



OPEN Functional and structural clustering of social relationship layers among college students for link prediction with applications to perceived drinking networks

Eun Lee^{1,18}✉, Ovidia Stanoi^{2,6,18}, Xie He³, Yoona Kang⁴, Mia Jovanova⁵, Amanda L. McGowan^{3,6,7,8}, David M. Lydon-Staley^{6,7,8}, Dani S. Bassett^{9,10,11,12,13,14}, Kevin N. Ochsner², Zachary M. Boyd¹⁵, Emily B. Falk^{6,16,17} & Peter J. Mucha³✉

Studies on college drinking—a behavior associated with health risks and reduced productivity—often rely on broadly defined peer or friendship networks. Yet, friendship networks can be further divided into more specific relational types, and the associations between such diverse social ties and drinking ties remain poorly understood. This study adopts a multilayer network framework to conduct a fine-grained examination of ten distinct types of social networks (i.e., layers)—including friendship, leadership, emotional support, and perceived drinking—by analyzing their structural similarity across three levels (node, link, and triad) and along a relationship-axis spectrum. We cluster networks based on these multi-level similarities, identifying three primary clusters (Affiliation, Leadership, and Drinking) and assess whether these structurally coherent clusters contribute to improved link prediction performance in perceived drinking networks. Both *k*-means clustering based on multi-level structural similarity and projection-based analysis along the hierarchical–horizontal spectrum revealed that perceived drinking nominations are structurally closest to leadership layers, followed by emotional support layers within the Affiliation cluster. Assessing whether these functional clusters can improve link prediction in perceived drinking networks, we find that layers within the same cluster yield predictive performance close to that of models using all layers. Notably, emotional support layers, which may reflect their structural proximity and distributed connectivity, offer the highest link prediction accuracy for drinking ties. Taken together, these findings demonstrate that coherent clusters of fine-grained, functionally and structurally aligned social layers facilitate more efficient inference in sparse or partially observed drinking-related networks. This, in turn, highlights their potential utility in predicting and preventing health-risk behaviors—such as alcohol use—that are typically difficult to observe or measure directly, and clarifies the types of relationships most strongly associated with such behaviors.

¹Department of Scientific Computing, Pukyong National University, Nam-gu, Busan 48513, South Korea. ²Department of Psychology, Columbia University, New York, NY 10027, USA. ³Department of Mathematics, Dartmouth College, Hanover, NH 03755, USA. ⁴Department of Psychology, Rutgers, The State University of New Jersey, Camden, NJ 08102, USA. ⁵School of Medicine, University of St. Gallen, 9000 St Gallen, Switzerland. ⁶Annenberg School for Communication, University of Pennsylvania, Philadelphia, PA 19104, USA. ⁷Department of Bioengineering, School of Engineering and Applied Sciences, University of Pennsylvania, Philadelphia, PA 19104, USA. ⁸Leonard Davis Institute of Health Economics, University of Pennsylvania, Philadelphia, PA 19104, USA. ⁹Department of Bioengineering, University of Pennsylvania, Philadelphia, PA 19104, USA. ¹⁰Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104, USA. ¹¹Department of Neurology, University of Pennsylvania, Philadelphia, PA 19104, USA. ¹²Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA. ¹³Department of Psychiatry, University of Pennsylvania, Philadelphia, PA 19104, USA. ¹⁴The Santa Fe Institute, Santa Fe, NM 87501, USA. ¹⁵Department of Mathematics, Brigham Young University, Provo, UT 84602, USA. ¹⁶Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104, USA. ¹⁷Wharton Marketing Department, University of Pennsylvania, Philadelphia, PA 19104, USA. ¹⁸Eun Lee and Ovidia Stanoi contributed equally to this work. ✉ email: eunlee@pku.ac.kr; Peter.J.Mucha@dartmouth.edu

Alcohol use among college students poses a significant public health concern, contributing to increased risks of injury and chronic health problems, as well as impairments in academic performance and productivity^{1–3}. To mitigate the risks associated with alcohol use, it is important to understand the factors related to drinking behaviors among college students. Multiple studies on alcohol use in college students have found that drinking behaviors are heavily influenced by social relationships^{4–9}, defined by ‘friends’ or ‘peers’ broadly. However, peer relationships can be categorized into various types depending on their underlying social purposes. For instance, you might consider one person your mentor (leader), and go to someone else when you have a hard time and want to talk (emotional supporter); that is, a person’s perceived leadership network and emotional support network might be configured with different ties (though of course, the same individual might serve multiple functions).

Given this diversity of social ties, one may ask: does knowing the structure of a person’s emotional support or leadership network also tell us something about the shape of their perceived drinking network? This question concerns whether certain types of social ties are more strongly associated with alcohol use. Prior evidence suggests that individuals tend to respond more to alcohol-related posts from close connections in their online networks, offering a potential clue in this regard⁹. Building on this, a more granular classification of perceived social networks may help identify specific relationship types that are more closely associated with, or reflective of, patterns of alcohol consumption. Such a perspective aligns with theoretical frameworks like Primary Socialization Theory (PST), which, in this context, highlight how alcohol-related behaviors might develop through primary social group interactions¹⁰. While our analysis does not aim to test PST directly, our findings regarding the alignment of certain social layers (e.g., emotional support or leadership) with perceived drinking behaviors may offer structural insights into the peer dynamics underlying socialization processes.

While several studies have examined the structural similarities among various types of social ties—often using node characteristics, centralities, edge overlap, or entropy measures—most of these focus on generalized or online networks^{11–17}, such as smartphone proximity data¹¹ or online gaming activity¹³. Consequently, the structural similarities among offline multi-social layers (i.e., networks) have not been thoroughly examined, which has in turn limited our understanding of how specific social relationships are associated with perceived drinking-related networks¹⁸.

To address this gap and understand how different types of social relationships relate to college students’ perceived peer drinking behaviors, we examined the following research questions using a survey dataset on offline social networks among college students. First, do social networks that are structurally similar—based on systematic comparisons of network features across different types of social relationships—also exhibit functional similarity in terms of the roles they play within the group? Second, how are structural and functional similarities among social networks associated with perceived drinking behavior within peer groups? Third, can structural similarities across different types of social relationships be leveraged to predict missing links or infer unobserved attributes associated with peers’ perceived alcohol consumption?

Grounded in these research questions, we analyzed diverse social relationship types surveyed within nine college student groups (e.g., clubs, teams, etc.), including perceived drinking nominations (351 participants across nine student groups started the survey, 336 participants finished the survey [69 males, 266 females, one unknown]; see Methods). To ensure that we captured a broad range of interpersonal dynamics, we included ten peer nomination items reflecting affective, behavioral, and structural ties. These items were informed by prior network studies on emotional cohesion, social support, leadership, and behavioral norms^{19–21}, and have been widely used in peer network research. In the survey, students were asked to nominate group members in response to each of these ten questions—such as “Which group members do you like the most?” (Like), “Who would you say are the leaders of the group?” (Leaders), and questions regarding perceived alcohol consumption—in order to examine how perceived drinking networks are coupled with other layers of social connection (see Table 1 in Methods for all questions).

All peer nominations for each type of social relationship were represented as separate networks (i.e., layers), where links indicate the presence of a nomination between individuals (i.e., nodes). We refer to each of these relationship-specific networks—such as the network formed by nominations from the “Like” question or the network from the “Leaders” question—as a network layer to adopt the language of multilayer network analysis. Each layer corresponds to a single question in our social network survey and captures one distinct aspect of group interactions (see Fig. 1a, b). To answer the first research question, we analyzed the structural characteristics of

Abbreviation	Question
Like	"Which group members do you like the most?"
Closest	"Which group members are you closest to?"
Interact	"Which group members did you interact with most recently?"
Good News	"Which group members are you most likely to go to when something good happens to you?"
Bad News	"Which group members are you most likely to go to when something bad happens to you?"
Offer Support	"Which group members are most likely to come to you when something bad happens to them?"
Leaders	"Who would you say are the leaders in this group?"
Influence	"Who would you say are among those with the most influence over the other members?"
Drinking Most	"Which group members drink the most alcohol?"
Drinking Least	"Which group members drink the least alcohol?"

Table 1. Questions and their abbreviations for ten social relationships of college student groups.

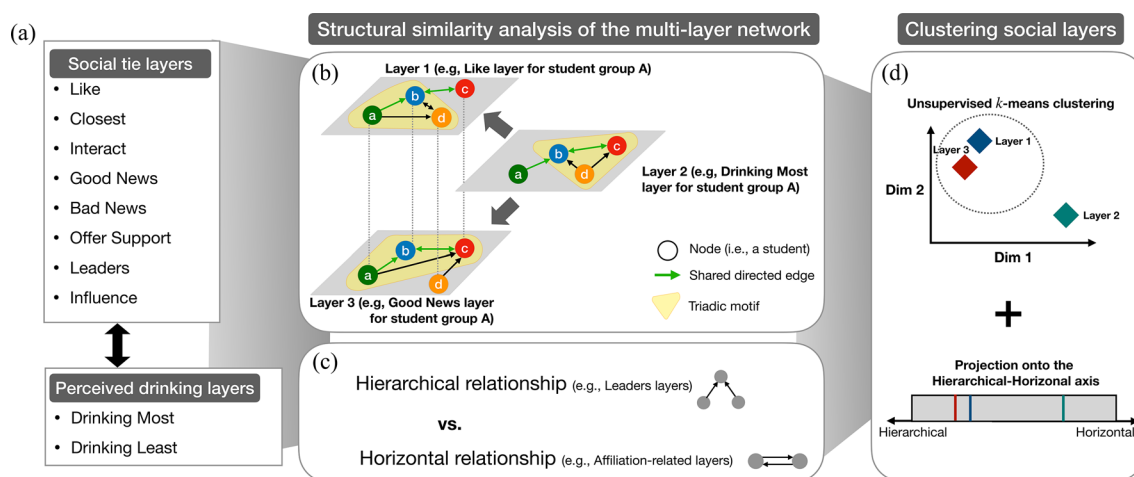


Fig. 1. Overview of the structural similarity analysis and clustering of social network layers. (a) Ten peer-nominated social layers, including two layers related to perceived drinking behavior, were collected from college student clubs. (b) Example illustration of three directed network layers (e.g., Like, Drinking Most, and Good News) representing different social relationships within a single student group. Each layer represents a specific type of social network (e.g., Like, Drinking Most, and Good News) and contains directed links between students (nodes), based on peer nominations. Green arrows indicate directed ties that are shared across multiple layers, and yellow highlights mark local triadic structures. The elements used for node-, edge-, and triad-level analysis are shown. (c) Social layers are positioned along a conceptual axis from hierarchical (e.g., Leaders layer) to horizontal (e.g., Affiliation-related layers), reflecting the degree of directionality and reciprocity in each network. (d) Node-, edge-, and triad-level similarity measures are used to perform unsupervised *k*-means clustering of network layers. Structural similarity scores are projected onto the first two dimensions (Dim 1 and Dim 2) of a low-dimensional embedding (e.g., PCA), where Dim 1 and Dim 2 capture dominant patterns of structural similarity across layers. Each layer is also positioned along a conceptual hierarchical–horizontal axis. The alignment between these two independent perspectives—an unsupervised clustering and a projection onto a conceptual relationship axis—highlights groupings of social layers that are both structurally and functionally similar.

each layer by calculating network measures at three levels, which were then compared across layers (Fig. 1b). Such measures include evaluating individuals' positional importance through centrality measures (i.e., node-level), counting common links between layers (link-level), and identifying triadic motifs (triad-level) between different layers. These features were then grouped into structurally similar clusters using unsupervised *k*-means clustering (Fig. 1d).

In addition to comparing structural similarity across layers, we further examined how these social relationships vary in their underlying group structure by projecting each layer onto a spectrum ranging from hierarchical (vertical) to horizontal (egalitarian) configurations as shown in Fig. 1c, d²². This distinction reflects how different relationship types are structured to meet different social needs: hierarchical relationships often involve clear directionality and ordering of roles, whereas horizontal relationships are characterized by reciprocity and mutual exchange. Understanding these structural dimensions offers a theoretical basis for interpreting the positioning of different layers within the spectrum and the behavioral dynamics they may represent. The clusters identified through *k*-means clustering and projection-based classification were subsequently used to evaluate

how well each social layer could predict missing links in drinking-related relationships. This analysis, which addresses the second and third research question, was conducted under the assumption that prediction accuracy would improve when more strongly correlated social layers are incorporated.

Collectively, we aim to establish a framework for identifying structural and functional similarities across multiple layers of social relationships and to examine how these structural characteristics relate to perceived alcohol-related networks among college students. By doing so, we assess whether structurally and functionally similar social layers can be leveraged to infer missing information in drinking-related networks, thereby enhancing our understanding of peer influence dynamics and informing future intervention strategies.

Methods

Multilayer social network data from college student groups

We initially recruited 15 student clubs from the University of Pennsylvania and Columbia University, offering an Amazon gift card as an incentive for participation through the Social Health Impact of Network Effects (SHINE) study²³. To ensure data quality, we included only those groups in which more than 25 participants completed at least 60% of the survey, resulting in a final sample of 351 participants across nine student groups. All research, methods, and study protocols were approved by the Human Subjects Electronic Research Application (HSERA) Institutional Review Board (IRB) at the University of Pennsylvania and were acknowledged by the Human Research Protection Office of the Department of Defense, with reliance agreements and local context review approvals at Columbia University. All research, methods, and study protocols were conducted in accordance with the Human Subjects Institutional Review Board (IRB) at the University of Pennsylvania and the Human Research Protection Office of the Department of Defense. All methods were performed in accordance with relevant guidelines and regulations. All participants provided informed consent before taking part in the study and were financially compensated.

Age and gender distributions are shown in Supplementary Figure S1. For the present analysis, we selected nine student groups with more than 25 members, ensuring that over 60% of their members had completed the survey to maintain the group members' representativeness. These groups have different foci (e.g., sports teams, music, dance performance), differ in size from 25 to 87 people, and include four woman-only and five mixed-gender groups. Participants were compensated \$20 (via Amazon gift card) for completing a social network survey and an additional \$20 (via gift card) if more than eighty percent of their student group completed the survey. While our sample was limited to student organizations from two private universities in the northeastern United States, the groups represented a range of club types (e.g., academic, social, athletic), gender compositions, and group sizes. This diversity offers some variation in social contexts, which may support cautious generalization to other college settings with similar organizational structures. However, further research is needed to evaluate whether these findings hold in broader institutional or cultural contexts.

Building on previous research indicating that the social life of a group can be mapped through multiple relational dimensions—as individuals adopt different roles within the group^{24,25}—we collected participants' peer nominations in response to ten relational questions designed to capture these various aspects. Prior work in social network analysis highlights the importance of distinguishing among affective (e.g., liking, closeness), behavioral (e.g., interaction, support), and structural (e.g., leadership, influence) ties to better understand the complexity of group interactions²⁵. For instance, the "Like" and "Closest" items assess affective bonds that are central to emotional cohesion¹⁹, while "Good News" and "Bad News" reflect asymmetrical trust and social support-seeking behaviors²⁵. Items such as "Leaders" and "Influence" capture power dynamics and perceived centrality in group decision-making^{20,26}. In addition, "Drinking Most" and "Drinking Least" reflect health-related behavioral norms within peer groups, which are particularly relevant to studies on college socialization and risky behavior. Although previous studies did not use the same set of items, they consistently emphasize the importance of capturing diverse types of peer relationships in adolescent and young adult networks. Motivated by these insights, we asked ten peer nomination questions to assess multiple dimensions of social relationships within peer groups. The ten survey questions are presented in Table 1.

Participants were given the option to select as many group members as they felt were appropriate in response to each question, but we excluded from our analyses self-nominations or nominations of members who did not complete the social network survey. We then aggregated participants' answers to these questions to specify ten social networks for each group. In each network, a directed link from individual A to individual B exists if A nominates B on the corresponding relationship question. For instance, if Aspen nominates Bailey on the question "Which group members are you closest to?", a link is created from Aspen to Bailey in the Closest social network of their group. The basic network characteristics are summarized in Supplementary Table S1.

Node-level centrality correlations

One can hypothesize that participants' positions would demonstrate overlap across functionally similar layers²⁷. To evaluate this, we computed node-level similarity by comparing centrality values across layers while preserving node identities. Specifically, we measured the in-degree, eigenvector, and betweenness centrality for each individual in each social network layer, treating each layer as a separate single-layer network, in order to capture each individual's relative prominence or positional importance within different types of social ties. These centralities were computed using the *python-NetworkX* library, and Spearman's rank correlation coefficient was calculated using *python-scipy*. The Spearman correlation coefficient quantifies the similarity in rank order of centrality values across layers, with a value of 1.0 indicating perfect rank agreement between two layers.

In this context, in-degree reflects the number of nominations received by each node—indicating how many peers identify a person in a given role (e.g., how many people view someone as a leader or emotional supporter), which can be interpreted as perceived popularity. Eigenvector centrality captures how well-connected a person is to other well-connected individuals, and can be interpreted as their overall influence within the group—for

example, someone whom influential peers also tend to nominate. Betweenness centrality quantifies how often an individual lies on the shortest paths between others, serving as a proxy for brokerage or gatekeeping—for instance, a person who connects otherwise separate subgroups within the network.

Link-level similarity

To measure the similarity of the edges between two different layers of a social group (e.g., Like and Closest), we calculate the Jaccard similarity between the pair of layers. Here, $J(L_{g,\alpha}, L_{g,\beta})$ represents the Jaccard index between two layers, α and β , in student group g . As one means of performing this calculation, we use indicators on all ordered node pairs specified by two binary vectors $E_{g,\alpha}$ and $E_{g,\beta}$, where a 1 entry indicates the presence of the corresponding directed edge in a given layer while a 0 entry represents its absence. Equivalently, $E_{g,\alpha}$ indicates the edge list of layer α for student group g . The Jaccard index is then given by

$$J(L_{g,\alpha}, L_{g,\beta}) = \frac{|E_{g,\alpha} \cap E_{g,\beta}|}{|E_{g,\alpha} \cup E_{g,\beta}|}. \quad (1)$$

That is, the Jaccard index measures the number of directed edges that the two layers $L_{g,\alpha}$ and $L_{g,\beta}$ share with each other (size of intersection) relative to the total number of directed edges in their union. This value reflects the extent to which the same directed ties—that is, the same student nominating the same peer—occur in both layers (e.g., whether student A nominates student B as both a “Good news” and “Bad news” friend). After calculating the Jaccard similarity indices for all possible pairs of the ten layers for each of the nine college student groups, we averaged the results over student groups to obtain a single Jaccard similarity matrix for a visual representation of the similarities between layers. Finally, we applied standard hierarchical clustering with the cluster map function in the *Seaborn* package of Python to identify relevant layer groupings, with results presented in Fig. 2b.

Triad census analysis

The triad census is useful for summarizing social relationship properties in directed networks²⁸. Various other characteristics of the network can be inferred from the triad counts—including network density, transitivity, and dyadic properties—and they also inherently give more involved information characterizing the different transitive patterns of pairwise interactions possible between three actors²⁹. In our context, this triad census analysis helps distinguish how different types of social ties (e.g., Like vs. Leaders) differ in structural patterns such as reciprocity, hierarchy, or transitivity. A full triad census includes counting the instances of the sixteen possible three-node motifs in a directed graph as shown in Fig. 3a. The label of each triad motif is written with three integers (indicating the number of reciprocal links, unidirectional links, and non-connected node pairs in the motif) and sometimes with an additional letter among $\{C, D, U, T\}$ to indicate Cyclic, Down, Up, and Transitive patterns, respectively (again, see Fig. 3a). We then applied the Triad Significance Profile (TSP) methodology suggested by Milo et al.³⁰ to obtain a normalized vector of Z-scores which can be written as

$$\text{TSP}_j = \frac{Z_j}{\sqrt{\sum Z_j^2}}, \quad (2)$$

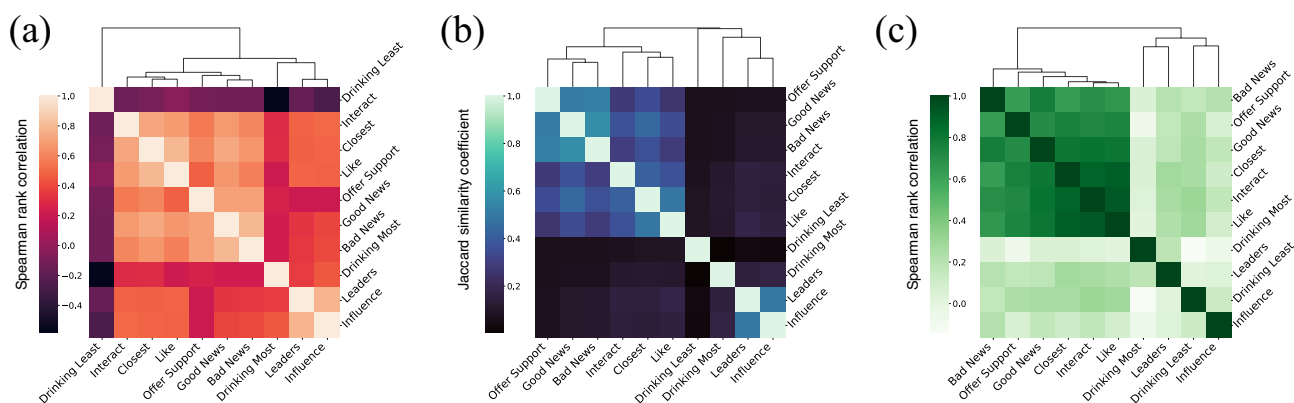


Fig. 2. Averaged layer-to-layer matrices for node-, link-, and triad-level similarities. (a) Spearman rank correlation coefficients of node in-degree between each pair of layers. (b) Jaccard similarity coefficients of overlapped edges between pairs of layers (i.e., pair counting directed edges that appear in both, one, or the other). (c) Spearman rank correlation coefficients of the 16 parts of the Triad Significance Profile (TSP) between each pair of layers. All coefficients were computed separately within each student group and then averaged across the nine student groups. Dendrograms were computed using *clustermap* in the *Seaborn* Python library. In all panels, colors closer to 1.0 indicate stronger similarity or correlation between layers. For example, in (a), a lighter shade between the “Like” and “Closest” layers indicates that students who receive many nominations in the “Like” layer also tend to receive many in the “Closest” layer.

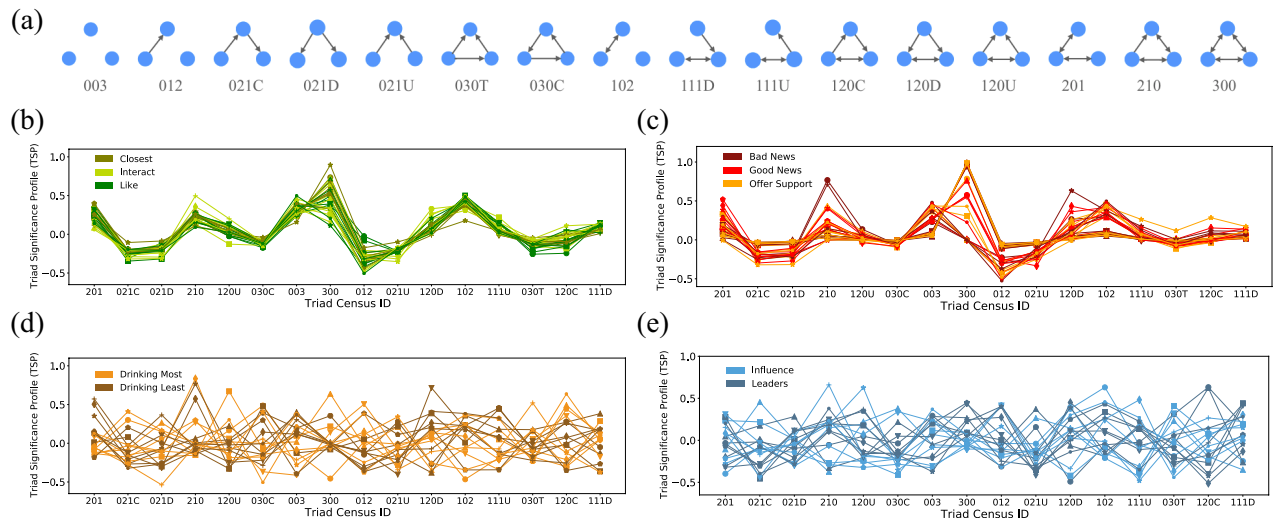


Fig. 3. Structural comparison of social layers using triad significance profiles. Triad Significance Profiles (TSP) of a layer in each student group are constructed from counts of the 16 possible directed triad motifs visualized in (a). TSPs of each layer are plotted here and grouped into the (b) friendship cluster (closest, interact, like), (c) emotional support cluster (bad news, good news, offer support), (d) drinking nominations (drinking most and drinking least), and (e) leadership cluster (influence and leaders layers). Each colored line represents the TSP of a single student group for the corresponding social network layer, capturing the triadic structure unique to that group's social context.

for each layer in a student group, where Z_j is the Z-score of triad motif j (in layer $L_{g,\alpha}$) relative to the counts of that motif in randomly rewired versions of the network layer. That is, $Z_j = (N_j - \langle R_j \rangle) / \sigma(R_j)$, where N_j denotes the count of triad motif j in the empirical network, R_j is a set of random realizations of this count (in our case, from an ensemble of 50 randomly rewired networks with the same degree sequence), angled brackets indicate the average and $\sigma(\cdot)$ the standard deviation of the set. The denominator in Eq. (2) ensures the collection of TSP scores for a given empirical network is normalized to length 1 (in a sum of squares sense).

Unsupervised k -means clustering based on multi-level similarity features

We average the node, link, and triad similarities of a student group by layers to define a collective feature associated with each social network layer α for the student group g . Specifically, we averaged together each group's structural similarities corresponding to the three matrices in Fig. 2—for node-level similarity based on the Spearman correlation (in-degree, eigenvector, betweenness centrality similarity), link Jaccard similarity, and TSP similarity, so each student group g has 10×10 matrix of the averaged similarity. We then applied Principal Component Analysis (PCA) to this $9 \times 10 \times 10$ matrix (10 layers for each of the 9 groups, each with 10 features). When we check the scree plot by varying the number of PCA components (see Fig. 4a), we confirm that the first two components together explain around 69% of the variance, while the first three components together explain around 77% of the variance. Reducing the similarity matrix to these three PCA directions, we subsequently applied k -means unsupervised clustering to further visualize how multi-level similarity can capture the major differences between the Leadership, Affiliation, and Drinking layers. We used five clusters for the visualization in Fig. 4c–e.

Cosine similarity along the hierarchical-to-horizontal axis

We select the closest layers of the nine student groups to represent the most 'horizontal' of our observed network layers (notated here as $\mathcal{X}^+ = \{\mathbf{v}_1^+, \mathbf{v}_2^+, \dots, \mathbf{v}_8^+, \mathbf{v}_9^+\}$), and similarly select the Leaders layers to represent the 'hierarchical' aspect ($\mathcal{X}^- = \{\mathbf{v}_1^-, \mathbf{v}_2^-, \dots, \mathbf{v}_8^-, \mathbf{v}_9^-\}$). We then measure the centroid position of each by averaging the corresponding representation vectors over the nine student groups: $\bar{\mathbf{V}}^+ = \frac{1}{9} \sum_{i=1}^9 \mathbf{v}_i^+$ and $\bar{\mathbf{V}}^- = \frac{1}{9} \sum_{i=1}^9 \mathbf{v}_i^-$. We then define the Hierarchical-to-Horizontal axis vector by subtracting the centroid of the hierarchical (Leaders) vectors from the centroid of the horizontal (Closest) positions: $\mathbf{v}_{\text{axis}} = \bar{\mathbf{V}}^+ - \bar{\mathbf{V}}^-$. To examine the composition along this axis, we calculate the cosine similarity between the axis vector and a vector \mathbf{v}_α of a social network layer as $\mathcal{S}_\alpha = \mathbf{v}_\alpha \cdot \mathbf{v}_{\text{axis}} / (|\mathbf{v}_\alpha| |\mathbf{v}_{\text{axis}}|)$, where α denotes a social layer of the Closest, Interact, Like, Good News, Bad News, Offer Support, Influence, Leaders, Drink Most, or Drink Least question in a selected student group.

Link prediction with functional clusters

Our link prediction calculations employ a stochastic block model method developed by DeBacco et al.³¹ that uses expectation maximization to compute the likelihood of an edge existing in a particular layer. Specifically,

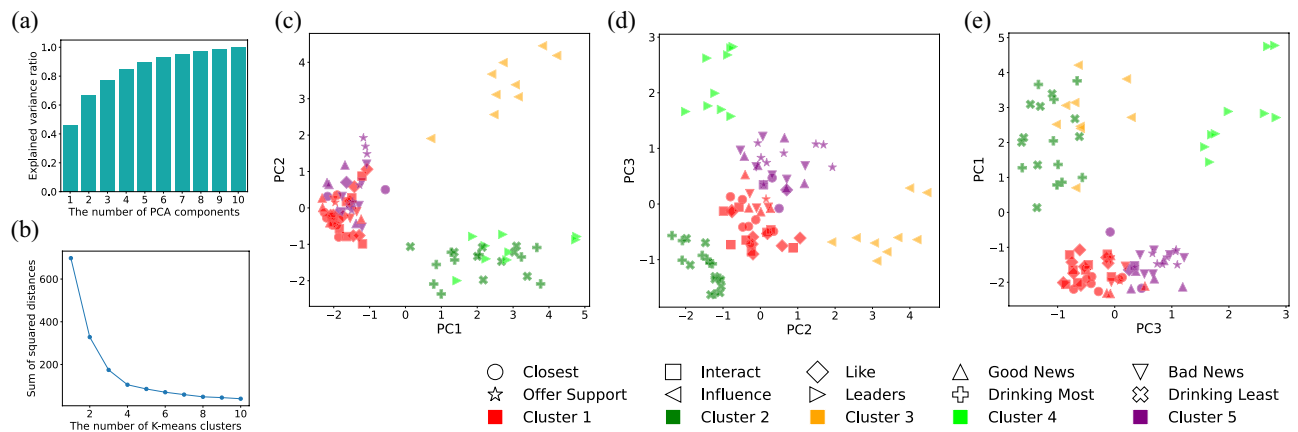


Fig. 4. Clustering of social network layers based on multi-level structural similarity using PCA and k -means. Unsupervised k -means clustering is applied to classify 10 different social network layers in each of the nine student groups. Node-, link-, and triad-level similarities are averaged for each pair of layers within a student group, separately. Principal Component Analysis (PCA) is used to reduce the 10 resulting similarity dimensions to three components. The scree plot in (a) visualizes the amount of variance explained over different numbers of components. Unsupervised k -means clustering scores on these three PCA components with different numbers of clusters are plotted in (b). The results of clustering into five groups are shown in (c–e), projected onto different combinations of the first three PCA axes.

$$M_{ij}^{\alpha} = \sum_{k,l=1}^K u_{ik} v_{jl} w_{kl}^{\alpha}, \quad (3)$$

where M_{ij}^{α} represents the likelihood of the existence of a directed edge between individuals i and j in layer α , obtained by a sum over blocks k and l with mixing matrix w_{kl}^{α} for layer α , and the membership vectors for node i in block k and node j in block l given by u_{ik} and v_{jl} , respectively (which can be different for edges in and out of the nodes; an undirected network would have $u = v$). This model requires knowledge of or other means of specifying the number of blocks, K , to be used in the link prediction. Because we are focused on the *potential* for accurate link prediction, we explore a range $5 \leq K \leq 11$ and in each case select the K with the highest AUC (as in ³¹) under 5-fold cross-validation (hiding 20% of the edges in the adjacency matrix at a time). In particular, K is selected case by case, specific to each prediction performed; that is, K could (and we expect it to) be different for different target-support combinations of layers. In this study, link prediction estimates how well the structure of one layer can inform or recover the missing links in another, such as using emotional support nominations to infer perceived drinking ties.

Results

We collected social network data on nine college student groups across ten questions listed in Table 1 (see Methods). We represented each student group as a multilayer network, with each node representing an individual member and with directed links between nodes in a layer encoding a specific social relationship according to the nominations on the corresponding survey question (e.g., friendship, leadership, etc.; see Methods). To compare and contrast layers (questions), we consider three different structural levels: (i) how individuals' positions differ across layers (node-level similarity), (ii) whether relationships between pairs of individuals co-occur in different layers (link-level similarity), and (iii) how measures on node triples correlate across layers (triad-level similarity). Across all three levels, we find consistent functional clusters of layers based on their structural similarities corresponding to identifiable three-level structural similarities. The consistency of these layer/question clusters is further supported by unsupervised clustering on a principal component representation of the full set of similarity comparisons, as well as by the projection-based clusters along the hierarchical-horizontal relationship axis. Finally, we show how they impact the potential for accurate link prediction, with a special focus on predicting missing links in the alcohol-related layers of our survey data.

Node-level and link-level similarity

Using the node-level similarity (see Methods), we hierarchically clustered the average in-degree correlations (averaged across student groups) between layers, revealing two primary clusters: an Affiliation cluster—comprising Closest, Like, Interact, Good News, Bad News, and Offer Support—and a Leadership cluster, consisting of Leaders and Influence layers, as shown in Fig. 2a. Within the Affiliation cluster, we further observed a subdivision into two sub-clusters: one composed of Good News, Bad News, and Offer Support (which we refer to as the Emotional Support cluster), and the other including Closest, Like, and Interact (the Friendship cluster in this study). Although the specific hierarchical structure slightly varies depending on the centrality measure used, the clustering into Emotional Support, Friendship, and Leadership groups remains consistent (see Fig. S2).

These results suggest that individuals occupy clearly differentiated roles within Affiliation and Leadership clusters in the social networks of the surveyed college student groups. Moreover, even within the broader Affiliation category, we observe distinct structural node-level characteristics between layers associated with everyday social interaction—such as Closest and Like—and those reflecting emotionally supportive relationships, such as Good News, Bad News, and Offer Support. This finding highlights a meaningful divergence in node-level roles between friendship-oriented and emotional support-oriented ties, even though both stem from perceived peer relationships.

In terms of the relationships with other social networks, the Drinking Least layer (nominates related to the question of “Which group members drink the least alcohol?”) has very little node-level similarity with any other layer (see Fig. 2a and Figure S2), whereas the Drinking Most layer (nominates related to the question of “Which group members drink the most alcohol?”) is closely grouped with the Leadership cluster (the Leaders and Influence layers). These findings suggest a potential alignment between social status and perceived drinking behavior, with individuals seen as leaders or influential figures also more frequently nominated as heavy drinkers.

Building on the patterns observed at the node-level similarity, we found that similar clustering structures also emerged at the link-level similarity. As shown in Fig. 2b, hierarchical clustering of the average Jaccard coefficients again revealed a Leadership cluster and an Affiliation cluster, with the latter further subdividing into Friendship and Emotional Support layers—mirroring the structure observed in the node-level analysis. Within the Leadership cluster, the strong overlap between the Leaders and Influence layers is reflected in a high Jaccard score of 0.49, indicating that if a participant nominates someone as a leader, they are also likely to nominate the same person as influential. At the same time, we observe notable variability in edge overlap across other functional clusters, with Jaccard scores ranging from 0.5 to 0.57 within the Emotional Support cluster, and from 0.37 to 0.49 within the Friendship cluster. These patterns highlight both the internal coherence of functional clusters and the nuanced differences in tie formation within and across them.

In addition to the main functional clusters, we also observe notable characteristics regarding the placement of the Drinking Most layer. Although its average Jaccard similarity with the Leaders and Influence layers is relatively low—0.15 and 0.17, respectively—it still clusters with them, likely due to its even lower similarity with other layers. This shows that while Drinking Most does not strongly overlap with any one layer, its structural position is more closely aligned with leadership-related ties than with affiliative ones. To further explore these relational patterns beyond dyadic overlap, we next turn to triadic motif analysis, which captures higher-order structural features between layers.

Triad-level similarity

While node- and link-level analyses revealed consistent clustering across social layers, they primarily capture pairwise relationships or individual-level prominence. To further investigate higher-order structural patterns, particularly those that may distinguish nuanced differences in group dynamics, we applied triadic motif analysis using the Triad Significance Profile (TSP)³⁰ (see Methods).

Triadic motifs offer insight into the underlying relational dynamics of each network layer, particularly in terms of reciprocity and local transitivity. For example, we hypothesized that layers in the Affiliation cluster would exhibit greater reciprocity and triangle closure, as individuals in these layers are more likely to form mutual or transitive ties (i.e., friends of friends tend to be friends). This expectation is confirmed in the TSPs of the Friendship and Emotional Support clusters (see Fig. 3b, c), which show increased prevalence of motifs such as ‘201’, ‘210’, ‘102’, and ‘300’, all of which reflect mutual nominations and closed triads. In particular, the Friendship cluster demonstrates strong overrepresentation of these motifs, consistent with the hypothesized highly reciprocal structure.

While Emotional Support layers show similar patterns, they also exhibit relatively higher proportions of motifs such as ‘210’ and ‘120D’, which indicate hierarchical structuring in triads³². This suggests that emotional support ties may involve an implicit sense of directionality or social differentiation, distinguishing them from more egalitarian friendship networks. Notably, these motif patterns are consistently observed across nearly all student groups, as shown by the tightly aligned TSP profiles in both the Friendship and Emotional Support clusters (see Fig. 3b, c). This cross-group consistency further supports the structural distinctiveness and functional coherence of these relationship layers.

In contrast to the consistent TSP patterns observed in the Friendship and Emotional Support clusters, no clear motif-level regularities were found in the layers of the Leadership cluster. As shown in Fig. 3e, the Leaders and Influence layers display considerable variability in their triad profiles across student groups, with no dominant motifs emerging. A similarly scattered pattern is observed in the Drinking Most and Drinking Least layers (Fig. 3d), suggesting a lack of coherent triadic structure in how students perceive these roles. One possible explanation is that these layers are shaped by highly localized or asymmetric nomination behaviors—often centered around a few prominent individuals—making them less likely to exhibit consistent motif signatures across groups. Additionally, the TSP method may be less effective in detecting triadic patterns in sparse or hub-dominated layers.

Unsupervised *k*-means clustering of structural similarities

To gain a more integrated view of how social layers relate to one another, we applied unsupervised *k*-means clustering to the combined node-, link-, and triad-level similarity measures for each layer within each student group (see Methods). As shown in Fig. 4c–e, the clustering based on the combined multi-level similarity data closely aligns with the hierarchical clustering results observed at each separate level of structural analysis. Specifically, we assigned each social layer in a student group to one of five clusters identified through *k*-means clustering, which clearly captures the separation between Leadership (Clusters 3 and 4) and Affiliation (Clusters 1 and 5) layers (see Fig. 4). Furthermore, the earlier subdivision of the Affiliation cluster into Friendship and

Emotional Support sub-clusters is reflected in the separation between Clusters 1 and 5, which appear close to each other in the low-dimensional projection. In contrast, the two Drinking layers form their own distinct group, constituting Cluster 2.

While the first principal component (“PC1”) distinguishes Affiliation from Drinking and Leadership clusters, “PC2” and “PC3” help further differentiate between the subtypes of Leadership and Drinking. Nevertheless, the Drinking layers remain structurally closer to the Leadership cluster, with an average Euclidean distance of 3.41 in the three-dimensional PC space, compared to a greater distance of 4.17 from the layers in the Affiliation cluster. This structural proximity between drinking and leadership layers, alongside their relative distance from affiliation layers, suggests that perceptions of drinking—particularly heavy drinking—may not merely reflect social affinity, but may also serve as a signal of social status or visibility within peer groups, rather than indicating everyday closeness.

Projection to the two general relationship axes of social relations: hierarchical and horizontal

To complement the clustering of layers based on three levels of structural similarity, we next examine how social relationships differ in their underlying organizational patterns—specifically, whether they are more *hierarchical* or *horizontal* in nature. To this end, we project each layer into a reduced structural feature space to determine its position along the hierarchical–horizontal relationship axis (see Methods). In Fig. 5a, the horizontal end of the axis is dominated by social layers characterized by mutual relationships, while the other end of the axis includes hierarchical relationships primarily with asymmetric interdependent patterns. The spectrum for a randomly selected student group in Fig. 5a identifies most of the Affiliation-related layers in the horizontal part of the spectrum, whereas the Leadership and Drinking clusters are placed in the hierarchical part of the spectrum.

Figure 5b displays each social layer from all student groups positioned individually along this spectrum. These differences in projection suggest that layers involving support or emotional disclosure—such as offering support (i.e., Offer Support layer) or sharing bad news (i.e., Bad News layer)—tend to be structurally more

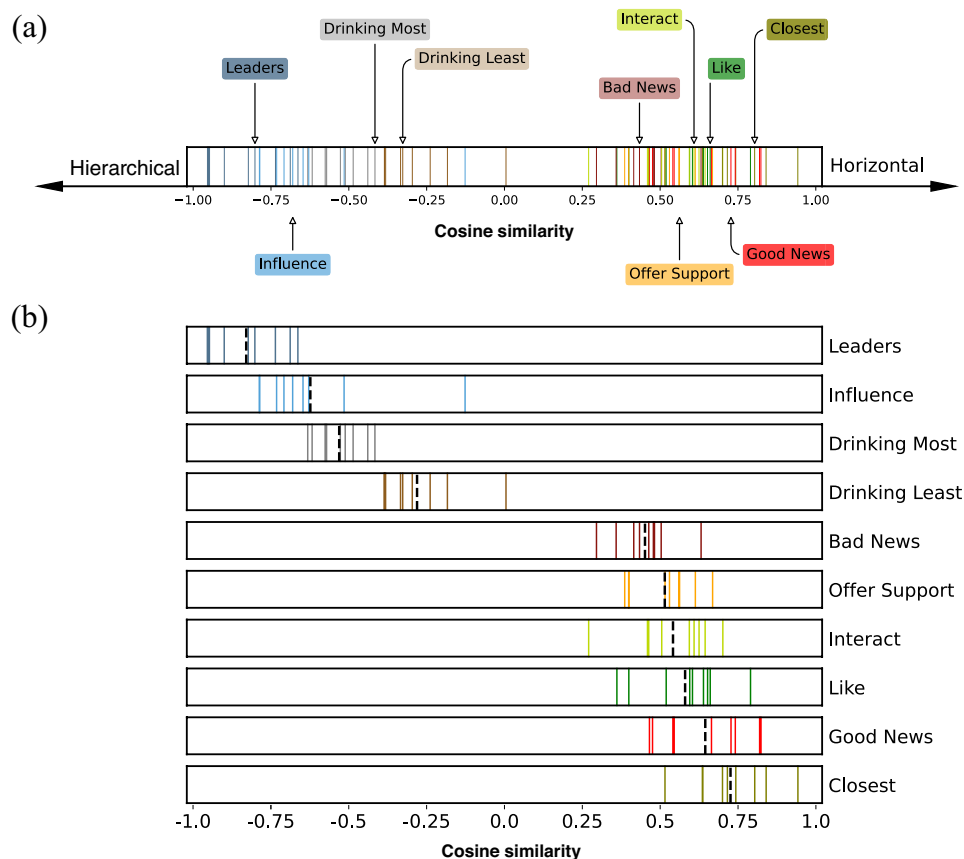


Fig. 5. Projections of social layers onto the hierarchical and horizontal axes of relational types. (a) Each of the 10 layers from the nine student groups is projected onto a hierarchical–horizontal axis, defined by the line connecting the centroid of the Leaders layers and the centroid of the Closest layers, using cosine similarity. A single selected student group’s projected social layers are indicated here by the arrows and annotations as an example. (b) The projection spectrum of each layer type is shown separately, with colored lines indicating the position for each group and black dashed lines representing the average projection across all groups. By comparing this result with that depicted in Fig. 4, we note that qualitatively similar spectra would also be obtained along PC1 (although in the opposite direction).

hierarchical than layers reflecting general social affinity or casual interaction, such as liking (i.e., Like layer) or sharing good news (i.e., Good News layer). This pattern suggests that not all forms of emotional exchange are structurally equivalent, even when categorized under support-related interactions. We also observe that the Drinking Most layers consistently appear near the hierarchical end of the spectrum, similar to Leadership and Influence, while Drinking Least layers are more variably positioned near the center. This distinction may indicate that heavy drinking behaviors are more commonly observed among students occupying structurally prominent or visible positions, which could potentially contribute to the reinforcement of social norms around alcohol use. In contrast, non-drinking behaviors may be less tied to structural positions and therefore less prominent in the social landscape.

Link prediction for drinking nominations using structurally similar social layers

Having identified structurally and functionally coherent clusters of social layers, we next evaluated whether these groupings can enhance link prediction, particularly for alcohol-related relationships with missing data. Specifically, we tested whether incorporating information from other layers, especially those within the same cluster, improves the accuracy of predicting missing links (see Methods).

Figure 6a shows the average area under the ROC curve (AUC) for link prediction across all target layers (vertical axis), using complete information from a second layer (horizontal axis) within the same student group as input. Diagonal blocks in the figure contain two different AUC scores: the score obtained using only the remaining 80% of the target layer (upper half), and the score obtained using all 10 layers of information (lower half, continuing only to use 80% of the target layer). We observed that AUC scores were comparably high when information from a layer within the same functional cluster as the target layer was provided, approaching the performance achieved using all layers combined. For instance, the Emotional Support layers (Offer Support, Bad News, Good News) yielded high scores when another layer from the same Emotional Support cluster was used as input.

Notably, despite their structural similarity, the Friendship and Emotional Support clusters exhibit an asymmetric influence on link prediction, with one direction providing more predictive value than the other. For instance, while layers in the Affiliation cluster (including Friendship) enhance the prediction of Emotional Support ties, the reverse direction yields only marginal improvement. These results suggest that knowing Friendship information helps narrow down which node pairs might and might not be in an Emotional Support layer but not vice versa. This asymmetry indicates that Friendship networks may contain broader structural cues that encompass but go beyond the more intense support-related connections.

In addition to the asymmetric improvements in prediction performance observed across layers, our analysis highlights that structurally and functionally coherent clusters can effectively substitute for the full multilayer network in link prediction tasks. As shown in Fig. 6a, incorporating a single additional layer from the same

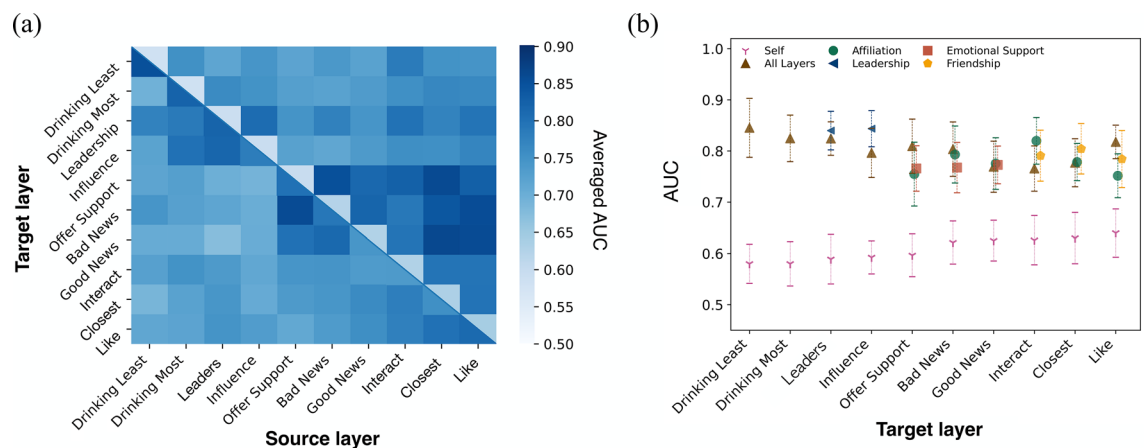


Fig. 6. AUC scores of link prediction with the functionally clustered social networks. (a) Average of 5-fold cross-validated AUC scores of link prediction using two layers (10 runs). The vertical axis labels indicate the target layer being predicted (with 20% of edges held out at a time), while the horizontal axis labels indicate the source of additional information for the link prediction. The diagonal blocks contain two different AUC scores: the upper half value is obtained using only the remaining 80% of the target layer; the lower half value indicates the AUC when this 80% of the target layer is supplemented with full information from all nine other layers. (b) AUC scores using functionally clustered layers to predict missing links in a target layer (horizontal). In the legend, each item describes the source of additional information used for link prediction: 'Self' refers to using only the remaining 80% of the target layer, corresponding to the upper half of the diagonal in (a). 'All' (brown triangle) indicates the use of the remaining 80% of the target layer plus full information from all other nine layers, corresponding to the lower half of the diagonal in (a). 'Affiliation' (green circle), 'Leadership' (blue triangle), 'Emotional Support' (maroon square), and 'Friendship' (yellow-orange pentagon) indicate the use of additional information from the corresponding layers within each functionally similar cluster. Error bars represent the standard deviation across the nine student groups.

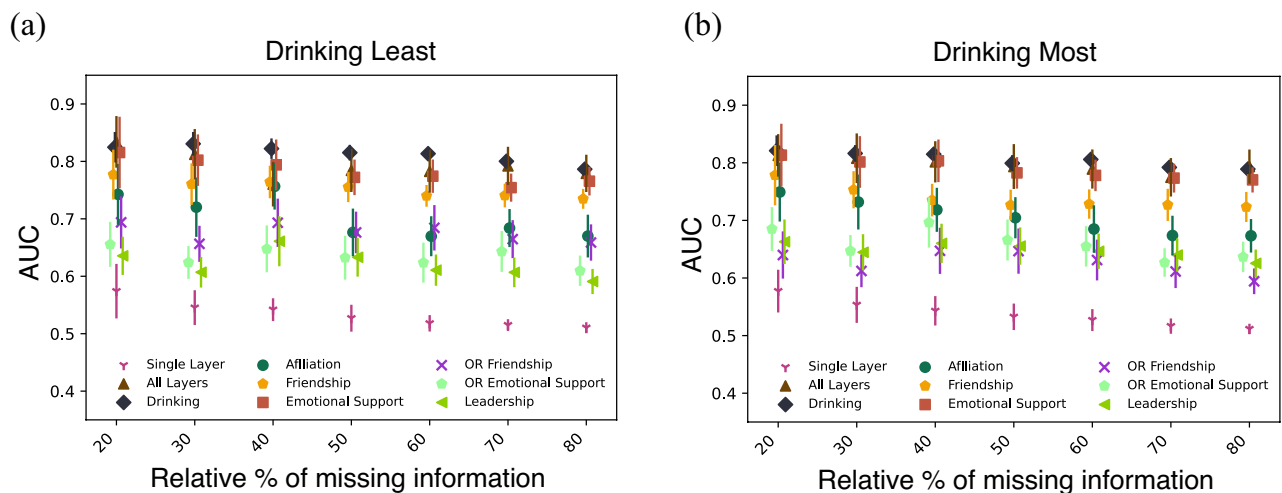


Fig. 7. AUC of link prediction on alcohol-related networks using other social layers. AUC of link prediction after removing information from the drinking least (a) and drinking most (b) layers, with varying levels of missing information (from 20% to 80%; x-axis). Each prediction is performed by first randomly removing a specified proportion $p\%$ (shown on the x-axis) of links from the target layer (i.e., drinking least in panel (a), drinking most in (b)), and then using the remaining $(100 - p)\%$ of links in the target layer together with all links from the designated source layers indicated in the legend. The source layers refer either to individual layers within a functional cluster (e.g., the Friendship cluster) or to composite layers constructed by merging the relevant layers using a logical OR operation, depending on the analysis condition (e.g., “OR Friendship” as shown in the legend). Specifically, OR denotes the logical OR operation over links in the listed layers. That is, a link exists in the “OR Friendship” layer if it appears in any of the friendship-related layers. ‘Single Layer’ uses only the remaining portion of the target layer for prediction. ‘All Layers’ adds full information from all other layers. Error bars represent the standard deviation across the nine student groups.

functional cluster provides predictive performance that, although somewhat lower than when all layers are used, remains relatively close—demonstrating the utility of structurally coherent layers even in reduced settings.

However, combining multiple layers within the same cluster substantially enhances performance, achieving results comparable to those using the entire multilayer network (Fig. 6b). Specifically, average AUC scores achieved by using cluster-based layer sets fall within the range of 0.77 to 0.85, nearly matching the full-information condition. These findings suggest that most of the predictive gain can be captured by leveraging structurally and functionally coherent subsets of layers, offering a practical and efficient alternative to using all layers—particularly in contexts with missing or incomplete network data.

We next examine whether structurally coherent clusters can improve link prediction performance in drinking-related layers. As shown in Fig. 7, incorporating the Emotional Support layers substantially improves prediction performance for both Drinking Most and Drinking Least networks, with accuracy approaching that of using all available layers. However, despite their structural proximity, the Leadership layers yield only marginal improvements in prediction. It is worth noting that this effect remains robust even when a large proportion of links in the target layer is held out.

This contrast raises the question of why Emotional Support layers contribute more to prediction accuracy than Leadership layers, despite their structural similarity. A plausible explanation is that these layers are not only structurally more aligned with the drinking-related layers—closer even than Affiliation as a whole—but also exhibit a more even distribution of nominations across individuals. The interplay between this structural proximity and the broad coverage of ties may together enhance their effectiveness in predicting missing links. In contrast, Leadership-related nominations tend to concentrate on a small number of peers, which may limit their generalizability and predictive utility. These results suggest that structural and functional alignment across specific types of social relationship layers can serve as an efficient basis for recovering missing information in alcohol-related networks, offering a data-efficient approach to understanding behavioral dynamics within peer groups.

Discussion

In this study, we examined multilayer social networks formed by breaking down peer relationships into ten distinct types, based on peer nomination survey data. Our goal was to examine the structural alignment among distinct types of peer relationships and their utility in recovering missing links—particularly in perceived drinking-related networks. Through node-, link-, and triad-level analyses, we identified clusters of structurally and functionally similar social layers, which were further differentiated along a spectrum from hierarchical to horizontal relationship types. These observed patterns offer conceptual and methodological insights into the ways individuals differentiate among social relationship types—such as affiliative, leadership, and drinking-related ties—by revealing underlying structural similarities in their network patterns.

In particular, the results demonstrate that structural information from one type of social relationship cluster—such as Emotional Support or Friendship—can be used to infer the structure of other layers, including those related to perceived drinking behavior. This finding suggests that different types of social relationships are not structurally independent, but instead organized in patterned and functionally meaningful ways, with implications for theories of social role differentiation and influence. Building on these insights, our results show that providing additional information from layers within the same functional cluster consistently improves link prediction performance. This improvement offers a practical pathway for recovering missing ties in peer networks by leveraging structurally coherent and socially meaningful layers.

While the preceding results highlight the benefits of leveraging structurally coherent layers within the same cluster, our analysis also suggests that structurally adjacent yet distinct clusters may exhibit asymmetric predictive relationships. A salient example of this is the relationship between the Friendship and Emotional Support clusters. While these clusters are functionally aligned and often grouped under the broader category of “peer networks,” our analysis reveals subtle but consistent structural distinctions that shape their predictive roles. Specifically, Friendship layers are characterized by high reciprocity and broad connectivity across all groups, whereas Emotional Support layers are more selective and hierarchically structured. These differences help explain the asymmetric link prediction results: Friendship layers are more effective in predicting Emotional Support ties than the reverse. This asymmetry highlights how nuanced variations in social structure—even among closely related relationship types—can have significant implications for network inference and intervention design.

Theoretically, prior work on social structure and exchange suggests that providing support or sharing news can create implicit role asymmetries, especially when such exchanges involve sensitive or trust-dependent information^{33,34}. These theoretical perspectives offer a possible explanation for the structural differences we observed between Friendship and Emotional Support layers. Future research could investigate how such structural asymmetries influence practical outcomes like link prediction, potentially guiding more effective intervention strategies in health-related behaviors.

We also confirmed an improvement in link prediction accuracy on perceived alcohol-behavior nominations when using data from layers in the Emotional Support cluster. While these layers are not the most structurally similar to drinking-related layers in terms of triadic or link-level similarity, *k*-means and projection-based analysis indicate that they are among the closest in low-dimensional space—second only to Leadership layers. This proximity, combined with the distributed structure of Emotional Support nominations, may help explain their effectiveness in predicting links. By contrast, although Leadership layers exhibit strong structural similarity to Drinking layers, their nomination patterns tend to be highly concentrated around a few individuals. This may limit their predictive value, as such concentrated structures might not generalize across the network. Further investigation is needed to clarify how the combination of structural similarity and the distribution of nominations influences link prediction performance, particularly for the Drinking and Leadership layers.

A better understanding of how these features contribute to inferring missing links in drinking-related networks may inform the design of more efficient intervention strategies that draw on selected social layers, such as Emotional Support or Leadership, rather than relying on broadly defined peer networks. In addition to informing interventions, our results may align with theoretical perspectives on social influence, such as Primary Socialization Theory (PST)¹⁰, which highlights the formative role of peer networks in shaping deviant or health-related behaviors. Building on this perspective, our multilayer framework offers a complementary structural lens for further examining how such socialization processes unfold within differentiated peer networks.

We also note some limitations of our study. To characterize the social lives within student groups, we asked participants to nominate other members in response to ten prompts. These prompts were selected to cover the most important dimensions of people's social lives as identified in previous research^{27,35–37}. Despite our efforts to include diverse aspects of students' social lives, the set of nomination items may not fully encompass the entire spectrum of meaningful relationships in the college context. Future work might address this limitation by asking participants to nominate other peers in response to a more comprehensive list of social network prompts and then applying dimension reduction methods to this large dataset^{38,39}.

We also note that the Drinking Most and Drinking Least links in our data are only nominations made by group members about others in the group. In particular, these perceptions can vary from the actual drinking behaviors of those nominated individuals. At the same time, a given nomination of an individual as someone who is typically Drinking Most or Drinking Least in the social group is undoubtedly the result of a complex interplay of different social elements, including the relative visibility of the person in the group as well as potential public/group demonstrations related to alcohol use or attitudes by the nominated target of the response.

Even considering these limitations, understanding the structural similarities and predictive effectiveness of social ties can inform the development of socially focused health interventions. In particular, these insights hold potential for predicting and preventing crisis-related behaviors, such as alcohol use, where direct data collection is often challenging. Specifically, based on a study showing that individuals distribute their emotion-regulation needs across multiple, discrete emotion-specific relationships⁴⁰, one could test the intervention effect of diverse peer models to promote healthy behaviors and buffer negative peer influence effects on unhealthy behaviors, such as drinking⁴¹.

Future work could also explore similar situations where different social relationship layers interact with missing information about group or individual behaviors, as seen with the two drinking layers considered in this study. To do these tasks accurately in various settings, extensive exploration of performance across group sizes and different levels of missing information is undoubtedly needed. Such investigations may be particularly useful for designing efficient sampling procedures that accurately capture such comparative behavior nominations even with lower total response rates from the group. Extending these investigations may provide the foundation for developing strategies for uncovering discrepancies within behavioral nominations so that such data can be better used to target positive interventions within social groups.

Data availability

The multilayer network representation of the data studied here can be obtained from the authors, in compliance with the relevant IRB, by request to Eun Lee <mailto:eunlee@pknu.ac.kr> and Peter Mucha <mailto:peter.j.mucha@dartmouth.edu>.

Received: 6 June 2024; Accepted: 10 October 2025

Published online: 24 November 2025

References

- Wechsler, H., Davenport, A., Dowdall, G. W., Moeykens, B. A. & Castillo, S. Health and behavioral consequences of binge drinking in college. *JAMA* **272**, 1672–1677. <https://doi.org/10.1001/jama.1994.03520210056032> (1994).
- Mundt, M. P., Zakletskaia, L. I. & Fleming, M. F. Alcohol-induced injury risk thresholds: A cross-sectional analysis of college students in 10 universities. *Alcohol. Clin. Exp. Res.* **33**, 2011–2015. <https://doi.org/10.1111/j.1530-0277.2009.00977.x> (2009).
- Meda, S. et al. Longitudinal effects of alcohol consumption on academic performance during college: A prospective investigation. *PLOS ONE* **12**, e0172213. <https://doi.org/10.1371/journal.pone.0172213> (2017).
- Baer, J. S., Stacy, A. & Larimer, M. Biases in the perception of drinking norms among college students. *J. Stud. Alcohol* **52**, 580–586. <https://doi.org/10.15288/jsa.1991.52.580> (1991).
- Perkins, H. W., Meilman, P. W., Leichter, J. S., Cashin, J. R. & Presley, C. A. Misperceptions of the norms for the frequency of alcohol and other drug use on college campuses. *J. Am. Coll. Health* **47**, 253–258. <https://doi.org/10.1080/07448489909595656> (1999) (PMID: 10368559).
- Perkins, H., Haines, M. & Rice, R. Misperceiving the college drinking norm and related problems: A nationwide study of exposure to prevention information, perceived norms and student alcohol misuse. *J. Stud. Alcohol* **66**, 470–8. <https://doi.org/10.15288/jsa.2005.66.470> (2005).
- Rosenquist, J. N., Murabito, J., Fowler, J. H. & Christakis, N. A. The spread of alcohol consumption behavior in a large social network. *Ann. Intern. Med.* **152**, 426–433. <https://doi.org/10.7326/0003-4819-152-7-201004060-00007> (2010).
- Huang, G. C. et al. Peer influences: The impact of online and offline friendship networks on adolescent smoking and alcohol use. *J. Adolesc. Health* **54**, 508–514 (2014).
- Strowger, M. et al. The association between social network members sharing alcohol-related social media content and alcohol outcomes among college student drinkers. *Alcohol. Clin. Exp. Res.* **46**, 1857–1866. <https://doi.org/10.1111/acer.14922> (2022).
- Oetting, E. R. & Donnermeyer, J. F. Primary socialization theory: The etiology of drug use and deviance. *Subst. Use Misuse* **33**, 995–1026. <https://doi.org/10.3109/10826089809062218> (1998).
- Mollgaard, A. et al. Measure of node similarity in multilayer networks. *PloS one* **11**, e0157436–e0157436 (2016).
- Brodka, P., Chmiel, A., Magnani, M. & Ragozini, G. Quantifying layer similarity in multiplex networks: A systematic study. *R. Soc. Open Sci.* **5**, 171747–171747 (2018).
- Szell, M., Lambiotte, R. & Thurner, S. Multirelational organization of large-scale social networks in an online world. *Proc. Natl. Acad. Sci. PNAS* **107**, 13636–13641 (2010).
- Morelli, S. A., Ong, D. C., Makati, R., Jackson, M. O. & Zaki, J. Empathy and well-being correlate with centrality in different social networks. *Proc. Natl. Acad. Sci. PNAS* **114**, 9843–9847 (2017).
- Kao, T.-C. & Porter, M. A. Layer communities in multiplex networks. *J. Stat. Phys.* **173**, 1286–1302. <https://doi.org/10.1007/s10955-017-1858-z> (2018).
- Schieber, T. A. et al. Quantification of network structural dissimilarities. *Nat. Commun.* **8**, 13928–13928 (2017).
- De Domenico, M., Nicosia, V., Arenas, A. & Latora, V. Structural reducibility of multilayer networks. *Nat. Commun.* **6**, 6864–6864 (2015).
- Faust, K. A puzzle concerning triads in social networks: Graph constraints and the triad census. *Soc. Netw.* **32**, 221–233. <https://doi.org/10.1016/j.socnet.2010.03.004> (2010).
- Marsden, P. V. & Campbell, K. E. Measuring tie strength. *Soc. Forces* **63**, 482–501 (1984).
- Krackhardt, D. Assessing the political landscape: Structure, cognition, and power in organizations. *Admin. Sci. Q.* **35**, 342–369 (1990).
- Borsari, B. & Carey, K. B. Peer influences on college drinking: A review of the research. *J. Subst. Abuse* **13**, 391–424 (2001).
- Moody, J. & White, D. R. Structural cohesion and embeddedness: A hierarchical concept of social groups. *Am. Sociol. Rev.* **68**, 103–127 (2003).
- Cosme, D. et al. Study protocol: Social health impact of network effects (SHINE) study. *PsyArXiv* <https://doi.org/10.31234/osf.io/cj2nx> (2022).
- Sparrowe, R. T., Liden, R. C., Wayne, S. J. & Kraimer, M. L. Social networks and the performance of individuals and groups. *Acad. Manag. J.* **44**, 316–325 (2001).
- Wellman, B. & Wortley, S. Different strokes from different folks: Community ties and social support. *Am. J. Sociol.* **96**, 558–588 (1990).
- Dumas, T. M., Wells, S., Flynn, A. & Graham, K. The influence of status on group drinking by young adults: A survey of natural drinking groups on their way to and from bars. *Subst. Use Misuse* **49**, 269–279. <https://doi.org/10.3109/10826084.2013.841243> (2014).
- Morelli, S. A., Ong, D. C., Makati, R., Jackson, M. O. & Zaki, J. Empathy and well-being correlate with centrality in different social networks. *Proc. Natl. Acad. Sci.* **114**, 9843–9847. <https://doi.org/10.1073/pnas.1702155114> (2017).
- Milo, R. et al. Network motifs: Simple building blocks of complex networks. *Science* **298**, 824–827. <https://doi.org/10.1126/science.298.5594.824> (2002).
- Holland, P. W. & Leinhardt, S. Local structure in social networks. *Sociol. Methodol.* **7**, 1–45. <https://doi.org/10.2307/270703> (1976).
- Milo, R. et al. Superfamilies of evolved and designed networks. *Science* **303**, 1538–1542. <https://doi.org/10.1126/science.1089167> (2004).
- De Bacco, C., Power, E. A., Larremore, D. B. & Moore, C. Community detection, link prediction, and layer interdependence in multilayer networks. *Phys. Rev. E* **95**, 042317. <https://doi.org/10.1103/PhysRevE.95.042317> (2017).
- Davis, J. A. Clustering and hierarchy in interpersonal relations: Testing two graph theoretical models on 742 sociomatrixes. *Am. Sociol. Rev.* **35**, 843–851 (1970).
- Coleman, J. S. Social capital in the creation of human capital. *Am. J. Sociol.* **94**, S95–S120 (1988).
- Blau, P. M. Exchange and Power in Social Life (1964).
- Zerubavel, N., Hoffman, M. A., Reich, A., Ochsner, K. N. & Bearman, P. Neural precursors of future liking and affective reciprocity. *Proc. Natl. Acad. Sci.* **115**, 4375–4380. <https://doi.org/10.1073/pnas.1802176115> (2018).
- Bayer, J. B., Hauser, D. J., Shah, K. M., O'Donnell, M. B. & Falk, E. B. Social exclusion shifts personal network scope. *Front. Psychol.* **10**, 1619. <https://doi.org/10.3389/fpsyg.2019.01619> (2019).
- Rucker, D. D. & Galinsky, A. D. Social power and social class: Conceptualization, consequences, and current challenges. *Curr. Opin. Psychol.* **18**, 26–30. <https://doi.org/10.1016/j.copsyc.2017.07.028> (2017) (Inequality and social class).

38. Jolly, E., Tamir, D. I., Burum, B. & Mitchell, J. P. Wanting without enjoying: The social value of sharing experiences. *PLOS ONE* **14**, 1–23. <https://doi.org/10.1371/journal.pone.0215318> (2019).
39. Peer, M., Hayman, M., Tamir, B. & Arzy, S. Brain coding of social network structure. *J. Neurosci.* **41**, 4897–4909. <https://doi.org/10.1523/JNEUROSCI.2641-20.2021> (2021).
40. Cheung, E. O., Gardner, W. L. & Anderson, J. F. Emotionships: Examining people's emotion-regulation relationships and their consequences for well-being. *Soc. Psychol. Pers. Sci.* **6**, 407–414. <https://doi.org/10.1177/1948550614564223> (2015).
41. Jovanova, M. et al. Brain responses and susceptibility to peer influence on drinking. <https://doi.org/10.31234/osf.io/kahyg> (2024).
42. Zurn, P., Bassett, D. S. & Rust, N. C. The citation diversity statement: A practice of transparency, a way of life. *Trends Cognit. Sci.* **24**, 669–672. <https://doi.org/10.1016/j.tics.2020.06.009> (2020).
43. Zhou, D. et al. Gender diversity statement and code notebook (dalejn/cleanbib: v1.1.1). <https://doi.org/10.5281/zenodo.4104748> (2022).

Acknowledgements

This work was supported by the Army Research Office (MURI award W911NF-18-1-0244), Global-Learning & Academic research institution for Master's-PhD students, and Postdocs(LAMP) Program of the National Research Foundation of Korea(NRF) grant funded by the Ministry of Education(No.RS-2023-00301702), and the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No.RS-2025-21073011). Additional support was provided by NSF award 2137511. The content is solely the responsibility of the authors and does not necessarily represent the official views of any agency supporting this research.

Author contributions

E.L., O.S., X.H. and P.J.M. designed the analysis and selected the calculations to be performed. E.L., O.S. and X.H. conducted the numerical calculations and developed the figures. E.L., Y.K., M.J., A.L.M., D.M.L-S. and Z.M.B. provided ideas for additional calculations. E.L., Y.K., D.S.B., K.N.O., E.B.F. and P.J.M. designed the study that developed the data studied here. All authors discussed the results at multiple points in time and contributed to writing the paper.

Declarations

Competing interests

The authors declare no competing interests.

Citation diversity statement

Recent work has identified a continuing bias in scientific citation practices such that papers from women and other minority scholars are under-cited relative to the number of their papers (see, e.g., the discussion in⁴²). We proactively aim to consider choosing references that reflect the diversity of the field in thought, form of contribution, and different aspects of identity. As one effort in this direction, we used a publicly available code⁴³ to estimate the gender identities of the first and last author of each reference by using databases that store the probability of a first name being carried by a woman. By this measure, our references (excluding self-citations to any of the authors of the present paper) contain 13.7% woman(first)/woman(last), 9.4% man/woman, 20.7% woman/man, and 56.2% man/man. This method is limited in that (a) names, pronouns, and social media profiles used to construct the databases may not, in every case, be indicative of gender identity and (b) it cannot account for intersex, non-binary, or transgender people (see⁴³ for concise discussion of the limitations of this and similar methods). We look forward to future work to help all of us better understand and implement equitable practices in science.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-24049-w>.

Correspondence and requests for materials should be addressed to E.L. or P.J.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025