., Raposo, E. P. & Stanley, H. E. *The physics of foraging: an introduction to random searches and biological encounters* (Cambridge University Press, 2011).

68. Wosniack, M. E., Santos, M. C., Raposo, E. P., Viswanathan, G. M. & da Luz, M. G. The evolutionary origins of lévy walk foraging. *PLoS computational biology* **13**, e1005774 (2017).

69. Hills, T. T. Animal foraging and the evolution of goal-directed cognition. *Cogn. science* **30**, 3–41 (2006).

70. Todd, P. M., Hills, T. T. & Robbins, T. W. *Cognitive search: Evolution, algorithms, and the brain*, vol. 9 (MIT press, 2012).

71. Berlyne, D. E. Conflict, arousal, and curiosity. (1960).

72. Day, H. An instrument for the measurement of intrinsic motivation. *An interim report to Dep. Manpow. Immigr.* (1969).

73. Leherissey, B. L. The development of a measure of state epistemic curiosity. (1971).

74. Litman, J. A. & Spielberger, C. D. Measuring epistemic curiosity and its diversive and specific components. *J. personality assessment* **80**, 75–86 (2003).

75. Zurn, P. & Bassett, D. S. Philosophy of biology: Seizing an opportunity. *eLife* **8**, e48336 (2019).

76. Johnson, K. T. Neurodiversity, autism, and curiosity. In Zurn, P. & Shankar, A. (eds.) *Curiosity studies: A new ecology of knowledge* (University of Minnesota Press, Minneapolis, MN, forthcoming).

77. Carver, C. S. & White, T. L. Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the bis/bas scales. *J. personality social psychology* **67**, 319 (1994).

78. Costa, P. T. & McCrae, R. R. Normal personality assessment in clinical practice: The neo personality inventory. *Psychol. assessment* **4**, 5 (1992).

Acknowledgements

We thank Jordan Dworkin, Christopher W. Lynn, and Shubhankar Patankar for feedback on earlier versions of the manuscript. D.S.B., D.M.L, D.Z., and A.S.B acknowledge support from the John D. and Catherine T. MacArthur Foundation, the Alfred P. Sloan Foundation, the ISI Foundation, the Paul Allen Foundation, the Army Research Laboratory (W911NF-10-2-0022), the Army Research Office (Bassett-W911NF-14-1-0679, Grafton-W911NF-16-1-0474, DCIST- W911NF-17-2-0181), the Office of Naval Research, the National Institute of Mental Health (2-R01-DC-009209-11, R01 – MH112847, R01-MH107235, R21-M MH-106799), the National Institute of Child Health and Human Development (1R01HD086888-01), National Institute of Neurological Disorders and Stroke (R01 NS099348), the National Science Foundation (BCS-1441502, BCS-1430087, NSF PHY-1554488 and BCS-1631550), and the National Institute on Drug Abuse (1K01DA047417-01A1). All authors acknowledge support from the Center for Curiosity. The content is solely the responsibility of the authors and does not necessarily represent the official views of any of the funding agencies.

Author contributions statement

D.M.L. designed the research with input from D.Z., P.Z., and D.S.B; D.M.L., A.S.B., and D.Z. analyzed the data; D.M.L. wrote the paper. A.S.B., D.Z., P.Z., and D.S.B. edited the paper.

Data and code availability

All data and code used in the manuscript are available upon request from the corresponding author.