Published in final edited form as:
Nat Hum Behav. 2021 March ; 5(3): 327–336. doi:10.1038/s41562-020-00985-7.

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

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

[a]Annenberg School for Communication, University of Pennsylvania, Philadelphia, PA 19104 USA

[b]Department of Bioengineering, School of Engineering & Applied Science, University of Pennsylvania, Philadelphia, PA 19104 USA

[c]Department of Neuroscience, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104 USA

[d]Department of Philosophy, American University, Washington, DC 20016 USA

[e]Department of Electrical & Systems Engineering, School of Engineering & Applied Science, University of Pennsylvania, Philadelphia, PA 19104 USA

[f]Department of Neurology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104 USA

[g]Department of Physics & Astronomy, College of Arts & Sciences, University of Pennsylvania, Philadelphia, PA 19104 USA

[h]Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104 USA

[i]Santa Fe Institute, Santa Fe, NM 87501 USA

Abstract

The open-ended and internally driven nature of curiosity makes characterizing the information seeking that accompanies it a daunting endeavor. We use an historicophilosophical taxonomy of information seeking coupled with a knowledge network building framework to capture styles of information seeking in 149 participants as they explore Wikipedia for over 5 hours spanning 21 days. We create knowledge networks in which nodes represent distinct concepts and edges represent the similarity between concepts. We quantify the tightness of knowledge networks using

* dsb@seas.upenn.edu.

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.

Competing interests

The authors declare no competing interests.

Data availability statement

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

Code availability statement

Analyses in the manuscript used code available through R, MATLAB, and the Brain Connectivity Toolbox. Code associated with the generative model is available at: https://github.com/dalejn/kinestheticCuriosity.

graph theoretical indices and use a generative model of network growth to explore mechanisms underlying information seeking. Deprivation curiosity, 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. With this framework in hand, future research can readily quantify the information seeking associated with curiosity.

Introduction

Curiosity is characterized by intrinsically motivated information seeking[1–3]. The information sought while acting on one's curiosity often has no immediate, tangible benefit[4–6]. Despite a lack of immediate benefits, the tendency to frequently experience curiosity is associated with positive well-being[7–9]; curiosity facilitates 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[10–12]. And irrespective of its immediate or potential utility, curiosity may well be valuable in itself[13].

Characterizing how individuals seek information when internally driven is fundamental to understanding how curiosity leads to the shoring up of resources that impact well-being. Historicophilosophical studies tracing the use of the word curiosity have identified styles of information seeking that span millennia, cultures, and languages[14]. 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"[15] (p.161). The busybody will "frisk about, and rove about, at random, wherever they please"[16] (sec. 34). The information seeking of the hunter is characterized by sampling closely connected concepts. The hunter does not "turn aside and follow every scent"[17] (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"[18] (p.59) in a targeted information search. Both styles are considered expressions of curiosity, but there are individual differences in the extent to which each style is expressed[19]. Tendencies to exhibit one style over another will lead to the accumulation of different types of 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 diverse information seeking styles a daunting endeavor. Existing approaches include the examination of saccadic exploration of visual scenes and responses to trivia questions designed to evoke curiosity[20,21]. Experimental paradigms are shedding 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 knowledge network building framework[22]. From this perspective, network nodes represent distinct concepts and network edges represent the manner in which the concepts are related. While seeking information, an individual traverses edges on knowledge networks, moving 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[23] and allows an application of the mathematical language of graph theory[24,25] to quantify complex manifestations of curious behavior. The easily distracted busybody will create 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 hunter, in contrast, will create tight networks consisting of closely-connected concepts and their networks will have large edge weights, high clustering, and low characteristic path length (Fig. 1).

To determine how trait curiosity manifests in knowledge network building, we can measure the associations between network structure and existing curiosity measurement instruments. Deprivation curiosity refers to the extent to which one's information seeking is motivated to overcome the feeling of being deprived of knowledge[26]. Deprivation curiosity has emerged as a robust individual difference[27,33], with several valid and reliable measures available[28,30,33,71]. Individuals high in deprivation curiosity possess a strong drive to know, and seek information that eliminates gaps in their knowledge[29,30]. There is an element of compulsiveness to deprivation curiosity. Deprivation curiosity does not motivate individuals to learn new things just for fun. Instead, when individuals decide that they are missing information required to better understand a concept, they experience a feeling of deprivation[31]. A determination to continue information seeking until a knowledge gap is filled results in a persistent and effortful form of specific exploration that resolves an unknown[32,33].

This compulsiveness may reflect motivational mechanisms of incentive-salience[31,34]. Incentive salience refers to a motivational feeling of 'wanting' in anticipation of an outcome that is separate from the hedonic response of 'liking'[35,36]. Mesolimbic dopamine activity is implicated in 'wanting' and motivates approach behaviors, often in the form of response perseveration and pursuit of a stimulus until consumption and satiation[37,38]. There is evidence that 'wanting' plays a role in deprivation curiosity, with the expected acquisition of knowledge associated with activation in the dopaminergic reward system[39–41]. Activation associated with deprivation curiosity shares activity implicated in incentive salience in the processing of other rewards (e.g., food, money)[39]. Due to the targeted and effortful nature of information seeking associated with deprivation seeking and its likely association with 'wanting', we hypothesize that individuals high in deprivation curiosity will create tighter networks as they encounter new information, recognize gaps in their knowledge, and search for closely related concepts in an iterative cycle of filling in knowledge gaps[26,42].

Here, we operationalize curiosity as a knowledge network building practice. We monitor the 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.,[43]), its use in existing work to examine information seeking[44,45], and findings that intrinsic learning forms a major motivation for use of Wikipedia[46,47]. We treat each Wikipedia page as a distinct concept or node in a knowledge network, and we quantify the semantic similarity between any two pages to create network edges. We use graph theory to quantify general notions of tight and loose networks to realize busybody and hunter styles of information seeking. We uncover potential mechanisms underlying knowledge network growth by developing a generative model of information search. We examine associations between knowledge networks and trait deprivation curiosity, with the hypothesis that individuals high in deprivation curiosity will create relatively tight, hunter-like knowledge networks. We examine the extent to which the tightness of knowledge networks changes over time. Curiosity exhibits fluctuations over relatively short time-scales (e.g., from day to day[7,48]) and hunter and busybody styles of knowledge network building are thought to be expressed to different degrees across time within-person[22]. We hypothesize that loose knowledge networks, reflecting the pursuit of novel, diverse, and varied information, will be created during periods of heightened sensation seeking tendencies that promote the "seeking of varied, novel, complex, and intense sensations and experiences" (p.26[49];[50]).

Results

We operationalize busybody and hunter styles of information seeking in a specific instance of knowledge network building. Using the 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 and Fig. 2B). 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 curiosity. In all models, we include four other facets of curiosity (joyous exploration, social curiosity, thrill seeking, and stress tolerance[30,33]) as covariates to examine the extent to which the hypothesized associations are specific to deprivation curiosity, rather than being driven by other facets of curiosity. 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 Methods. Descriptive statistics of key variables can be found in Supplementary Table 1. Two-tailed tests were used throughout. Throughout, b indicates unstandardized regression coefficients and β indicates standardized regression coefficients.

Deprivation curiosity 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 curiosity is positively associated with average edge weight (b=0.004, 95% CI=[0.001,0.007], p=0.01; Cohen's d = 0.44; a moderate effect size[51], Supplementary Table 2), indicating that participants high in deprivation curiosity are more hunter-like in their knowledge network building relative to participants low in deprivation curiosity, who in contrast are more busybody-like in their information seeking (Fig. 2D).

Deprivation curiosity 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[52]. 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 networks is 0.09 (SD=0.02). To conceptually link this metric to knowledge network building, 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, 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[53]. The mean characteristic path length in these networks is 0.99 (SD=0.03). To conceptually link this metric to knowledge network building 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 used multiple regression analysis to test the extent to which deprivation curiosity is associated with average clustering coefficient, while controlling for the other four facets of curiosity as well as network density and size, which are known confounds in network studies[54] (Fig. 3A). 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;

Supplementary Table 3). Deprivation curiosity is positively associated with the average clustering coefficient (b=0.003, 95% CI=[0.001, 0.006], p=0.01, β=0.23; a small effect size; Fig. 3B), suggesting that participants high in deprivation curiosity examine closely related concepts during information seeking to a greater extent than participants low in deprivation curiosity.

We next regressed the characteristic path length on deprivation curiosity while controlling for the other four facets of curiosity as well as both network density and size (Fig. 3C). 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, $p < 0.001$ (Supplementary Table 4). Deprivation curiosity is negatively associated with the characteristic path length (b=−0.001, 95% CI=[−0.001, −0.0001], p=0.02, β=−0.10; a small effect size; Fig. 3D) such that participants high in deprivation curiosity, while exploiting local information, also have networks that are easily traversable from one end to the next.

Principles of knowledge network growth and associations with curiosity

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 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. 4A). 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. 4B). The effect of reinforcement and regularity on knowledge network growth of simulated participants is illustrated in Supplementary Movie 1 and described in Figure 4.

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. 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 curiosity 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 (Supplementary Table 5). Deprivation curiosity is positively associated with reinforcement (b=1.36, 95% CI=[0.28, 2.44], p=0.01, β=0.24; a small effect size; Fig. 4C). This association indicates that participants with high values of deprivation curiosity have a greater tendency to return to previously visited concepts during knowledge network building.

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 curiosity 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; Supplementary Table 6). Deprivation curiosity is not significantly associated with regularity (b=0.01, 95% CI=[−0.01, 0.04], p=0.35, β=0.09). 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[55–57]. 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. 4B).

Variability in hunter and busybody styles

In order to examine the extent to which participants exhibit variability in their styles of information seeking across time, we partitioned the time series of Wikipedia browsing data into thirds to create early, middle, and late information seeking knowledge networks. Intraclass correlations indicate that 35% of the variance in the average edge weight, 26% of the variance in average clustering coefficient, and 64% of the variance in characteristic path length 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(297)=−0.16, 95% CI=[−0.27,−0.05], p=0.004; a small effect size, Fig. 5B), 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

($r(297)=-0.14$, 95% CI=[−0.25, −0.03], $p=0.01$, a small effect size; Fig. 5C), and longer than usual characteristic path length ($r(297)=0.19$, 95% CI=[0.08, 0.30], $p<0.001$; a small effect size; Fig. 5D).

Robustness and additional analyses

Additional analyses confirm that the results for the association between deprivation curiosity and average edge weight (Supplementary Table 7), clustering coefficient (Supplementary Table 8), characteristic path length (Supplementary Table 9), and reinforcement (Supplementary Table 10) are robust to the removal of non-significant covariates. We also note that after adjusting the alpha rate to control for the testing of deprivation curiosity's association with the tightness of knowledge networks across three different indices (Bonferroni correction such that the alpha rate is 0.05/3=0.02), the associations remain significant. Our correction for deprivation sensitivity specifically and not all dimensions of curiosity reflects the manuscript's focus on deprivation sensitivity. Use of a multiple regression rather than a multilevel model to test the association between deprivation curiosity and average edge weight also revealed the hypothesized association (Supplementary Table 11, Supplementary Figure 1). In Supplementary Methods and Supplementary Results, we test two additional models of network growth[58], the preferential attachment and preferential acquisition models, but find no evidence that these models accurately describe knowledge network growth. Additional analyses to characterize the structure of knowledge networks indicate that knowledge networks have modular and small-world structure (see Supplementary Methods and Supplementary Results).

Discussion

Curiosity is characterized by intrinsically motivated information seeking and is strongly associated with well-being due to the many informational resources reaped by consistently acting on one's curiosity over time[7–9]. The open-ended and internally driven nature of curiosity makes it difficult to quantify the resources that are collected during information seeking, which in turn are theorized to promote well-being on extended timescales. Here, we overcome this challenge by integrating historicophilosophical styles of curious information seeking[14,19] with a knowledge network building approach to curiosity[22] to characterize and quantify the internally driven and idiosyncratic information seeking of individuals under minimal external constraints.

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 knowledge network building practice.

As well as describing the resulting knowledge networks, we formalized 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 the propensity to travel across different topological distances at each edge between concepts (i.e., regularity). 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. 4B). Fractal movement patterns make Lévy flights particularly apt for efficiently searching for resources embedded in complex environments with hierarchical, lattice, patchy, or heterogeneous organizations[59–61]. Lévy flights have been observed in the movement trajectories of diverse systems, including cells, animals, and humans[62–64]. 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]. Findings from the generative model suggest, then, that humans display a 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.

In addition to quantifying qualitative notions of loose and tight knowledge networks, we examined the role of deprivation curiosity in styles of knowledge network building. In line with the notion that individuals high in deprivation curiosity have a drive to eliminate the unknown as they encounter new information and recognize gaps in their knowledge[26,27,33], deprivation curiosity was consistently associated with three indices used to quantify the tightness and looseness of participants' knowledge networks. Greater deprivation curiosity was associated with higher average edge weights, higher clustering coefficients, and shorter characteristic path lengths. Our findings support propositions regarding deprivation curiosity and provide new insight into its expression during open-ended information seeking. When considering the findings from the generative model, reinforcement was associated with deprivation curiosity. The association underlines the importance of revisiting information in explaining the tendency for participants high in deprivation curiosity to create tight networks. Regularity was not statistically significantly associated with deprivation curiosity, suggesting that the mechanism underlying the relative tightness of the knowledge networks of participants high in deprivation curiosity 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.

In examining the association between deprivation curiosity and knowledge network architecture, we controlled for other facets of curiosity. Although treated as a covariate to examine the independent association between deprivation curiosity and knowledge network structure, joyous exploration, a facet of curiosity associated with pure enjoyment of novel stimuli[30], was negatively associated with both average edge weight and reinforcement. This pattern of associations 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, and that they exhibit little tendency to return to previously visited concepts. Joyous exploration lends itself to carefree, causal information seeking and is positively associated with ambiguity tolerance[32,71], in line with the association with loose networks observed in the present study. However, we caution that joyous exploration was not significantly associated with average clustering coefficient or path length, that few consistent associations between facets of curiosity beyond deprivation curiosity and knowledge network indices were observed, and that facets beyond deprivation curiosity were treated as covariates in the current study which motivated our corrections for multiple comparisons. 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.

We considered styles of knowledge network building and deprivation curiosity as both traits and states[7]. 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 time. 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. 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.

Our 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[72]. Specific curiosity refers to an aroused state experienced when confronted with ambiguous stimuli, leading to specific exploration to obtain depth of knowledge[73]. Diversive curiosity, by contrast, refers to the need to seek new experiences to obtain a breadth of knowledge[74]. These dimensions of curiosity continue to be probed[75]. The strength of this literature lies not only in its 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 curiosity[76]. 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.

It is important to interpret the study findings in the context of the study's limitations. The collection of data under few restrictions provides a more ecologically valid design relative to existing experimental laboratory paradigms to capture the internally directed information seeking that lies at the core of contemporary definitions of curiosity[1–3]. Yet, participants received incentives for completing the 15-minutes of Wikipedia browsing, raising the possibility that participants browsed to obtain the incentives rather than to satisfy their deprivation curiosity. However, the incentives were provided to encourage continued participation in the study protocol and were not contingent on seeking out information on specific topics or in particular ways. The motives behind information seeking over extended timescales are undoubtedly manifold. Yet, consistently observing associations between deprivation curiosity across three indices of knowledge network tightness suggests that a desire to satisfy one's deprivation curiosity was one motive behind the information seeking observed. Even though the observed associations are modest, it is notable that the associations can be detected in the real world in the midst of all the noise that exists outside of the laboratory.

We situate hunter and busybody styles of information seeking on a dimension ranging from loose to tight networks using continuous variables (e.g., average edge weight, clustering coefficient, and path length). Implicit in this formalism and our use of continuous variables 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 curiosity is an important individual difference[19]. By dichotomizing the data, we would fail to capture the extent of variation in hunter and busybody styles across both individuals and time. Note that both hunter and busybody styles are considered expressions of curiosity, aligning 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 framework may contribute to a broader appreciation of the diverse range of information seeking manifested across the spectrum of neurotypical and neuroatypical learners[77]. A next step will be to examine how tendencies to practice different styles are 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[7,9].

We focused on hunter and busybody styles of knowledge network building. Historicophilosophical studies have identified a third archetype known as the dancer[14]. The dancer experiments and breaks with traditional pathways of investigation, taking leaps of creative imagination and, in the process, produces new concepts and radically remodels knowledge networks. Paradigms beyond information seeking will be necessary to capture the work of the dancer due to the centrality of creation to the dancer's definition. Creativity paradigms, that capture searches in semantic memory networks or analyses of the structure of creative works themselves, would lend themselves to capturing the dancer. Although the source of the data and the processes underlying it would differ drastically from the current information seeking paradigm, the network based approach taken here may generalize and capture the knowledge network building and construction of the dancer. Indeed, network approaches have previously been applied to capture differences in network structures of semantic associations observed in people with varying levels of creativity[78,79].

In summary, we use a knowledge network building framework to capture and quantify styles of information seeking put forward in a historicophilosophical taxonomy of curious information seeking. Individuals' highly idiosyncratic, internally directed information seeking can be represented as knowledge networks and 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 curiosity in motivating distinct styles of information seeking by finding evidence that individuals high in deprivation curiosity 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[7,50]. 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 Supplementary Methods. All research was conducted in accordance with the Human Subjects Electronic Research Application institutional review board (IRB) at the University of Pennsylvania. The IRB board at the university declared the study exempt due to the minimal risk the study posed to participants. All participants provided informed consent before taking part in the study.

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[7,50]. No statistical methods were used to pre-determine sample sizes but our sample sizes are similar to those reported in previous 21-day intensive longitudinal studies[80]. 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.

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. Participants were compensated with Amazon gift cards at each study phase: $25 after completing the baseline assessment and the laboratory visit. For the daily assessment, completion was incentivized by making participant payment contingent on completion: completion of 3, 4, 5, 6, and 7 surveys each week was compensated with gift cards worth $10, $15, $20, $25, and $35, respectively. Continued participation through the daily assessment was further incentivized by using a raffle for which an iPad mini was the prize. Completion of all seven surveys each week resulted in one entry into the raffle drawing.

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. Specifically, during the laboratory visit, the investigator stated: "We would like you to open a new tab on your browser and visit https://www.wikipedia.org/. We would like you to spend 15 minutes each evening reading about whatever you want on Wikipedia. For example, if you wanted to learn more about Philadelphia, you could go to the Philadelphia Wikipedia page". At this point the researcher used the Wikipedia search bar to navigate to https://en.wikipedia.org/wiki/Philadelphia to ensure that all participants had familiarity with Wikipedia and its usage. "You can read through the page. You can also click on links you find interesting or you can use the search bar to search for new topics. There is no right or wrong way to do this. We are interested in what it is that people read about when they are not forced to read about anything in particular." We developed this set of instructions so as to ensure that people would browse according to their curiosity, and not in any particular manner suggested by the experimenter. 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. Note that the deprivation sensitivity subscale is highly similar to another commonly used deprivation curiosity scale[33].

We measured day's sensation-seeking using 2 items adapted from the Fun-Seeking subscale of the BIS/BAS scales[81] and the Excitement-Seeking subscale of the Revised Neuroticism, Extraversion, and Openness Personality Inventory[82]. 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. The term frequency for a document is given as

$$tf(t, d) = f_{t,d}, \tag{1}$$

where the term frequency of token t in document d is given by the frequency f with which it appears in the document. The inverse document frequency (*idf*) for a token is defined as:

$$idf(t, D) = \log\frac{|D|}{|\{d \in D: t \in d\}|}), \tag{2}$$

where the inverse document frequency of token t in the set of documents (the study corpus of all 18654 Wikipedia pages) D is given by the log of the number of documents in the set of documents D divided by the number of documents d in the set of documents D that contain the token t. The *tf-idf* is a product of the token's frequency and the token's inverse document frequency. Thus, common tokens appearing very frequently in the corpus will be down-weighted while rare terms will be associated with a relatively large number. To account for differences in document length, we applied a common normalization such that the Euclidean norm of the *tf-idf* vector for a document became 1. After calculating the normalized *tf-idf* for each token, we quantified the similarity between pairs of nodes by computing the cosine similarity between all possible pairs of the 18654 vectors. The cosine similarity resulted in a quantification of node similarity ranging from 0 to 1, with higher values indicating greater similarity of the text within each Wikipedia page.

We chose to define edges by their cosine similarity rather than by the binary hyperlink indicator for two main reasons. First, participants were not constrained to using hyperlinks to navigate Wikipedia and had the option to use the search bar, for example. Second, there were cases in which the text between two Wikipedia pages indicated high similarity between concepts as reflected in the cosine similarity, and there was face validity that two pages were similar to one another, yet no hyperlink was present between the pages. For example, the cosine similarity of the edge between the Mazda RX-8 and the Mazda RX-7 Wikipedia pages (both sports cars manufactured by Japanese automobile manufacturer Mazda) was 0.90 indicating high similarity yet no hyperlink existed between the pages. This lack of hyperlink between two pages with similar text may reflect the coarse, binary nature of hyperlinks relative to the more fine-grained measure of concept similarity available through the cosine similarity ranging between 0 to 1. It may also reflect the fact that hyperlink generation depends on users' subjective assessments of concept similarity rather than a text-based approach. Although there are differences between the cosine similarity and hyperlink approaches, 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 with a large effect size ($t(26976)=97.85$, $p<0.001$, Cohen's $d = 1.10$). Thus, there is some overlap between networks with edges defined by cosine similarity and hyperlinks. We constructed knowledge networks using Wikipedia hyperlinks rather than cosine similarity and repeated the main text analyses in Supplementary Methods and Supplementary Results to allow comparison. We find that the cosine similarity approach, but not the hyperlink approach, is sensitive to individual differences in deprivation curiosity.

Data Analysis

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 curiosity. See Supplementary Methods for definitions of network statistics and descriptions of analysis approaches. Data distribution was assumed to be normal but this was not formally tested.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

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). All authors acknowledge support from the Center for Curiosity. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

References

1. Gottlieb J, Oudeyer PY, Lopes M. & Baranes A. Information-seeking, curiosity, and attention: Computational and neural mechanisms. Trends Cogn. Sci. 17, 585–593 (2013). [PubMed: 24126129]
2. Gottlieb J. & Oudeyer PY Towards a neuroscience of active sampling and curiosity. Nat. Rev. Neurosci. 19, 758–770 (2018). [PubMed: 30397322]
3. Kidd C. & Hayden BY The Psychology and Neuroscience of Curiosity. Neuron 88, 449–460 (2015). [PubMed: 26539887]
4. Blanchard TC, Hayden BY & Bromberg-Martin ES Orbitofrontal cortex uses distinct codes for different choice attributes in decisions motivated by curiosity. Neuron 85, 602–614 (2015). [PubMed: 25619657]

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 (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 (2016).
7. Lydon-Staley DM, Zurn P. & Bassett DS Within-person variability in curiosity during daily life and associations with well-being. J. Pers. 88, 625–641 (2020). [PubMed: 31519052]
8. Park S, Kim M-S & Chun MM Concurrent working memory load can facilitate selective attention: evidence for specialized load. J. Exp. Psychol. Hum. Percept. Perform. 33, 1062–75 (2007). [PubMed: 17924807]
9. Peterson C, Ruch W, Beermann U, Park N. & Seligman ME Strengths of character, orientations to happiness, and life satisfaction. J. Posit. Psychol. 2, 149–156 (2007).
10. Fredrickson BL Positive Emotions Broaden and Build. In Advances in Experimental Social Psychology (eds. Plant EA, Devine PG) 1–53 (Academic Press, 2013).
11. Fredrickson BL The Role of Positive Emotions in Positive Psychology. Am. Psychol. 56, 218–226 (2001). [PubMed: 11315248]
12. Kashdan TB, Rose P. & Fincham FD Curiosity and exploration: Facilitating positive subjective experiences and personal growth opportunities. J. Pers. Assess. 82, 291–305 (2004). [PubMed: 15151805]
13. Nowotny H. Insatiable curiosity: innovation in a fragile future (MIT Press, 2010).
14. Zurn P. Busybody, hunter, dancer: three historical models of curiosity. In Toward new philosophical explorations of the epistemic desire to know: Just curious about curiosity (ed. Papastefanou M) 27–49 (Cambridge Scholars Press, 2019).
15. Heidegger M. Being and time (trans. Stambaugh J)(State University of New York Press, 1996).
16. Yonge CD et al. The works of Philo: Complete and unabridged (Hendrickson Pub, 1993).
17. Helmbold WC et al. Plutarch’s Moralia, vol. 1 (Harvard University Press, 1960).
18. Nietzsche F. Beyond Good and Evil (trans. Kaufmann W) (Vintage Books, 1996).
19. Zurn P. & Bassett DS On Curiosity: A Fundamental Aspect of Personality, a Practice of Network Growth. Pers. Neurosci. 1, e13 (2018).
20. Baranes A, Oudeyer PY & Gottlieb J. Eye movements reveal epistemic curiosity in human observers. Vis. Res. 117, 81–90 (2015). [PubMed: 26518743]
21. Risko EF, Anderson NC, Lanthier S. & Kingstone A. Curious eyes: Individual differences in personality predict eye movement behavior in scene-viewing. Cognition 122, 86–90 (2012). [PubMed: 21983424]
22. Bassett DS A network science of the practice of curiosity. Curiosity studies: Toward a new ecology of knowledge (eds. Zurn P. Shankar A) 57–74 (University of Minnesota Press, 2019).
23. West R. & Leskovec J. Human wayfinding in information networks. Proc. 21st international conference on World Wide Web 117, 619–628 (2012).
24. Gross JL & Yellen J. Handbook of graph theory (CRC press, 2004).
25. Newman M. Networks (Oxford university press, 2018).
26. Loewenstein G. The psychology of curiosity: A review and reinterpretation. Psychol. Bull. 116, 75–98 (1994).
27. Litman JA & Mussel P. Validity of the Interest-and Deprivation-Type Epistemic Curiosity Model in Germany. J. Individ. Differ. 34, 59–68 (2013).
28. Litman JA & Jimerson TL The Measurement of Curiosity As a Feeling of Deprivation. J. Pers. Assess. 82, 147–157 (2004). [PubMed: 15041521]
29. Golman R. & Loewenstein G. Information gaps: A theory of preferences regarding the presence and absence of information. Decision 5, 143–164 (2018).
30. Kashdan TB 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 (2018).
31. Litman J. Curiosity: Nature, dimensionality, and determinants. In The Cambridge handbook on motivation and learning (eds. Renniger KA, Hidi SE) 418–422 (Cambridge University Press, 2019).

32. Lauriola M. et al. Epistemic curiosity and self-regulation. Pers. Individ. Differ. 83, 202–207 (2015).
33. Litman JA Interest and deprivation factors of epistemic curiosity. Pers. Individ. Differ. 44, 1585–1595 (2008).
34. FitzGibbon L, Lau JKL & Murayama K. The seductive lure of curiosity: Information as a motivationally salient reward. Curr. Opin. Behav. Sci. 35, 21–27 (2020).
35. Berridge KC From prediction error to incentive salience: mesolimbic computation of reward motivation. Eur. J. Neurosci. 35, 1124–1143 (2012). [PubMed: 22487042]
36. Berridge KC The debate over dopamine's role in reward: the case for incentive salience. Psychopharmacology 191, 391–431 (2007). [PubMed: 17072591]
37. Robinson TE & Berridge KC Incentive-sensitization and drug 'wanting'. Psychopharmacology 171, 352–353 (2004).
38. Berridge KC & Robinson TE What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? Brain Res. Rev. 28, 309–369 (1998). [PubMed: 9858756]
39. Lau JKL, Ozono H, Kuratomi K, Komiya A. & Murayama K. Shared striatal activity in decisions to satisfy curiosity and hunger at the risk of electric shocks. Nat. Hum. Behav. 1–13 (2020). [PubMed: 31965067]
40. Kobayashi K. & Hsu M. Common neural code for reward and information value. Proc. Natl. Acad. Sci. 116, 13061–13066 (2019). [PubMed: 31186358]
41. Gruber MJ, Gelman BD & Ranganath C. States of Curiosity Modulate Hippocampus-Dependent Learning via the Dopaminergic Circuit. Neuron 84, 486–496 (2014). [PubMed: 25284006]
42. Shin DD & Kim S.-i. Homo curious: Curious or interested? Educ. Psychol. Rev. 31, 853–874 (2019).
43. Scholtes I. When is a Network a Network? Multi-Order Graphical Model Selection in Pathways and Temporal Networks. Preprint at arXiv: https://arxiv.org/abs/1702.05499 (2017).
44. West R. & Leskovec J. Automatic versus human navigation in information networks. In Proceedings of the 6th international AAAI conference on weblogs and social media, 362–369 (AAAI, 2012).
45. Lamprecht D, Lerman K, Helic D. & Strohmaier M. How the structure of wikipedia articles influences user navigation. New Rev. Hypermedia Multimed. 23, 29–50 (2017). [PubMed: 28670171]
46. 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).
47. 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).
48. Kashdan TB & Steger MF Curiosity and pathways to well-being and meaning in life: Traits, states, and everyday behaviors. Motiv. Emot. 31, 159–173 (2007).
49. Zuckerman M. Behavioral expressions and biosocial bases of sensation seeking (Cambridge University Press, 1994).
50. Lydon-Staley D. & Bassett D. Within-person variability in sensation-seeking during daily life: Positive associations with alcohol use and self-defined risky behaviors. Psychol. Addict. Behav. 34, 257–268 (2020). [PubMed: 31815502]
51. Kleiman E. EMAtools: Data Management Tools for Real-Time Monitoring/Ecological Momentary Assessment Data (2017). R package version 0.1.3.
52. Onnela JP, 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 (2005).
53. Latora V. & Marchiori M. Efficient behavior of small-world networks. Phys. review letters 87, 198701 (2001).
54. Van Wijk BC, Stam CJ & Daffertshofer A. Comparing brain networks of different size and connectivity density using graph theory. PloS one 5, e13701 (2010).

55. Viswanathan GM et al. Optimizing the success of random searches. Nature 401, 911–914 (1999). [PubMed: 10553906]
56. Viswanathan G. et al. Lévy flights in random searches. Phys. A: Stat. Mech. its Appl 282, 1–12 (2000).
57. Sims DW et al. Scaling laws of marine predator search behaviour. Nature 451, 1098–1102 (2008). [PubMed: 18305542]
58. Hills TT, Maouene M, Maouene J, Sheya A. & Smith L. Longitudinal analysis of early semantic networks: Preferential attachment or preferential acquisition? Psychol. science 20, 729–739 (2009).
59. Santos M, Viswanathan G, Raposo E. & da Luz M. Optimization of random searches on regular lattices. Phys. Rev. E 72, 046143 (2005).
60. 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).
61. Raposo EP et al. How landscape heterogeneity frames optimal diffusivity in searching processes. PLoS Comput. Biol. 7, e1002233 (2011).
62. Harris TH et al. Generalized lévy walks and the role of chemokines in migration of effector cd8+ t cells. Nature 486, 545–548 (2012). [PubMed: 22722867]
63. Rhee I. et al. On the levy-walk nature of human mobility. IEEE/ACM transactions on networking (TON) 19, 630–643 (2011).
64. Rhodes T. & Turvey MT Human memory retrieval as lévy foraging. Phys. A: Stat. Mech. its Appl 385, 255–260 (2007).
65. Bartumeus F. Lévy processes in animal movement: an evolutionary hypothesis. Fractals 15, 151–162 (2007).
66. Humphries NE et al. Environmental context explains lévy and brownian movement patterns of marine predators. Nature 465, 1066–1069 (2010). [PubMed: 20531470]
67. Viswanathan GM, Da Luz MG, Raposo EP & Stanley HE The physics of foraging: an introduction to random searches and biological encounters (Cambridge University Press, 2011).
68. Wosniack ME, Santos MC, Raposo EP, Viswanathan GM & da Luz MG The evolutionary origins of lévy walk foraging. PLoS Comput. Biol. 13, e1005774 (2017).
69. Hills TT Animal foraging and the evolution of goal-directed cognition. Cogn. Sci. 30, 3–41 (2006). [PubMed: 21702807]
70. Todd PM, Hills TT & Robbins TW Cognitive search: Evolution, algorithms, and the brain, vol. 9 (MIT press, 2012).
71. Litman JA Relationships between measures of I- and D-type curiosity, ambiguity tolerance, and need for closure: An initial test of the wanting-liking model of information-seeking. Pers. Individ. Differ. 48, 397–402, (2010).
72. Berlyne DE Conflict, arousal, and curiosity. (McGraw-Hill,1960).
73. Day H. An instrument for the measurement of intrinsic motivation. An interim report to Dep. Manpow. Immigr. (1969).
74. Leherissey BL The development of a measure of state epistemic curiosity. (ERIC, 1971).
75. Litman JA & Spielberger CD Measuring epistemic curiosity and its diversive and specific components. J. Pers. Assess. 80, 75–86 (2003). [PubMed: 12584070]
76. Zurn P. & Bassett DS Philosophy of biology: Seizing an opportunity. eLife 8, e48336 (2019).
77. Johnson KT Neurodiversity, autism, and curiosity. Curiosity studies: Toward a new ecology of knowledge (eds. Zurn P. Shankar A) 129–146 (University of Minnesota Press, 2020).
78. Kenett YN, Anaki D. & Faust M. Investigating the structure of semantic networks in low and high creative persons. Front. Hum. Neurosci. 8, 407 (2014). [PubMed: 24959129]
79. Kenett YN & Faust M. A semantic network cartography of the creative mind. Trends Cogn. Sci. 23, 271–274 (2019). [PubMed: 30803872]
80. Fosco GM & Lydon-Staley DM A within-family examination of interparental conflict, cognitive appraisals, and adolescent mood and well-being. Child Dev. 90, e421–e436 (2019). [PubMed: 29171662]

Author Manuscript

Author Manuscript

Author Manuscript

81. Carver CS & White TL Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the bis/bas scales. J. Pers. Soc. Psychol. 67, 319–333 (1994).

82. Costa PT & McCrae RR Normal personality assessment in clinical practice: The neo personality inventory. Psychol. Assess. 4, 5–13 (1992).

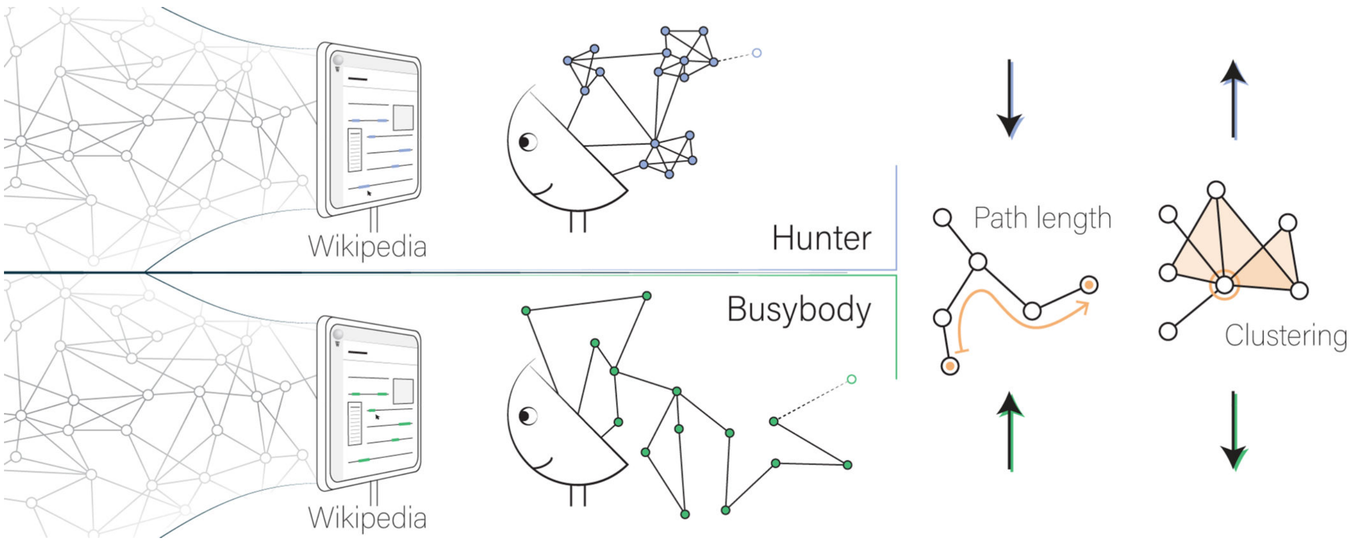

Figure 1. Hunter and busybody styles of information seeking.

Participants explored Wikipedia for 15 minutes every day for 21 days. We represent participants' information seeking as knowledge networks[22]. 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[14] 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.

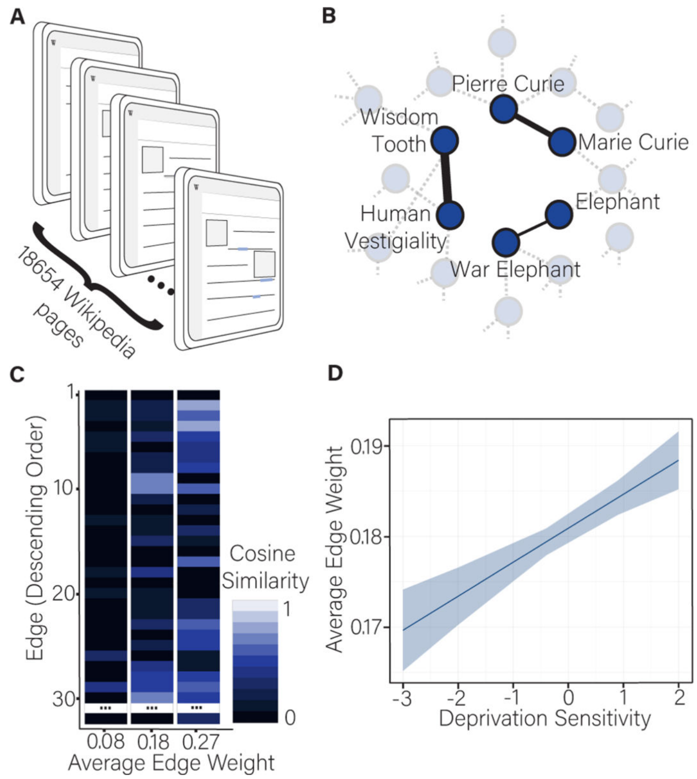

Figure 2. Knowledge network construction and the association between deprivation curiosity and edge weight.

(*A*) The sample (n=149) 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 using 27967 observations nested in 149 participants show that participants high in deprivation curiosity had higher average edge weights, indicating that they tended to visit similar concepts as they traversed Wikipedia (b=0.004, 95% CI=[0.001,0.007], p=0.01, Cohen's d = 0.44; a moderate effect size[51]; Supplementary Table 2). The ribbon around the model estimated association represents the standard error.

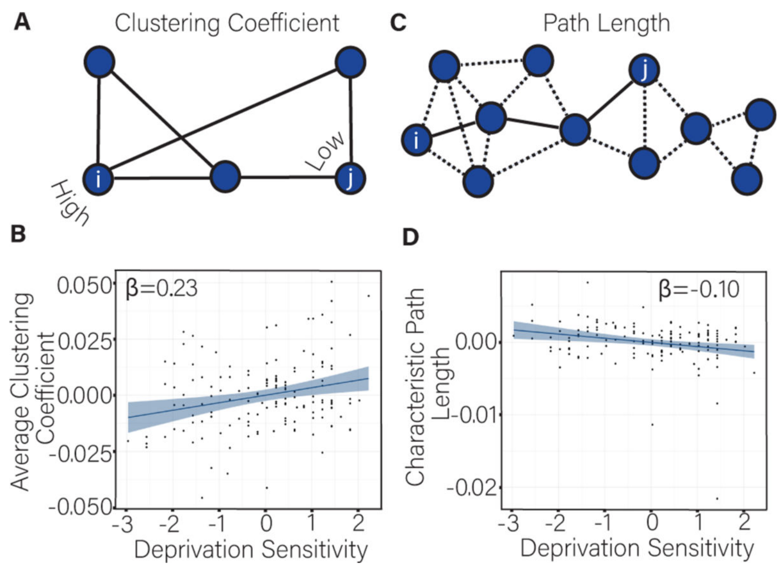

Figure 3. Deprivation curiosity and the clustering and path length of knowledge networks.

(*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 from a regression analysis with 148 participants shows that deprivation curiosity is positively associated with the average clustering coefficient (b=0.003, p=0.01, β=0.23, 95% CI=[0.001, 0.006]). 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 from a regression analysis with 148 participants shows that deprivation curiosity was negatively associated with the characteristic path length (b=−0.001, p=0.02, β=−0.10; a small effect size; 95% CI=[−0.001, −0.0001]). 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*: β=standardized regression coefficient.

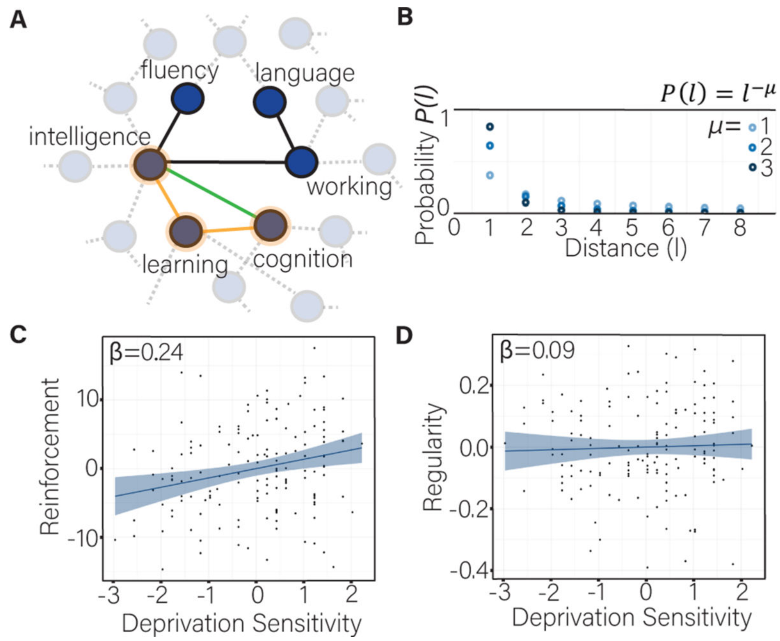

Figure 4. Generative model and associations with deprivation curiosity.

Our generative model of knowledge network growth consists of two growth rules that capture individual differences in how people seek information on Wikipedia. During a random walk (*A*), an individual starting at the node "intelligence" draws length $l=2$ from a Pareto distribution with $\mu=1$ (*B*), meaning that the agent will traverse two edges. In this case, they travel from "intelligence" to "learning" (edge 1) and then from "learning" to "cognition (edge 2) indicated by the orange lines. Between-person differences in the preference for taking short versus long steps (operationalized as the number of edges traversed) is captured by our first growth rule, regularity. High regularity values are associated with a preference for taking shorter steps 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 ($\mu=3$, dark blue, *B*) have a higher probability of taking steps of distance 1 relative to participants with low values ($\mu=1$ 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. Our second growth rule, reinforcement, can be described by considering why the individual visited "learning" after "intelligence" and what happens to the network after the participant travels from "intelligence" to "cognition". The probability of visiting "learning" versus other neighbors of "intelligence" (e.g., "fluency" or "working") is related to the weight of the edges between "intelligence" and its neighboring nodes. Edges with high weights, indicating greater similarity in concepts, have a higher probability of being traversed than those with lower weights. After walking to the nodes "learning" and "cognition" (orange edges, *A*), the edge between the initial ("intelligence") and target ("cognition") nodes, in green, gets strengthened by the reinforcement value δw. In other words, the weight of the edge increases. Reinforcement is used to capture individual differences in the tendency to seek similar and previously sought information while traversing Wikipedia by determining how much the weight of an edge should increase. For participants with high values of reinforcement, the green edge in panel *A* would be reinforced to a relatively high extent, leading to a greater likelihood of the individual returning to previously visited concepts and resulting in tight networks characteristic of the hunter. See Supplementary Video 1 for dynamic illustrations. Partial residual plots from a regression analysis with 149 participants indicate that deprivation curiosity is positively associated with reinforcement (*C*; b=1.36, 95% CI=[0.28, 2.44], p=0.01, β=0.24; a small effect size). The ribbon around the line of best fit represents the standard error. We observe no statistically significant association between deprivation curiosity and regularity (*D*; b=0.01, 95% CI=[−0.01, 0.04], p=0.35, β=0.09) in a regression analysis with 149 participants. The ribbon around the line of best fit represents the standard error. *Note*: β=standardized regression coefficient.

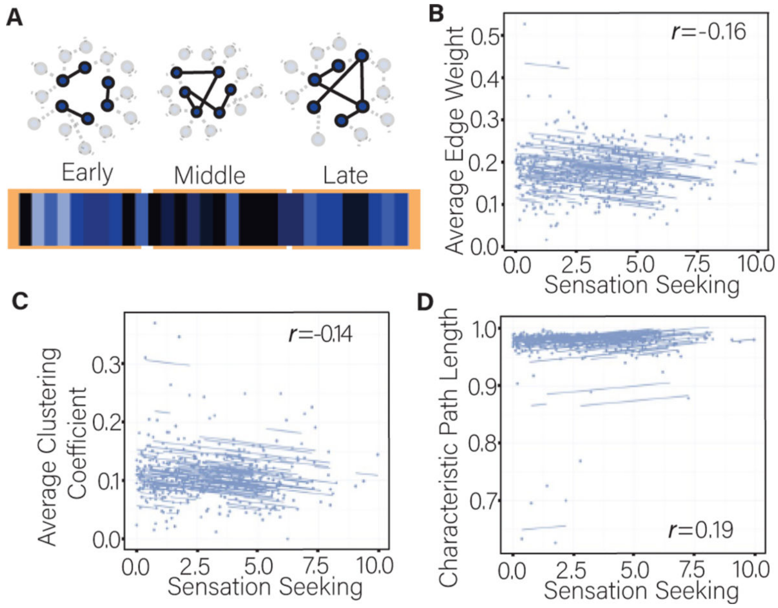

Figure 5. Within-person variability in hunter and busybody styles.

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*; $r(297)=-0.16$, 95% CI=[−0.27,−0.05], $p=0.004$), lower than usual average clustering coefficients (*C*; $r(297)=-0.14$, 95% CI=[−0.25, −0.03], $p=0.01$), and longer than usual characteristic path lengths (*D*; $r(297)=0.19$, 95% CI=[0.08, 0.30], $p<0.001$) are created. Each dot represents one of three observations for a participant ($n=149$) and lines represent the repeated measures correlation fit for each participant (panels *B*-*D*).