Network structure explains intellectual discourse across human history

Cody Moser[1,2,✉], **Alyssa Ortega**[1,3], **and Tyler Marghetis**[1,✉]

[1]Cognitive and Information Sciences, University of California, Merced
[2]School of Collective Intelligence, Université Mohammed VI Polytechnique
[3]Psychology, University of Wisconsin Madison

The production of knowledge is a collective endeavor. Scientific discovery, for instance, reflects not only the insights of individual scientists but also interactions among scientists. This is often taken to reflect social influences on collective epistemic vitality, the capacity to generate and synthesize new ideas. However, in science, it is difficult to separate intellectual influence from access to material resources, such as equipment and grant funding, since networks of collaboration influence the circulation of both intellectual and non-intellectual resources. Here, as a strict test of how social structure shapes intellectual discourse, we use the three-thousand-year history of a human debate in communities that relied on intellectual argumentation (rather than, say, empirical experiments). Building on the work of historians and sociologists, we digitized and quantified the time-evolving network structure of interaction among intellectuals (*N* = 3187), broadly construed, from religious debate in ancient India (c. 800 BCE) to 20th century debates about the logical foundations of mathematics in Europe and North America. We find that the production or preservation of knowledge by a community is explained by its network structure but not with overall levels of antagonism, suggesting that how communities are organized matters more for intellectual progress than how contentious they are. Extending tools from collective intelligence to intellectual history, we call for an integration of the science of science, the philosophy of science, and the history of ideas to forge a comprehensive understanding of the social dynamics of knowledge.

Collective Intelligence | Computational History | Social Networks | Science of Science

Correspondence: ***cmoserj@gmail.com and tyler.marghetis@gmail.com***
Department of Cognitive and Information Sciences, University of California, Merced, CA 95343, USA

Introduction

The production of knowledge is often a social endeavor (1–7). It follows that the dynamics of knowledge production may depend not only on individual genius but on the distributed interpersonal interactions that give rise to collective thinking. Within science, for instance, the collective dynamics of discovery are predicted by the structure of social interactions among scientists, as captured by co-authorship (8–14). While this is often taken as evidence for social influences on the dynamics of *ideas*, scientific collaborations also shape access to *non-intellectual* resources like equipment and grant funding, which can shape scientific productivity (15–19). This confounding of intellectual and non-intellectual exchange has made it difficult to isolate the influence of social structure on the dynamics of collective knowledge.

Here, to investigate the relationship between social structure and collective knowledge, we leverage a three-thousand-year record of intellectual traditions that foreground systematic reasoning and argumentation (Fig. 1). From ancient times to the modern era, in societies across the globe, communities of thinkers have engaged in sustain discourse with their own influential thinkers, intellectual foci, and debates (20). Here, we focus on abstract argumentation by specialized intellectuals — forms of knowledge production that encompass not only debates in contemporary departments of philosophy, mathematics, and theology, but also a much wider range of historical debates: metaphysical and logical disputes among early Buddhist and Jain thinkers in India, scholastic theology in Medieval Europe and the Islamic world, debates over virtue and governance in Confucian and Legalist schools in China, controversies surrounding the formal foundations of mathematics and logic in the modern era, and the natural philosophy that gave rise to modern science during the European scientific revolution (21). These traditions represent some of the oldest forms of collective knowledge production, in some cases predating the emergence of modern institutionalized science by thousands of years.

While these intellectual communities cover much of the ambit of human thought, they share in common an engagement with abstract ideas through a process of argumentation. They vary, however, in their epistemic outcomes — that is, the way they preserved or transformed knowledge. Throughout history, some intellectual communities are characterized by the emergence and synthesis of new ideas, effectively "changing the conversation" in a way analogous to disruptive scientific discoveries (12). We refer to this as epistemic vitality. Other intellectual communities primarily focus on taking existing ideas and preserving, interpreting, or elaborating upon them, thus acting as a system of collective memory. We refer to this as epistemic stasis. Both vitality and stasis are valuable — one generating and integrating new ideas, the other preserving established ones — analogous to the distinction between paradigm-shifting "revolutionary" science and paradigm-preserving "normal" science (3). A central question, then, is what explains why some communities become more epistemically vital, while others tend toward stasis. One possibility is that these collective epistemic dynamics reflect the underlying social structure of a community.

Historically, several frameworks have emphasized the social dynamics of intellectual thought. The 19th century the-

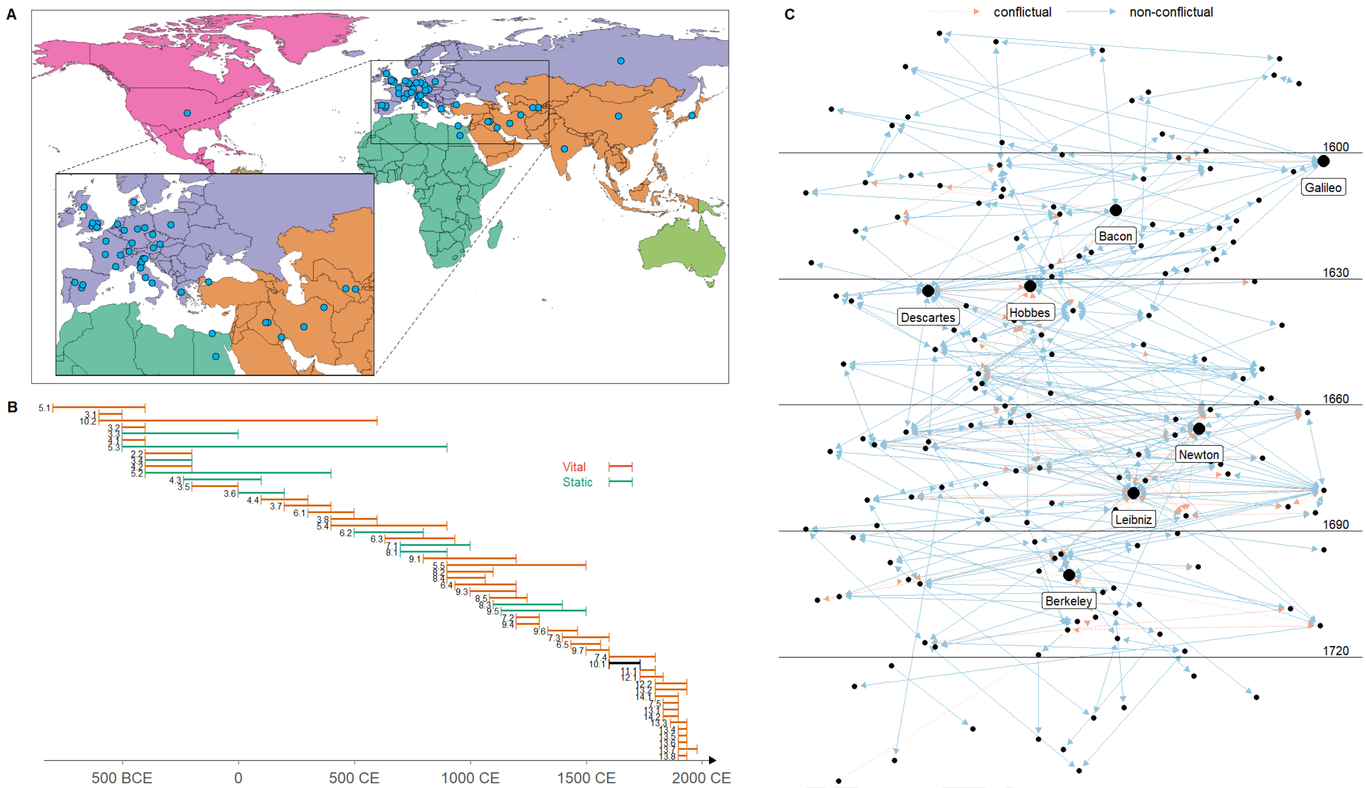

Fig. 1. Global intellectual communities throughout history. **(A)** All geographic locations of all intellectual networks included in the study. Inset shows the region of Europe and the Middle East with the highest density of intellectual communities, although there are communities from throughout Africa, America, Asia, and Europe. Map created using `ggmap` (22). **(B)** The timeline of each intellectual community, with lines indicating their start and end dates. Numbers next to the lines correspond to figures in (21); colors indicate the vitality of each community, except for the black line that designates the network visualized in panel C; the location and identity of each numbered network is available in the Supplemental Materials. **(C)** An example of one network, described in (21) as, "The European Network: The Cascade of Circles," which lasted from 1600 to 1735 CE. Nodes represent individual intellectuals (key figures labeled); edges represent intellectual relationships (solid blue = non-conflictual; dotted orange = conflictual). Each network was divided into non-overlapping temporal periods of equal duration (indicated by horizontal black lines, with years on the right), following the temporal divisions introduced in Randall Collins's visual representations of the history of intellectual networks (21).

ologian and historian John Henry Newman (23) likened epistemic systems to living organisms, gaining maturity through sustained internal debate. The 20th century philosopher of science Thomas Kuhn (3) described intellectual traditions as paradigm-bound, periodically restructured when crises spur the adoption of new frameworks.. At the turn of the current century, the sociologist Randall Collins (21) argued that intellectual traditions are structured by competition for collective attention among individual thinkers and their apprenticeship lineages. Thus, while knowledge generation is sometimes caricatured as the solitary pursuit of understanding, individuals are nearly always operating within larger intellectual communities.

It remains unclear, however, which features of a community are responsible for its epistemic vitality. One family of views holds that vitality arises from individual-level tension, disagreement, or diversity (24–26). According to these accounts, the most productive communities are those marked by persistent disagreement (27) or by a high degree of individual autonomy, where thinkers are relatively isolated and free to pursue independent lines of inquiry (28, 29). Others emphasize the placement of individuals within the larger social network, such as centralizing disruptors (12, 30) or intellectual outsiders in the periphery of the social network (31, 32). Still other accounts highlight more system-level factors such as the balance between tight-knit collaboration and intellectual distance (21, 33). Productive communities, for instance, have been proposed to consist of a small number of highly integrated groups that generate intellectual energy through competition and debate (21). Synthesizing these perspectives, recent work emphasizes the dual importance of innovation within independent sub-communities and the role of 'bridging nodes,' individuals or institutions that connect otherwise disconnected subgroups (12, 33–35). According to these accounts, epistemically vital communities should exhibit structural convergence, in which different subgroups are brought together. These proposals are not mutually-exclusive, but they highlight different structural features that may contribute in parallel to a community's epistemic vitality.

Answering these questions requires knowledge of an epistemic community's social network. Purely theoretical debates rely on the exchange, critique, and collaborative exploration of ideas. In science, social networks can be inferred from patterns of co-authorship (7, 11, 12, 14, 36). For intellectual communities throughout history, however, quantifying the network of intellectual exchange is more difficult. Co-authorship was unheard of in many intellectual communities (37), even though intellectuals engage deeply with each other through social interaction (21). Thus, while many intellectual traditions share with science and other epistemic

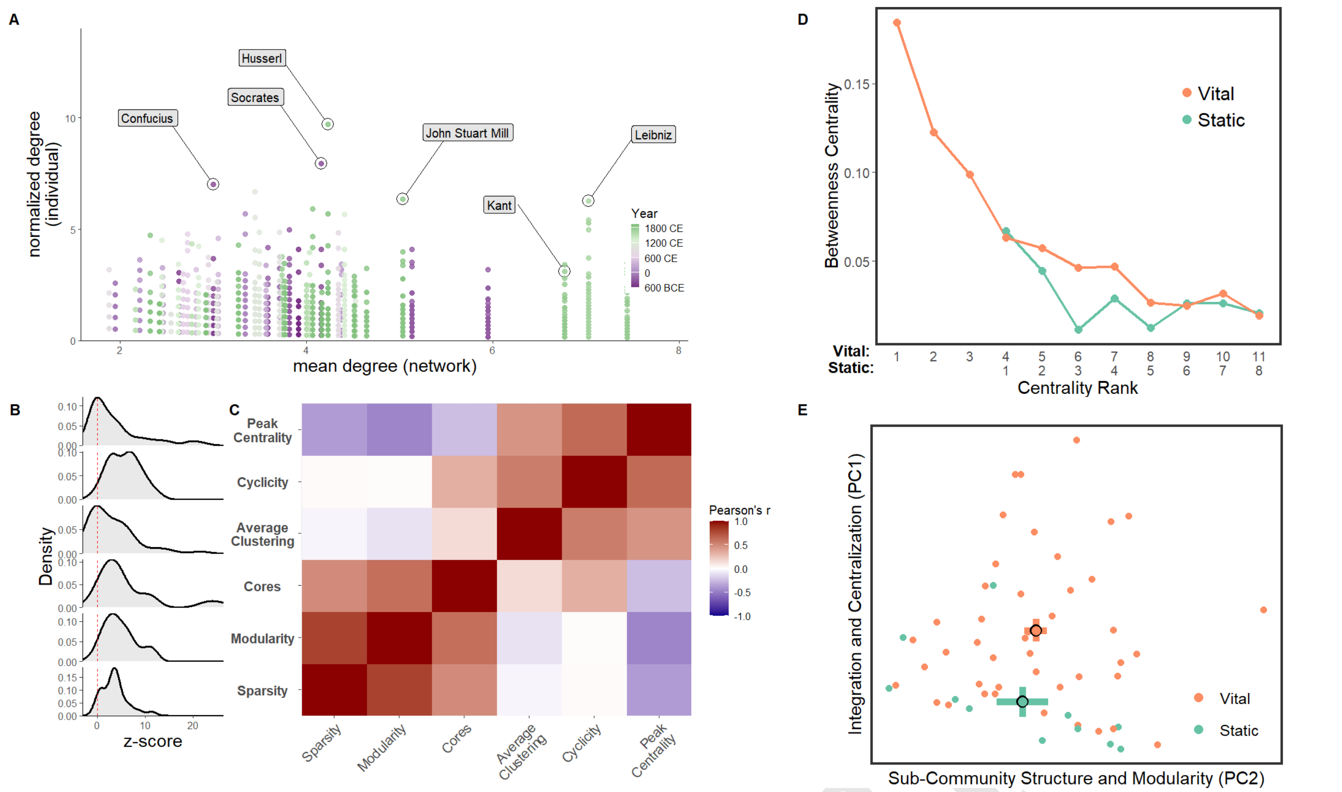

Fig. 2. Quantifying the structure of intellectual communities. **(A)** Historically notable figures were highly connected within their communities. Each point ($N = 3187$) represents one intellectual's normalized degree (y-axis) within their community, plotted against the mean degree of that community (x-axis). Columns of points represent intellectuals within a particular community. Historically transformative intellectuals (annotated) appear consistently as high-degree nodes relative to their peers. (Color indicates each community's mean year.) **(B)** Intellectual communities were highly structured. Density plots of key network measures. Measures for each network were z-scored relative to the network's null distribution derived from time-respecting null networks (see Methods for details). Red lines at 0 thus represent the expected value of the time-respecting null networks, and density plots represent z-scored values for real-world communities. (See Figure S2 for additional network measures). **(C)** Correlation heatmap of the same key network measures shown in (B). (See Figure S1 for additional network measures.) **(D)** Rank ordering of agent betweenness centrality in vital and static networks. For each network, the agent with the nth highest betweenness centrality was identified and points represent the group average at each rank across all networks. Static group points are slightly right-shifted, highlighting convergence between network types. Color indicates the network's epistemic vitality.) **(E)** Vital and static communities differed in network structure. Each network was embedded in a 2-dimensional structure space, created using Principle Component Analysis. PC1 primarily captures network integration and centralization, while PC2 relates more to community fragmentation. Color indicates the network's epistemic vitality. Crosses represent means $\pm$ standard errors.

practices a reliance on intellectual exchange, quantifying that intellectual change has been difficult to do at scale. In principle, traditions of abstract reasoning should serve as model systems for studying the social life of knowledge, but doing so in practice has been difficult.

To investigate the structural features of social networks that generate epistemic vitality, we digitized and analyzed the structure of social exchange within intellectual communities spanning nearly three thousand years (see Methods for details). Since co-authorship was infrequent in many of these communities, social networks must therefore be reconstructed from historical evidence. This was undertaken by the sociologist Randall Collins, who reconstructed the social structure of historical intellectual communities on the basis of a large-scale synthesis of 691 historical texts and biographies (21). We digitized and analyzed this dataset of social networks, which consists of individual intellectuals ($N = 3187$ nodes) and their discursive relationships ($N = 5415$ edges) across multiple historical communities ($N = 55$ networks) spanning nearly three millennia (Fig. 1). Within each network — which often spanned multiple generations — individuals are connected through various relationships, such as student-teacher ties, alliances, or rivalries. Each community is thus characterized by a social network that emerges over time from a mix of differing relationships. These networks vary in complexity, with some containing over a hundred nodes (intellectuals) and hundreds of edges (relationships). Additionally, drawing on Collins's synthetic assessments based on hundreds of historical documents (21), we categorized communities based on their historically documented innovation (i.e., epistemic vitality) or emphasis on preservation (i.e., epistemic stasis).

Using this corpus, we quantified the structure of intellectual social networks to examine how different patterns of social connection relate to the dynamics of collective thought. Drawing inspiration from the theoretical accounts surveyed above, we measured a variety of factors that have been proposed to facilitate epistemic vitality. We analyzed the temporal evolution of these measures within each network, investigating how structure emerged over time. Finally, we examined how individual actors and broader community structures have related to epistemic vitality throughout the history of knowledge production, showing that the temporal trajectory of a community's network structure predicts its epistemic vitality.

Results

Quantifying the social structure of intellectual debate. Based solely on local connectivity, foundational figures from the history of intellectual debate stood out as highly influential relative to their peers (Fig. 2A). For each intellectual, we operationalized their relative influence within a network as their normalized degree (i.e., the number of links to other intellectuals, z-scored within each community). As a few illustrative examples, highest-degree nodes included Confucius, Socrates, Husserl, John Stuart Mill, and Leibniz, all major figures from the history of thought; Kant was second only to Schiller.

The contributions of individual intellectuals, however, should depend on the larger communities in which they are embedded. We thus quantified the social structure of intellectual debate across history by characterizing each community's social network using a set of structural measures. These included macroscale properties (e.g., modularity, number of k-cores, sparsity), microscale features (e.g., peak centrality, cyclicity, average clustering), and global distance-based features (e.g., average path length, network diameter). These measures tended to group into three correlated clusters related to community structure, network centralization, and path length (Figure 2C; Figure S1; see Table S1 for descriptions of each measure and their correlations). To contextualize these measures, we generated time-respecting null models by randomly rewiring each network's edges while preserving temporal structure (see Methods for details).

Intellectual networks generally exhibited greater community structure and network centralization than expected by chance. Intellectual communities exhibited significantly greater modularity and sparsity than their null counterparts, with 96.2% of networks scoring above the median of their respective null distributions for both measures. Binomial tests confirmed that the proportion of networks exceeding their null counterparts was significantly greater than chance for both modularity ($p < 0.001$) and sparsity ($p < 0.001$), suggesting that intellectual communities tend to be both well-partitioned and loosely connected. Real networks also contained more communities (79% above nulls, $p < 0.001$), indicating a tendency toward intellectual factionalism; were more cyclical (96%, $p < 0.001$), reflecting discursive feedback between scholars; and more clustered (68%, $p = 0.006$), a sign of tight-knit local interactions. Additionally, peak centrality was higher in 70% of networks ($p = 0.003$), suggesting the presence of individuals who were especially influential within their respective systems.

The network structure of epistemic vitality. We next investigated whether certain network structures facilitate epistemic vitality. We classified intellectual communities into periods of epistemic vitality or of stasis, based on the account of global intellectual exchange in Collins' *Sociology of Philosophies* (21), which synthesized historical research on intellectual communities throughout history. Epistemically vital periods were characterized by the emergence and synthesis of new ideas, while static periods either maintained existing ideas or involved unresolved disputes that did not lead to synthesis (Figure 1B). We assessed the reliability of this classification of epistemic vitality using a pretrained machine learning model (see Methods); in the Supplemental Materials (Supplementary Note S3), we show that our results are robust to whether we use human or machine classification of epistemic vitality.

Overall, vital and static communities did not differ significantly in size ($M_{static} = 62$ vs. $M_{vital} = 57$; Student's t-test: $t = 0.405, p = .69$) or number of edges ($M_{static} = 78$ vs. $M_{vital} = 105, t = -0.914, p = .37$). The question, then, is whether these nodes and edges were *organized* differently.

Several theoretical accounts suggest that a community's epistemic vitality is shaped by its social structure, including factors such as interconnectedness (33), levels of disagreement (27), centralizing individuals (34), and productive fragmentation (24). We evaluated these proposals using measures of the social network's structure: connectivity and average degree (representing interconnectedness of each community), peak centrality (indicating the existence of centralizing figures), and the scaled number of sub-communities (capturing fragmentation versus integration). As these were correlated (Fig. 2), below we also reduced these variables to two dimensions using Principal Component Analysis and examined whether networks differing in vitality occupied distinct regions in this reduced feature space.

To test the role of interconnectedness, we examined network connectivity and degree. Vital communities differed reliably from static ones (Fig. 2D, E). Specifically, individual intellectuals in vital communities were more integrated across the whole network (network connectivity: $M_{static} = 0.18, M_{vital} = 0.49, p =< 0.001$) and were more locally connected (average degree: $M_{static} = 1.33, M_{vital} = 1.88, p =< 0.001$).

To test the role of centralizing figures, we examined the peak centrality of intellectuals within each community. As predicted, vital communities contained individuals who were more centralizing (peak centrality: $M_{static} = 0.07, M_{vital} = 0.18, p =< 0.001$). Inspection of the rank-centrality distribution confirmed that vital communities were characterized by a few extreme outliers who were highly central; in the absence of those highly central individuals, the distribution of node centrality was qualitatively the same for vital and static communities (Fig. 2D).

Epistemically vital communities were thus more globally integrated, more locally connected, and contained extreme individuals who were centralizing. These three features all loaded highly on the first principle component, which reliably distinguished epistemically vital communities from static ones (linear regression predicting PC1: $b = 1.624 \pm 0.543$ SE, $p = 0.004$; Fig. 2E).

By contrast, vital and static communities did not differ in the tendency of intellectuals to disagree among themselves or fragment into factions. To test the role of antagonism or disagreement, we examined the tendency to dis-

Table 1. Linear mixed effects model predicting network structural measures over time.

Predictor	Connectivity	Avg Degree	Peak Centrality	Sub-Communities	Disagreement Ratio
Intercept	0.23 (0.07) [3.22]	0.61 (0.14) [4.44]	0.04 (0.03) [1.22]	0.66 (0.06) [11.27]	0.63 (0.26) [2.42]
Vitality (Start)	-0.01 (0.08) [-0.08]	0.29 (0.16) [1.81]	0.06 (0.04) [1.64]	-0.06 (0.07) [-0.87]	0.18 (0.30) [0.62]
Vitality (End)	**0.33 (0.13) [2.60]***	**0.82 (0.22) [3.75]*****	**0.13 (0.04) [3.31]****	**-0.23 (0.06) [-4.11]*****	-0.07 (0.16) [-0.42]
Time	-0.04 (0.13) [0.29]	0.14 (0.21) [0.65]	0.01 (0.05) [0.25]	-0.07 (0.08) [-0.87]	-0.12 (0.33) [-0.36]
Vitality × Time	**0.34 (0.14) [2.35]***	**0.54 (0.25) [2.18]***	0.07 (0.06) [1.16]	**-0.18 (0.09) [-2.02]***	-0.25 (0.37) [-0.67]

Note: Vitality (Start) and Vitality (End) indicate the coefficient estimates for Vitality from two different models, one where Time was centered at each community's origin, and another where Time was centered at the community's culmination. These model estimates thus capture the difference between epistemically vital and static communities at their origin and culmination, respectively. Vitaliy × Time is the interaction between the Vitaliy and Time predictors; this captures the possible divergence or convergence of Vital and Static networks over their lifetime (e.g., top two panels of Figure 3A). Standard errors and t-statistics are in parentheses and square brackets, respectively. Statistical significance is indicated by asterisks (*, < .05; **, <.01; ***, <.001).

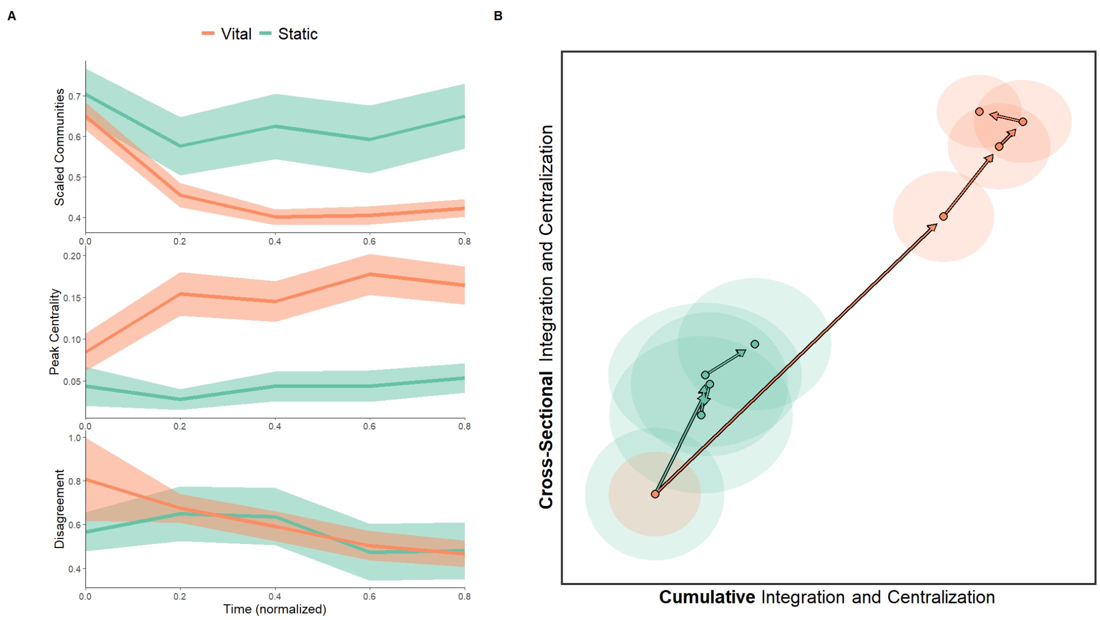

Fig. 3. The temporal emergence of epistemic vitality. **(A)** Temporal evolution of network structure within cross-sectional time periods. Vital and static communities did not differ at their origin. Over time, however, vital communities became less fractionated (scaled communities) and contained more highly central individuals (peak centrality). By contrast, vital and static communities had similar amounts of disagreement throughout their existence. (Lines = means; shaded ribbons = standard errors.) See Figure S4 for additional network measures. **(B)** Temporal differentiation of the network structure of vital (orange) and static (green) communities. The x-axis shows the network structure (integration and centralization) within each non-overlapping, cross-sectional time period. The y-axis shows the network structure accumulated from a community's origin up until that time period. This allows us to visualize the temporal evolution of network structure, both cross-sectional (x-axis) and cumulative (y-axis). Static communities show little change over time in the integration and centralization of intellectuals within a given time period (x-axis) and only a slight increase in cumulative structure. Vital communities, by contrast, grew more integrated and centralized both with each subsequent time periods (x-axis) and overall (y-axis). (Points = means. Shaded circles = standard errors.

agree within each community (ratio of conflictual to non-conflictual edges). Vital and static communities did not differ in antagonism ($M_{static} = 0.45, M_{vital} = 0.26, p = 0.19$). This suggests that it is not the sheer amount of disagreement that distinguishes vital communities, but rather how that disagreement is structured. To test the role of fragmentation, we measured the tendency of communities to cluster into sub-communities (number of sub-communities, scaled by network size). Vital and static networks did not differ along this dimension ($M_{static} = 0.19, M_{vital} = 0.16, p = 0.24$). This measure loaded highly on the second principal component, which did not reliably distinguish vital communities from static ones (linear regression predicting PC2: $b = 0.148 \pm 0.285$ SE, $p = 0.61$). Epistemic vitality may depend less on how much communities disagree or fragment into factions, and more on how disagreements and sub-communities are organized to support productive tensions at the collective level.

Thus, while intellectuals in vital and static communities were equally likely to fragment into distinct sub-communities (PC2 in Fig. 2D), in vital communities the overall structure was more integrated and centralized (PC1 in Fig. 2D). This is consistent with theories of epistemic vitality that foreground the importance of structural coherence and centralizing figures, who can bridge communities and facilitate the circulation of ideas across otherwise disconnected subgroups (33, 34).

.

The temporal dynamics of epistemic vitality. We next investigated the temporal emergence of these networks. To analyze the temporal trajectory of network structure, networks were divided into non-overlapping temporal periods

(Fig. 1c), following the temporal divisions introduced in Randall Collins's visual representations of the history of thought (21). Within each network, these periods were of equal duration, typically corresponding to approximately one generation, although sometimes of longer duration for communities that persisted for hundreds of years. To examine how centrality and community structure shape the dynamics of agreement and disagreement in systems with differing levels of epistemic vitality, we analyzed how these distinguishing measures evolved over time, including the scaled number of communities, peak centrality, average degree, and the ratio of disagreement to agreement. To investigate the temporal evolution of these measures within each network, for each measure we fit a mixed effects model, with fixed effects of epistemic vitality, time within each community, and their interaction (Table 1; see Methods for details).

Our analysis above of holistic community structure revealed that epistemic vitality was associated with greater network integration. We thus investigated how integration emerged over the life of an intellectual community (Table 1; plotted in Supplementary Note 5.). We find that, at their origin, vital and static networks did not differ in integration, as measured by connectivity (effect of epistemic vitality on connectivity at communities' origination ($b = -0.01 \pm 0.08$ SE, $p = 0.94$). Over time, however, the connectivity of non-vital communities remained stable (change over time in static communities: $b = -0.04 \pm 0.13$ SE, $p = 0.77$), while epistemically vital communities became more connected (vital communities: $b = 0.30 \pm 0.07$ SE, < 0.001). At their culmination, therefore, epistemically vital communities were more connected than static communities (effect of epistemic vitality on connectivity at the communities' culmination: $b = 0.33 \pm 0.13$ SE, $p = 0.013$).

There was a similar pattern for community density as measured by average degree (Table 1 and Fig. 3A, bottom left), with no significant difference between vital and static communities at the beginning ($b = 0.29 \pm 0.16$ SE, $p = 0.080$), but significantly increased in average degree in vital communities ($b = 0.68 \pm 0.12$ SE, $p < 0.001$), so that vital communities had higher degree at their culmination. In other words, while both static and vital communities *begin* with similar measures of local and global connectivity, vital communities became more integrated over time, with individual nodes being more connected to nodes in disparate parts of their communities and overall more connected as individuals.

Vital communities were also distinguished by by the temporal emergence of centralizing individuals who connect different parts of the network (Table 1 and Fig. 3A, top right). In the analyses above of the whole networks, we found that epistemic vitality was associated with the presence of highly centralizing figures, as measured by peak centrality. Initially, however, both vital and static communities exhibited similar peak centrality (centrality at origin: $b = -0.06 \pm 0.04$ SE, $p = 0.11$), but at their culmination the vital communities had higher peak centrality ($b = -0.13 \pm 0.04$ SE, $p = 0.002$). This emergent difference reflected a significant increase over time in the peak centrality of vital communities ($b = 0.08 \pm 0.03$ SE, $p = 0.008$) but not in static communities ($b = 0.01 \pm 0.05$ SE, $p = 0.81$), though this difference was not itself significant ($b = 0.07 \pm 0.06$ SE, $p = 0.254$). Thus, while vital and static communities were equally connected by centralizing figures at their origin, only in vital communities do highly centralizing figures emerge who bridge divisions and consolidate influence, occupying key positions within the network.

Vital discourses may emerge from the merging of many disparate communities into a few highly-concentrated ones. Our analyses of whole networks found no association between epistemic vitality and the tendency for individuals to organize into sub-communities. The epistemic benefits of tight-knit sub-communities, however, may be specific to sub-communities that co-exist within the same time period. We thus analyzed the change over of time of sub-communities calculated cross-sectionally (i.e., among individuals within the same time period; Table 1 and Fig. 3A, top left). When they first originated, epistemically vital communities had the same amount of sub-community structure as other communities (effect of epistemic vitality on number of sub-communities, scaled by community size of temporal period at the communities' origination: $b = -0.06 \pm 0.07$ SE, $p = 0.39$). Over time, however, the community structure of non-vital communities remained stable (change over time in static communities: $b = -0.07 \pm 0.08$ SE, $p = 0.39$), while epistemically vital communities gradually coalesced into a smaller number of communities (vital communities: $b = -0.24 \pm 0.04$ SE, $p < 0.001$). At their culmination, therefore, epistemically vital communities had significantly fewer communities than other communities (effect of epistemic vitality on scaled sub-communities at the communities' culmination: $b = -0.23 \pm 0.06$ SE, $p < 0.001$). These results suggest that while static and vital communities begin with comparable levels of fragmentation, only vital communities undergo a process of structural consolidation over time. This dynamic, not reflected in the results from the static whole network measures, arises as communities at each successive point in time became more integrated with one another compared to previous periods, showing how epistemically vital communities became increasingly integrated across generations.

The vital and static communities did not differ along all dimensions. Some theories of epistemic vitality, for instance, argue that epistemically vital communities are characterized by high levels of disagreement (27). To test such theories, we looked at the relative amount of disagreement and agreement within each network, measured as the ratio of conflictual edges to non-conflictual edges (Table 1 and Fig. 3A, bottom right). We found no evidence that epistemic vitality is associated with individual-level antagonism on its own. At their temporal origination, vital and static communities did not differ in the amount of disagreement ($b = 0.18 \pm 0.3$ SE, $p = 0.54$). Both types of communities typically decreased in disagreement over time, but the amount of decrease did not differ significantly ($b = -0.25 \pm 0.37$ SE, $p = 0.50$). As a result, at their culmination, vital communities and static communities did not differ in the amount of disagreement ($b = -0.07 \pm 0.16$ SE, $p = 0.67$). The amount of disagreement on its own, therefore, is insufficient for epistemic vitality. Vital communities are not merely cantankerous; they organize disagreement into tight-knit communities that are connected by centralizing figures.

To synthesize these changes in network structure over time, we projected these network measures onto the first principal component derived from the Principal Component Analysis of whole-network measures described in the previous section. We did this both for network measures calculated cross-sectionally (i.e., for each non-overlapping time period) and cumulatively (i.e., for the entire network, from its origin to the current time period). This yielded a two-dimensional embedding of the network's time-evolving structure (Fig. 3B). At their origin, vital (orange) and static (green) communities were indistinguishable on the basis of their cumulative (y-axis) or cross-sectional (x-axis). Over time, static communities changed only minimally in their cumulative and cross-sectional structure, with their core structural attributes remaining largely stable (green trajectory in Fig. 3B). In contrast, epistemically vital communities showed consistent structural change in both their cumulative and cross-sectional structure. As a result, at their culmination, vital communities differed significantly from static ones in both their cumulative structure (measured at the final time point: $M_{static} = -3.82$ vs. $M_{vital} = -1.36$, $p < 0.001$) and their cross-sectional structure ($M_{static} = -4.37$ vs. $M_{vital} = -2.08$, $p < 0.001$). This suggests that epistemic vitality is associated with increasing structural integration and centralization over time, while static systems maintain a more fragmented and diffuse configuration.

Discussion

Understanding the social drivers of intellectual progress is a long-standing challenge. Our study of three millennia of intellectual debate offers new insights by asking: What kinds of community structures facilitate epistemic vitality? By quantifying the social network structure of nearly 3,000 years of intellectual debate, we found that epistemic vitality is more likely to emerge in communities characterized by greater interconnectedness (mean degree), increasing integration (fewer fragmented subcommunities), and the presence of centralizing figures (individuals with high centrality). Intellectual communities were thus more likely to sustain epistemic vitality if they consisted of integrated groups connected by centralizing figures. Diffuse communities that lacked clear intellectual lineages or centralizing figures remained epistemically static. While epistemically vital and static communities began with similar social structures, the features associated with vitality, such as the presence of strong centralizing figures and connectivity, tended to emerge over time, pointing to the importance of bridging nodes and integrative figures. At their origin, both vital and static communities consisted of fragmented subgroups. But while static communities maintained this fragmentation, vital communities became integrated by centralizing individuals (Fig. 2A, D), suggesting that a key historical driver of epistemic vitality is the bridging of previously disconnected perspectives into a cohesive and "energetic" intellectual discourse (Fig. 3B) (21).

Studies of collective intelligence have proposed at least three accounts of epistemic vitality: (1) that vital communities are sustained by persistent, individual-level disagreement (24, 26); (2) that they emerge from loosely connected sub-communities pursuing parallel lines of inquiry (28, 29); and (3) that they depend on central actors who bridge these communities and integrate their insights (21, 33). Our findings inform all three accounts. Our analyses of historical intellectual communities suggested that — while disagreement may play a role — the structure of this disagreement, rather than its mere presence, is more critical. We also found that while most intellectual communities begin with multiple sub-communities, vital communities show evidence of consolidation of this community structure over time time. Furthermore, this consolidation appears to be driven, in part, by the emergence of centralizing individuals who play a key role in linking and synthesizing across sub-communities. Notably, the community structures associated with epistemically vital communities differ from those predicted by theories emphasizing sparsity and discord (24). Moreover, our findings highlight the degree to which network structure varies over time; this raises the possibility that features such as centralization may play different roles over the lifetime of a community, thus helping to resolve apparent contradictions in past work (e.g., (38, 39)).

Our approach was informed by research on the role that social structures play in shaping the vitality of other knowledge systems and was thus idea-agnostic, focusing not on the content of intellectual thought but on the structural dynamics that shape how epistemic systems evolve and interact over time (4, 40–43). We have not examined the specific content of different epistemic systems or the specific intellectual role that centralizing thinkers may have played in these systems, whether as radical innovators, synthesizers, or disruptors. This allowed us to first isolate the potential role of network topology itself, a factor that has been implicated in the dynamics of other epistemic systems (31, 33, 36). Future work may help bridge this gap between social structure and intellectual content by incorporating the sociopolitical contexts and ideological contents of these traditions, as emphasized by sociologists and historians of ideas (20, 21, 44). For instance, to clarify how epistemic commitments manifest in both discourse and community structure, one could integrate community-level structural analyses, such as those presented here, with natural language processing of the products of those communities (e.g., publications, letters, etc.) (45, 46).

One fertile direction for future research lies in a comparative project examining structural differences between different epistemic enterprises. Studies of epistemic communities of the sort pursued here and within the science of science could be extended to a wide range of discursive knowledge systems, including those found in medicine and other so-called "thought collectives" (2, 47). Such work could help clarify, for instance, whether science, philosophy, logic, and religion represent fundamentally distinct social processes and how each contributes to the broader social systems through which humans understand and engage with the world. Such a project holds promise to integrate the science of science (7–18), the philosophy of science (2–5, 48, 49), and the history of ideas (19–21) to forge a comprehensive understanding of the dynamics of collective knowledge.

Methods

Digitizing the social networks of intellectual exchange. Networks were digitized from Randall Collins's comparative history of global intellectual thought, *The Sociology of Philosophies* (21). On the basis of hundreds of historical documents ($N = 691$), Collins reconstructed networks of intellectual agreement, disagreement, and master-pupil relationships for epistemic communities across history ($N = 3187$ nodes representing intellectuals and $N = 5415$ edges across $N = 55$ networks). Communities in the dataset span over 2700 years, from 800 BCE to 1980 CE; cover a wide geographic range, stretching across Europe, North Africa, the Middle East, India, China, and Japan, and more; and encompassing epistemic movements from ancient Buddhism to modern French existentialism. In total, the dataset includes 55 networks:

- 24 from Asia, including 10 with intellectuals in ancient China, 7 in India, 8 in Japan, and 4 encompassing Persian, Middle Eastern, and Central Asian thinkers,
- 30 from Europe (9 from Ancient Greece, 5 from Rome, and the remainder spread across Western and Central Europe) and 5 from North Africa and the Middle East,
- 1 from North America (the American Pragmatist movement).

The dataset thus reflects a diversity of eras, regions, and intellectual traditions, across multiple languages and cultural contexts, from Classical Chinese, Sanskrit, and Arabic, to Latin, German, French, and English.

We digitized the network diagrams in Collins' text (21). For each network, two independent coders translated Collins' visualization into a directed, weighted edge list. In each edge list, intellectual disputes were assigned a directed negative weight (-1), master-pupil relationships received a directed positive weight ($+1$), and acquaintances were coded as bidirectional positive edges ($+1$ in both directions). In cases of coding discrepancies, a third coder adjudicated. We also divided each network into non-overlapping temporal periods, following the temporal divisions introduced in Randall Collins's visualizations (21). Within each network, these periods were of equal duration, typically corresponding to approximately one generation, although sometimes of longer duration for communities that persisted for hundreds of years. To situate each individual to points in time, we follow the ordering in Collins's text, which anchors them within critical periods of exchange in their broader networks. To assess the reliability of these edge lists, we used a pretrained large language model (OpenAI's GPT-4o) to independently assess relationships between intellectuals. The model agreed substantially with the edge lists (Cohen's $\kappa = 0.69$; see SM Methods and Fig. S5 for details).

Collins also evaluated the epistemic vitality of each community using a comparative methodology that spanned a large temporal and geographical range. This allows for a bird's eye evaluation of different intellectual communities' relative productivity and vitality. A coder reviewed Collins' discussion of each community and made a made a holistic determination of the community's epistemic vitality. To assess the reliability of these judgments, we used a pretrained machine learning model, OpenAI's GPT-4o, to independently assess the epistemic vitality of each network using Monte Carlo cross-validation (repeated random subsampling). We prompted the model with a definition of epistemic vitality, along with a a labeled random subset of communities (75% each of vital and static communities). Communities were described only by name (e.g., "Network of Greek Philosophers from Socrates to Chrysippus") and date range. We then asked the model to decide whether or not each of the remaining communities was epistemically vital. This process was repeated 101 times, and for each community we calculated how often it was classified as epistemically vital by the model. This ML-derived measure of epistemic vitality was significantly correlated with the human classification of epistemic vitality classification derived from Collins (2000) (21) ($r = .30$). In the Supplemental Materials, we reproduce our main results using this machine classification of epistemic vitality.

Quantifying network topology. For each network we calculated 13 measures of structural and relational properties: average degree, average clustering, peak centrality, connectivity, flow hierarchy, average path length, network diameter, sparsity, modularity, communities, cores, coreness, and the ratio of disagreement to agreement. (See Table S1 for a description of all network measures.) We calculated these measures for each entire network. Edge weights were only used to calculate the ratio of agreement to disagreement. We also calculated the temporal evolution of these measures within each community in two ways: (1) cross-sectionally, for each non-overlapping time period, and (2) cumulatively, incorporating all network interactions from the beginning up to the given time period.

To capture variation along theoretically-important dimensions of interconnectedness (33), centralizing individuals (34), and productive fragmentation (24), we performed Principal Component Analysis (PCA) on four key variables (connectivity, scaled communities, peak centrality, and average degree) calculated on the whole networks. The first principal component (PC1) explained 74% of the variance. We also used this principle component to capture the temporal evolution of network structure by projecting the cross-sectional measures and the cumulative measures on to PC1.

Time-respecting null models. For each of these growing networks, we constructed time-respecting null models ($N = 1000$) (50) in which edges are randomly shuffled while respecting the network's temporal structure. To do so, edges are shuffled using a Newman edge-rewiring algorithm (51) in two ways. Edges between nodes within the same time period are shuffled to connect random nodes within the same time period. Edges between a given time period and a preceding time period (i.e., between intellectuals and their predecessors) are shuffled so they connect the same two periods. This approach preserves the time-respecting nature of the networks and the balance of edges both within and between time periods, while randomizing the overall network structure and degree distribution.

Statistical analyses. To assess the structural distinctiveness of the real networks, we calculated all measures for each time-respecting null model and compared the real networks to these null distributions (i.e., z-scored using the null distributions).

To analyze the networks' temporal evolution, we used linear mixed effect models to predict each measure of network structure. Models included fixed effects for time (normalized within each network to range from 0 to 1), epistemic vitality (static = 0, vital = 1), and their interaction, and random intercepts and slopes for time for each network. To estimate the effect of epistemic vitality at the networks' culmination (i.e., when time = 1), we rebaselined time so it runs from -1 (origin) to 0 (culmination). To estimate the change over time for epistemically vital networks, we rebaselined epistemic vitality (i.e., vital = 0, static = 1).

Data Availability. All data generated and analyzed in this study will be publicly available via the Open Science Framework (OSF). This repository includes all network data, metadata, and supplementary materials necessary to replicate and extend the findings of this study.

Code Availability. All custom code used for data processing, network analysis, and visualization will be available in the same OSF repository.

ACKNOWLEDGEMENTS

This work received support from the Institute for Humane Studies under grant no. IHS017802 awarded to CM.

References

1. Robert K Merton. *The sociology of science: Theoretical and empirical investigations.* University of Chicago press, 1973.
2. Ludwik Fleck. *Genesis and Development of a Scientific Fact.* University of Chicago Press, 1981.
3. Thomas S. Kuhn. *The Structure of Scientific Revolutions.* University of Chicago Press, Chicago, 1962.
4. David L. Hull. *Science as a Process: An Evolutionary Account of the Social and Conceptual Development of Science.* University of Chicago Press, Chicago, 1988. ISBN 9780226360507.
5. Paul Thagard. Societies of minds: Science as distributed computing. *Studies in History and Philosophy of Science*, 24(49-67):251–257, 1993.
6. Edwin Hutchins. *Cognition in the Wild.* MIT press, 1995.
7. Stefan Wuchty, Benjamin F Jones, and Brian Uzzi. The increasing dominance of teams in production of knowledge. *Science*, 316(5827):1036–1039, 2007.
8. Barry Bozeman and Elizabeth Corley. Scientists' collaboration strategies: implications for scientific and technical human capital. *Research policy*, 33(4):599–616, 2004.
9. James D Adams, Grant C Black, J Roger Clemmons, and Paula E Stephan. Scientific teams and institutional collaborations: Evidence from us universities, 1981–1999. *Research policy*, 34(3):259–285, 2005.
10. Fuyuki Yoshikane and Kyo Kageura. Comparative analysis of coauthorship networks of different domains: The growth and change of networks. *Scientometrics*, 60:435–446, 2004.
11. Alexander Michael Petersen. Quantifying the impact of weak, strong, and super ties in scientific careers. *Proceedings of the National Academy of Sciences*, 112(34):E4671–E4680, 2015.
12. Lingfei Wu, Dashun Wang, and James A Evans. Large teams develop and small teams disrupt science and technology. *Nature*, 566(7744):378–382, 2019.
13. Mark EJ Newman. Who is the best connected scientist? a study of scientific coauthorship networks. *Phys. Rev. E*, 64(016131), 2001.
14. Santo Fortunato, Carl T Bergstrom, Katy Börner, James A Evans, Dirk Helbing, Staša Milojević, Alexander M Petersen, Filippo Radicchi, Roberta Sinatra, Brian Uzzi, et al. Science of science. *Science*, 359(6379):eaao0185, 2018.
15. Marcus Antonius Ynalvez and Wesley M Shrum. Professional networks, scientific collaboration, and publication productivity in resource-constrained research institutions in a developing country. *Research Policy*, 40(2):204–216, 2011.
16. Waverly W Ding, Sharon G Levin, Paula E Stephan, and Anne E Winkler. The impact of information technology on academic scientists' productivity and collaboration patterns. *Management Science*, 56(9):1439–1461, 2010.
17. Pierre Azoulay, Joshua S Graff Zivin, and Gustavo Manso. Incentives and creativity: evidence from the academic life sciences. *The RAND Journal of Economics*, 42(3):527–554, 2011.
18. Brian A Jacob and Lars Lefgren. The impact of research grant funding on scientific productivity. *Journal of public economics*, 95(9-10):1168–1177, 2011.
19. Charles de Dampierre and Hugo Mercier. The structure and evolution of scholarly interests from antiquity to the eighteenth century. *Scientometrics*, pages 1–21, 2025.
20. Hajime Nakamura. *A Comparative History of Ideas.* Kegan Paul International, London, 1992. ISBN 071030384X.
21. Randall Collins. *The sociology of philosophies: A global theory of intellectual change.* Harvard University Press, 2000.
22. David Kahle and Hadley Wickham. ggmap: Spatial visualization with ggplot2. *The R Journal*, 5(1):144–161, 2013.
23. John Henry Newman. *An essay on the development of Christian doctrine.* James Toovey, 1846.
24. Paul E Smaldino, Cody Moser, Alejandro Pérez Velilla, and Mikkel Werling. Maintaining transient diversity is a general principle for improving collective problem solving. *Perspectives on Psychological Science*, 19(2):454–464, 2024.
25. Jingyi Wu and Cailin O'Connor. How should we promote transient diversity in science? *Synthese*, 201(2):37, 2023.
26. Kevin JS Zollman. The epistemic benefit of transient diversity. *Erkenntnis*, 72(1):17–35, 2010.
27. Feng Shi, Misha Teplitskiy, Eamon Duede, and James A Evans. The wisdom of polarized crowds. *Nature human behaviour*, 3(4):329–336, 2019.
28. Yiling Lin, James A Evans, and Lingfei Wu. New directions in science emerge from disconnection and discord. *Journal of Informetrics*, 16(1):101234, 2022.
29. Christina Fang, Jeho Lee, and Melissa A Schilling. Balancing exploration and exploitation through structural design: The isolation of subgroups and organizational learning. *Organization science*, 21(3):625–642, 2010.
30. Jacob G Foster, Andrey Rzhetsky, and James A Evans. Tradition and innovation in scientists' research strategies. *American sociological review*, 80(5):875–908, 2015.
31. Jesse Milzman and Cody Moser. Decentralized core-periphery structure in social networks accelerates cultural innovation in agent-based modeling. In *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems*, pages 2213–2221, 2023.
32. Peter Csermely, András London, Ling-Yun Wu, and Brian Uzzi. Structure and dynamics of core/periphery networks. *Journal of Complex Networks*, 1(2):93–123, 2013.
33. Cody Moser and Paul E Smaldino. Innovation-facilitating networks create inequality. *Proceedings of the Royal Society B*, 290(2011):20232281, 2023.
34. Janet C Long, Frances C Cunningham, and Jeffrey Braithwaite. Bridges, brokers and boundary spanners in collaborative networks: a systematic review. *BMC health services research*, 13:1–13, 2013.
35. Daril A Vilhena, Jacob G Foster, Martin Rosvall, Jevin D West, James Evans, and Carl T Bergstrom. Finding cultural holes: How structure and culture diverge in networks of scholarly communication. *Sociological Science*, 1:221–23, 2014.
36. Mark EJ Newman. The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences*, 98(2):404–409, 2001.
37. Blaise Cronin, Debora Shaw, and Kathryn La Barre. A cast of thousands: Coauthorship and subauthorship collaboration in the 20th century as manifested in the scholarly journal literature of psychology and philosophy. *Journal of the American Society for information Science and Technology*, 54(9):855–871, 2003.
38. Valentin Danchev, Andrey Rzhetsky, and James A Evans. Centralized scientific communities are less likely to generate replicable results. *Elife*, 8:e43094, 2019.
39. David de la Croix, Rossana Scebba, and Chiara Zanardello. Flora, cosmos, salvatio: Premodern academic institutions and the spread of ideas. Discussion Paper DP20569, Centre for Economic Policy Research (CEPR), 2025. CEPR Press, Paris & London.
40. Damon Centola. The network science of collective intelligence. *Trends in Cognitive Sciences*, 26(11):923–941, 2022.
41. Pablo Contreras Kallens, Daniel J Hicks, and Carolyn Dicey Jennings. Networks in philosophy: Social networks and employment in academic philosophy. *Metaphilosophy*, 53(5): 653–684, 2022.
42. Pablo Contreras Kallens and Jeff Yoshimi. Bibliometric analysis of the phenomenology literature. In *Horizons of Phenomenology: Essays on the State of the Field and Its Applications*, pages 17–47. Springer International Publishing Cham, 2023.
43. Martin Davies and Angelito Calma. Australasian journal of philosophy 1947–2016: a ret-

rospective using citation and social network analyses. *Global Intellectual History*, 4(2): 181–203, 2019.

44. Dean Keith Simonton. The sociopolitical context of philosophical beliefs: A transhistorical causal analysis. *Social Forces*, 54(3):513–523, 1976.
45. Dan Edelstein, Paula Findlen, Giovanna Ceserani, Caroline Winterer, and Nicole Coleman. Historical research in a digital age: reflections from the mapping the republic of letters project. *The American Historical Review*, 122(2):400–424, 2017.
46. Ruth Ahnert and Sebastian E Ahnert. *Tudor networks of power*. Oxford University Press, 2023.
47. Jonathan Harwood. Ludwik fleck and the sociology of knowledge. *Social studies of science*, 16(1):173–187, 1986.
48. Patrick Grim, Frank Seidl, Calum McNamara, Isabell N Astor, and Caroline Diaso. The punctuated equilibrium of scientific change: a bayesian network model. *Synthese*, 200(4): 297, 2022.
49. Karin Knorr Cetina. Culture in global knowledge societies: Knowledge cultures and epistemic cultures. *Interdisciplinary science reviews*, 32(4):361–375, 2007.
50. Zhuo-Ming Ren, Manuel Sebastian Mariani, Yi-Cheng Zhang, and Matúš Medo. Randomizing growing networks with a time-respecting null model. *Physical Review E*, 97(5):052311, 2018.
51. Mark EJ Newman, Steven H Strogatz, and Duncan J Watts. Random graphs with arbitrary degree distributions and their applications. *Physical review E*, 64(2):026118, 2001.

Supplementary Note 1: Pairwise correlations of all network metrics

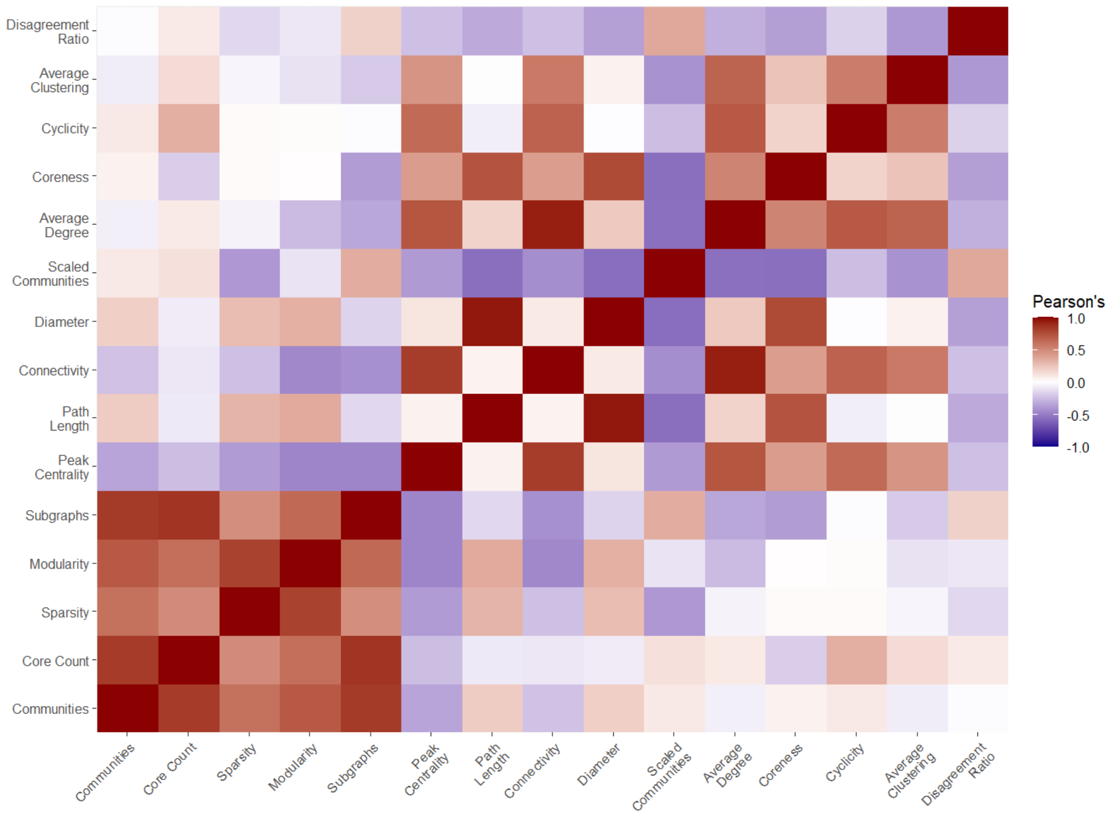

Fig. S1. Correlation heatmap of all network metrics used in the study, showing relationships among structural and community-level features. Red indicates positive correlations; blue indicates negative correlations (Pearson's r).

Supplementary Note 2: Observed and null distributions of network-based structural, community, and path length metrics

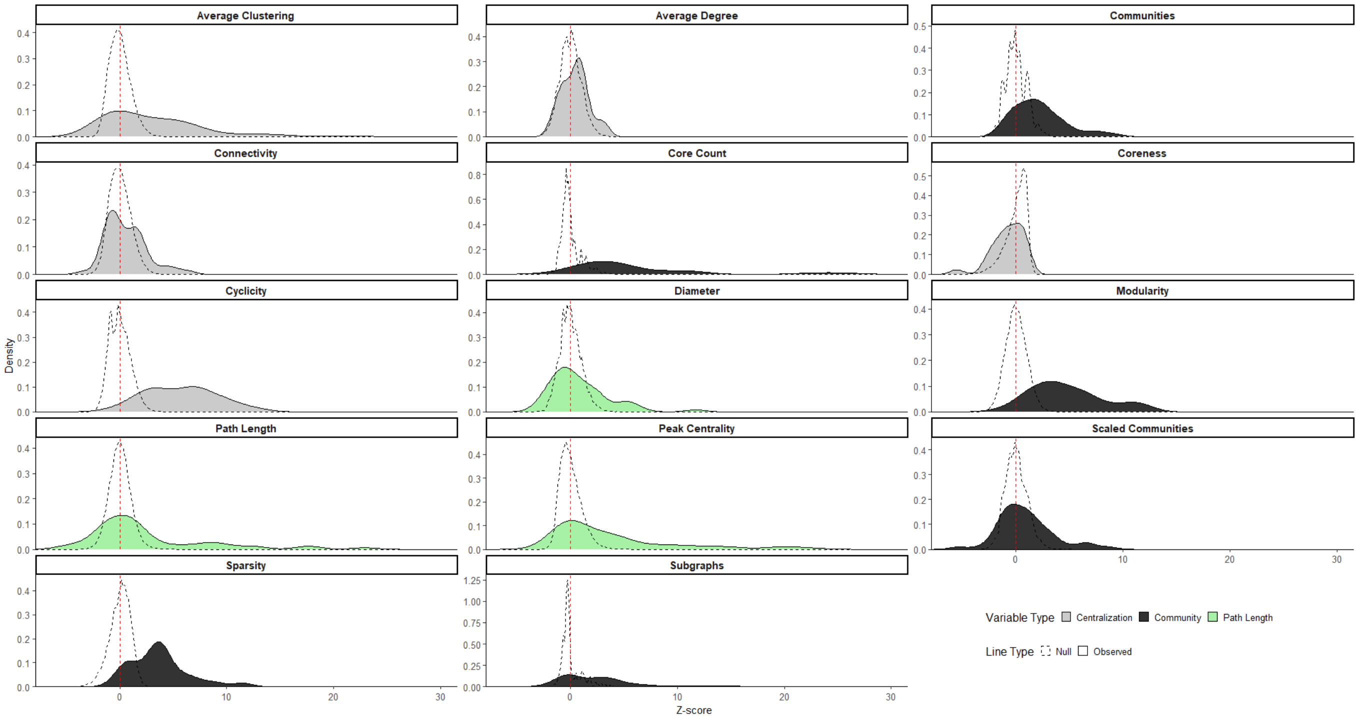

Fig. S2. Density plots of z-scored network metrics across graphs. Metrics are categorized by variable type from correlation metrics between all variables (Fig S1): Community (black), Centralization (gray), and Path Length (green), while dashed curves indicate null model distributions. Dashed vertical lines mark the mean of null distributions.

Supplementary Note 3: Reliability of Epistemic Vitality Classifications

In the Main Text, we use a measure of epistemic vitality based on human annotation of Randall Collins' global history of intellectual thought (21) (hereafter, "human-judged- vitality"). To assess the reliability this measure of epistemic vitality, we used a large language model (OpenAI's GPT-4o) to independently classify the epistemic vitality of every intellectual community (hereafter, "ML-vitality"). We employed Monte Carlo cross-validation (repeated random subsampling). In each iteration, the model was given a definition of epistemic vitality along with a labeled random subset of communities (75% each of vital and static communities). We then asked the model to classify the remaining communities as either epistemically vital or static. This process was repeated 101 times, and for each community, we calculated the proportion of runs in which it was classified as vital. These proportions were thus a continuous measure of each community's vitality assessment ("ML-vitality").

We first asked whether ML-vitality was consistent with the human-judged-vitality measures used in the Main Text. A linear regression predicting ML-vitality from human-judged-vitality showed a significant positive relationship ($b = 0.271 \pm 0.120$ SE, $p = 0.027$). A Welch two-sample t-test confirmed this association, with significantly higher ML-vitality values for human-judged-vital communities ($M_{\text{static}} = 0.39$ vs. $M_{\text{vital}} = 0.66$, $t_{53} = -2.2$, $p = 0.043$). We thus examined whether our main findings replicate when we use ML-vitality instead of human-judged-vitality.

At the network level, we found that ML-vitality predicted PC1 of the Principle Component Analysis of network structure, which captures integration and centralization (linear regression predicting PC1: $b = 0.894 \pm 0.270$ SE, $p = 0.002$; Fig. 2), thus replicating the finding in the Main Text. ML-vitality did not predict PC2, which captures fragmentation and antagonism (PC2: $b = 0.051 \pm 0.139$ SE, $p = 0.368$), once again replicating the result in the Main Text. Likewise, we find more directly that ML-vitality does not predict the levels of antagonism in communities (linear regression predicting antagonism: $b = -0.065 \pm 0.057$ SE, $p = 0.262$). Thus, epistemically vital and static communities differ significantly in their structural integration and centralization, but not in fragmentation or antagonism, whether communities' vitality was judged by human or ML model.

We next examined the temporal evolution of these networks. To convert the continuous ML-vitality measure into discrete categories, we used tertile-defined categories. The temporal evolution of these ML-defined vital and static communities was qualitatively similar to temporal evolution of human-judged categories (Fig. S3). At their culmination, ML-defined vital communities differed significantly from static ones in their cumulative structure ($M_{\text{static}} = -2.71$ vs. $M_{\text{vital}} = -0.69$, $t_{35} = -3.07$, $p = 0.004$). However, departing from the Main Text, they did not differ significantly in their cross-sectional structure at the very end ($M_{\text{static}} = -3.02$ vs. $M_{\text{vital}} = -1.60$, $t_{34} = -1.85$, $p = 0.071$).

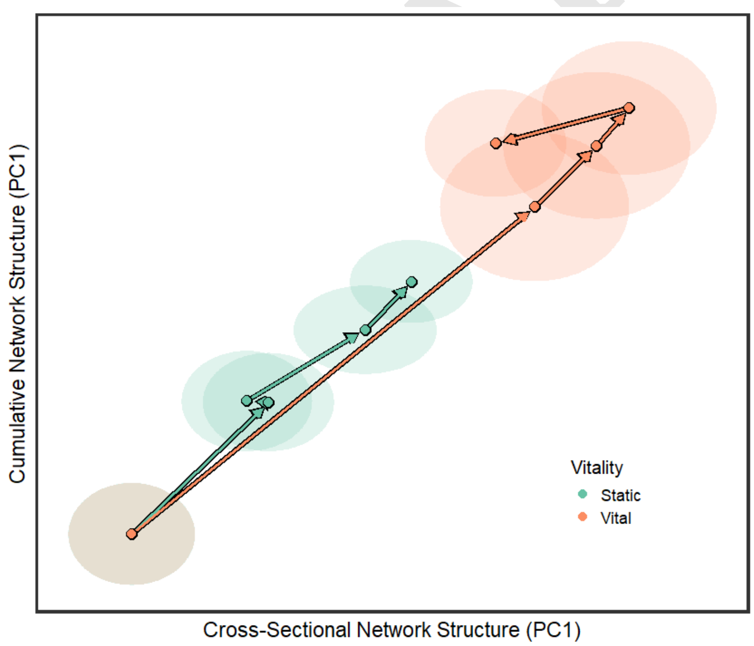

Fig. S3. Temporal differentiation of the network structure of vital (orange) and static (green) communities, using the first and third tertiles of GPT-4o's continuous vitality ratings. The x-axis shows the network structure (integration and centralization) within each time period. The y-axis shows the network structure accumulated from a community's origin up until that time period. This allows us to visualize the temporal evolution of network structure, both cross-sectional (x-axis) and cumulative (y-axis). Both static and vital communities increase in cumulative structure over time (y-axis), reflecting growing integration and centralization. However, vital communities show more consistent and pronounced changes across both cross-sectional structure (x-axis) and cumulative integration (y-axis), whereas static communities exhibit more modest or variable shifts in network structure across time periods. (Points = means. Shaded circles = standard errors.

Supplementary Note 4: Temporal evolution of network metrics by epistemic vitality

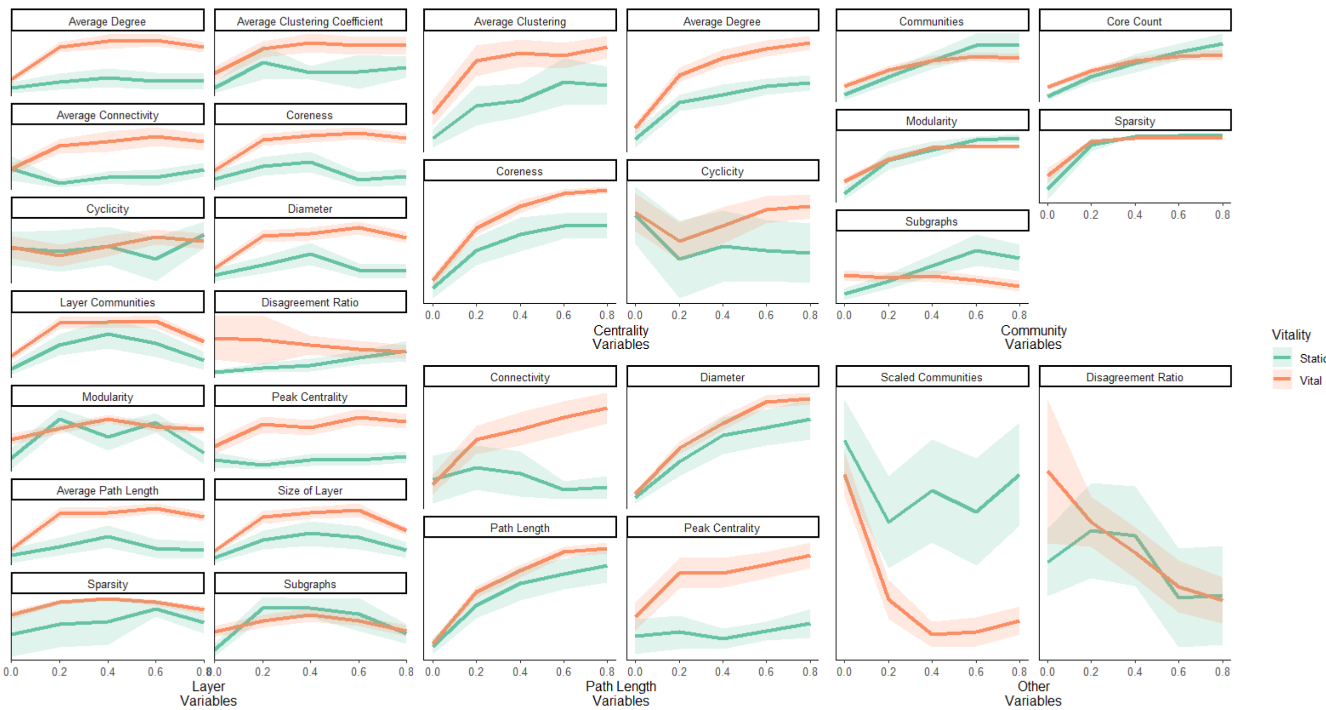

Fig. S4. Line plots of all network metrics used in the study. Shaded areas represent standard errors of the mean. Lines are grouped by network vitality condition (Static vs. Vital).

Supplementary Note 5: Assessing the reliability of historical edge lists using LLM prediction

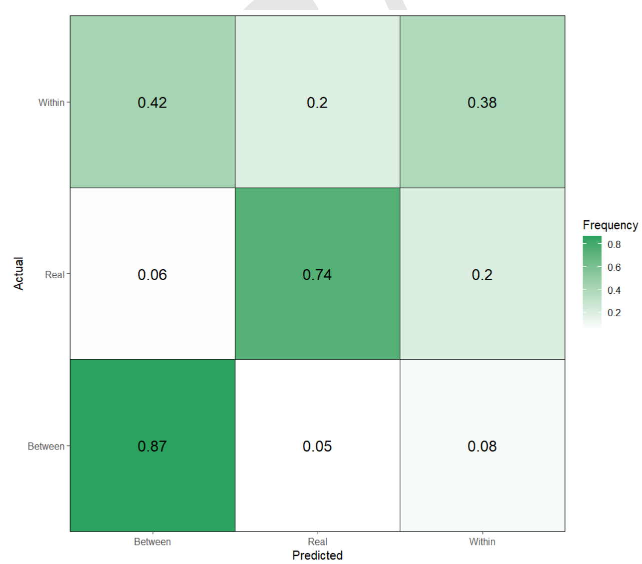

Fig. S5. To assess the reliability of the edge lists, we combined all edges between major intellectuals ($N = 884$), or those with names in the dataset, with a surrogate set of intellectual duos who were not connected. Approximately half of these surrogate duos consisted of intellectuals from within the same network ($N = 220$, “within"), with the rest from different networks ($N = 275$, “between"). We then queried a pretrained large language model (OpenAI's GPT-4o) about every pair, both real and surrogate (unconnected duos selected from within the same network or between different networks), asking whether the two intellectuals had interacted or influenced each other, belonged to the same community but had not interacted or influenced each other, or belonged to entirely different communities. There was substantial agreement between the model and the edge list we derived from Collins's text (21); (weighted Cohen's $\kappa = 0.69, p < .001$).

Supplementary Note 6: Features Table

Feature	Label (Whole/Temporal)	Description	Variable Type	Feature Cluster
Average Clustering Coefficient	average_clustering, layer_clustering	Measure of proportion of how many of a node's neighbors are connected to one another to form a complete clique.	Node average	Centralization
Average Degree	average_deg, average_deg_layer	Measure of the number of neighbors an individual node has.	Node average	Centralization
Coreness	cp_ratio, cp_layer	Graph-based measure of the number of nodes with shortest path lengths equivalent to the graph's diameter to those which are not.	Graph-based	Centralization
Cyclicity	cyclicity, cyclicity_layer	The inverse of a graph's flow hierarchy (or the fraction of edges not participating in a cycle), where a cycle is a path that starts and ends at the same node, in a graph.	Graph-based	Centralization
Sub-Communities	communities, layer_communities, comm_scale	Number of distinct communities detected in the graph using the modularity-based Louvain optimization procedure.	Graph-based	Community
Cores	cores	Graph-based measure of the number of cores in a k-core decomposition of the graph, wherein a k-core is a group of nodes that each possess at least k connections.	Graph-based	Community
Modularity	modularity, mod_layer	Measures the strength of division of a network into distinct communities based on the density of edges within communities compared to between them. Communities were defined using the Louvain optimization algorithm.	Graph-based	Community
Size of Layer	size_layer	The number of nodes in a temporal layer of a graph.	Layer-based	Community
Sparsity	sparsity, sparsity_layer	The number of edges in a network divided by the number of possible edges in the network.	Graph-based	Community
Subgraphs	subgraphs, subgraphs_layer	Number of connected components in the graph, where each subgraph is a set of nodes that are connected.	Graph-based	Community
Average Connectivity	connectivity, connectivity_layer	Measure of the minimum number of nodes between two pairs which must be removed to disconnect the pair.	Node average	Path-based
Average Path Length	path_length, pl_layer	Measure of the shortest number of edges between a given pair of nodes.	Node average	Path-based
Diameter	diameter, diameter_layer	The maximum eccentricity (the longest shortest path length) in a graph.	Graph-based	Path-based
Peak Centrality	peak_centrality, peak_centrality_layer	The highest betweenness centrality of a node in a graph, where betweenness centrality measures how often a node lies on the shortest paths between other nodes.	Graph-based	Path-based
Disagreement Ratio	disagreement_ratio, layer_disagreement	The ratio of antagonistic to non-antagonistic edges in a graph.	Graph-based	Edge-based

Table S1. Feature descriptions and types, sorted by Feature and cluster from correlation matrices. Bolded variables are those used in the Principal Components Analysis.