Network Openness as an Early-Career Indicator of Scholar Career Trajectory in Research Institutes

Dr. Alyson Grace Eggleston, The Pennsylvania State University

Alyson Eggleston is an Associate Professor in the Penn State Hershey College of Medicine and Director of Evaluation for the Penn State Clinical and Translational Science Institute. Her research and teaching background focuses on program assessment, STEM technical communication, industry-informed curricula, and educational outcomes veteran and active duty students.

Dr. Robert J. Rabb P.E., The Pennsylvania State University

Robert Rabb is the associate dean for education in the College of Engineering at Penn State. He previously served as a professor and the Mechanical Engineering Department Chair at The Citadel. He previously taught mechanical engineering at the United States Military Academy at West Point. He received his B.S. in Mechanical Engineering from the United Military Academy and his M.S. and PhD in Mechanical Engineering from the University of Texas at Austin. His research and teaching interests are in mechatronics, regenerative power, and multidisciplinary engineering.

Erika Swift

Erika Swift is a Professor of Practice and Center for Medical Innovation Director at The Pennsylvania State University College of Medicine. Before joining Penn State University in 2015, she worked in various industry sectors, including medical devices, dental equipment, and consumer goods. She has over 25 years of experience in regional economic development, technology commercialization, new business development, entrepreneurial coaching, and startups. She received her Bachelor's Degree in Marketing with a Concentration in International Business from Bloomsburg University, an MBA from York College of Pennsylvania, and an MHA from Penn State University.

Anne K DeChant

Anne DeChant is the Associate Director for the Center for Medical Innovation (CMI) at Penn State University Hershey College of Medicine. Ms DeChant leads education and training programs for CMI, providing specialized programming for healthcare innovation and commercialization. Ms DeChant also manages a portfolio of therapeutics and diagnostic technologies, and provides expertise and support for translational research development, strategic planning, and implementation of major projects. Ms DeChant is a dynamic leader with 20+ years of experience in higher education, and earned her MBA at Case Western Reserve University, her MS in Cell Biology at Cleveland State University, and her BS at Ohio State University.

Network Openness as an Early-Career Indicator of Scholar Career Trajectory in Research Institutes

Abstract
Identifying graduate students and early career faculty with strong potential for research productivity and funding success is a growing priority in institutional research. Bibliometric approaches, the study of trends in publication topics and co-authors, offer new ways to quantify the collaborative and topical breadth of scholars' early publishing activity.

Following researchers in team science, we hypothesize that centers for institutional research are likely to draw or select for scholars with more open, interdisciplinary, and multi-topic networks of coauthors. Using recent cohorts of Innovation Fellows' (IF) scholars who have completed an entrepreneurial and industry-focused mentoring program, we compared network graph heuristics IF scholars and their peers in the same environments. This paper describes a normalized composite measure of network openness and closedness derived from co-author graphs as an early indicator of academic career trajectory. To operationalize network openness, we introduce a composite heuristic based on established graph-theory measures: low Burt's constraint, high participation coefficients, and high entropy of collaborator communities. We demonstrate that these structural properties, which may be observable within the first five years of publishing, can be correlated with distinct developmental trajectories across academic appointment types. Existing literature shows that over time, it is natural for scholars to develop denser networks as their social capital grows, but much more difficult to expand in topic focus or collaborator networks.

We identify three canonical structures: (1) 'Cohesive Cores,' dense graphs with few clusters and high total link strength, typical of early-career researchers embedded in disciplinary labs or advisor-led teams; (2) 'Bridges,' semi-open networks marked by multiple moderately connected clusters, characteristic of faculty cultivating cross-unit collaborations or work within interdisciplinary Research Institutes; and (3) 'Exploratory Constellations,' open networks with diffuse ties and low density, often associated with interdisciplinary or translational research agendas.

Comparative analyses between institute scholars and tenure-track, professional, research-track faculty show differential value in each configuration: closed networks accelerate productivity and reputation within disciplinary domains critical for tenure, while open networks enable diversified funding, cross-institutional visibility, and adaptability crucial for non-tenure professional advancement. The proposed normalized openness index functions as a useful analytic for evaluating collaborative diversification, offering administrators and research professionals a predictive metric for mentoring and resource allocation in data-driven faculty evaluation systems.

Keywords: Bibliometrics**,** Coauthor Network Analysis**,** Early-Career Researchers**,** Interdisciplinarity**,** Research Evaluation

1. Introduction: Early-Career Evaluation in Engineering and Applied STEM Contexts

Colleges and universities are increasingly seeking more sophisticated pipeline analyses to identify where STEM talent is recruited, how scholars progress through training and development pathways, and ultimately, how these pathways influence retention, degree completion, and career outcomes. This study analyzes three cohorts of Innovation Fellows from the Penn State Center for Medical Innovation, which recruits graduate and post-doctoral scholars for an entrepreneurial training and industry engagement program for graduate students and postdoctoral scholars modeled on the NSF I-Corps framework [1-4].

Traditionally, bibliometric indices have been used to quantify faculty promotion and tenure portfolios and attempt to quantify the impact a scholar's work may have on the discipline, as well as denote prestige where publications occur in highly selective venues. Examples of these metrics include the *h*-index and Journal Impact Factor. The *h*-index was never intended by its creator to predict scholarly success or future funding capture, yet it is frequently repurposed for precisely these evaluative aims [5]. *H*-index measures also fail to predict future funding capture [6-9]. More fundamental critiques of the *h*-index factor include that it emphasizes pure counts with no regard for actual author contribution or underlying innovation. Additionally, disciplines, even in STEM, differ with regard to authoring practices, author ordering, and inclusion. Lastly, *h*-indices disadvantage early career researchers, and provide no predictive value for scholar development [10]. Further, citation counts have been found to be inversely related to methodological rigor and reproducibility [11].

In contrast, emerging work in team science demonstrates that research impact is maximized when scholars combine deep expertise with interdisciplinary coauthor networks. Analyzing large-scale publication networks, Uzzi and colleagues show that teams characterized by divergent disciplinary combinations of expertise produce higher-impact outcomes, an effect closely tied to the mathematical structural features of those networks [12], [13]. Arguing that these high-entropy networks are socially produced, Barley's work [14] describes the communicative and organizational infrastructure necessary to produce these network effects, drawing on longitudinal network analyses combined with qualitative inquiry. Longitudinal analyses further demonstrate that participation in interdisciplinary research is associated with measurable reductions in network constraint and increased diversity of collaboration over time, highlighting the role of institutional context in shaping network entropy [15]; however, not all scholars are a 'best fit' for this approach, as it appears to align most strongly with individuals who precociously pursue boundary-spanning, exploratory collaboration patterns of work, which is a fairly unique trait for early career scholars, many of whom are embedded in siloed laboratory environments and context dependent authorship networks that are explicitly encouraged by advisor-led research structures.

2. Network Features

Quantifying Collaboration

Because early-career collaborations are shaped by institutional context, advisor relationships, disciplinary norms, as well as opportunity, co-author network structure is best understood as a signal of prior conditions rather than an individual's choice. Accordingly, network signals here are understood as part of a scholar's trajectory information, and not causal determinants.

Grounded in network theory, three structural characteristics are particularly relevant for characterizing boundary-spanning and translational collaboration patterns: (1) low Burt's constraint; (2) high participation across multiple communities; and (3) and high entropy of collaborator affiliations. Burt's constraint quantifies the degree to which a node (scholar) has interconnected connections. Participation across communities is a concept that quantifies how evenly a node's links are distributed across network clusters or communities. Finally, entropy of collaborate affiliations is borrowed from information theory and measures the dispersion of collaborators across affiliation categories (disciplines, institutions, academic units). This constellation of features captures the degree to which a scholar's contributions are distributed across unique intellectual and institutional spaces rather than a singular, topic concentration. While these features have been associated with large-scale studies of impact, innovation, profit return, and team science in prior work, they are not optimal for all scholars, nor are they accessible at all career stages for everyone. Nevertheless, their presence functions as an early indicator of an orientation toward exploratory and collaborative work.

To examine these structural features empirically, this study employs a set of bibliometric and network-derived measures that approximate openness, participation, and entropy within individual scholars' co-authorship networks. Table 1 summarizes the metrics used, their definitions, and their theoretical alignment.

Table 1: Network features and operational definitions

Metric	**Source**	**Definition**	**Indicator**
Total Publication Count	Web of Science	Number of indexed publications	Greater productivity (not necessarily interdisciplinarity)
Institution Count	Web of Science	Number of unique institutional affiliations among co-authors	Cross-institutional engagement
Category Count	Web of Science	Number of distinct subject categories published in	Cognitive entropy
Core Field	Web of Science	Primary disciplinary field of publication	Anchoring domain
Structural Span	Derived	Institutions × Categories	Cross-boundary reach
Interdisciplinary (Binary)	Manual Heuristic	Yes/No classification based on conceptual distance across subject categories	Presence of interdisciplinary activity
Cluster Count	VOSviewer	Number of detected co-author clusters	Community diversity
Link Count	VOSviewer	Total co-authorship ties	Network activity
Total Link Strength	VOSviewer	Weighted co-authorship intensity	Tie concentration

Networks and Temporal Effects

Network density is informed by temporal path. Co-author networks tend to become denser over time as collaborators reappear on subsequent follow-on publications and research programs begin to cohere. This network density can be a sign of within-domain productivity. At the same time, network openness is necessarily both easier to observe earlier in a career, and also harder to acquire later in career trajectories. This asymmetry suggests that early-career network openness functions less as a mutable trait and more as a trajectory signal, reflecting early exposure to cross-boundary collaboration opportunities rather than later strategic repositioning.

Networks and Appointment Types
Network openness or cohesion does not necessarily correlate with particular career structures. Broadly, tenure-track pathways often prize disciplinary depth and recognized advisor and author 'lineages.' Professional, research track, and entrepreneurial roles may value adaptability and cross-unit collaboration. Each approach has a role in the larger institutional ecosystem.

3. Methods

This study draws on administrative records from the [Redacted] innovation fellows program across three cohorts (2023–2025). In total, 20 individual records were identified: five fellows from the 2023 cohort, seven from the 2024 cohort, and eight from the 2025 cohort. Publicly available scholarly records were linked using ORCID identifiers where available. One individual from the 2024 cohort and three individuals from the 2025 cohort did not have ORCID identifiers and were excluded from bibliometric analysis. Among fellows with ORCID identifiers, two individuals from the 2025 cohort did not yet have publication records indexed in Web of Science [16], reflecting their early career stage, and were also excluded. The final analytic sample therefore consisted of 14 fellows with complete bibliometric and network records across the three cohorts.

Bibliometric Measures
For each individual with a complete record, publication data were extracted from Web of Science. Measures included: (1) total number of publications, (2) number of distinct institutional affiliations represented in their co-author network, (3) number of discrete Web of Science subject categories in which they published, and (4) primary or core field of publication. Core fields spanned biomedical engineering, chemistry, chemical engineering, medicine, and friction and vibration, reflecting the interdisciplinary scope of the CMI program.

A binary indicator of interdisciplinarity (yes/no) was also assigned using a manual heuristic based on the conceptual distance among Web of Science subject categories represented in each scholar's publication record, rather than publication count alone.

Derived Network Measures
To capture cross-boundary reach, a Structural Span Score was computed for each scholar by multiplying the number of distinct Web of Science subject categories by the number of unique institutional affiliations represented in their publication record. This composite measure was intended to index the extent of cross-institutional and cross-disciplinary engagement.

Co-authorship Network Analysis
Each scholar's publication record was exported to VOSviewer [17] to generate ego-centric co-authorship networks. From these networks, three structural measures were extracted: (1) number of detected co-author clusters, (2) number of links among collaborators, and (3) total link strength, representing the weighted publication count between co-authorship ties. In VOSviewer co-authorship networks, a link represents the presence of at least one co-authored publication between two authors, while link strength reflects the frequency of co-authored publications

between those collaborators. Together, these measures characterize the size, cohesion, and diversity of each scholar's collaboration network.

Analogs to Canonical Network Measures

Low Burt's Constraint (LBC) conceptually maps to how often a scholar connects to unique collaborators. This can be approximated for each scholar by identifying all co-authors, followed by a cluster analysis, wherein a could be a home department or field. This is computed as:

$$\text{Independence} = 1 - \sum_k \left(\frac{n_k}{N}\right)^2$$

Where n_k = number of collaborators in cluster k. This formulation also maps to a Simpson-style Diversity Index, which has been used in ecological studies and adapted to economic, linguistics, and sociology [18].

Similarly, a Participation Coefficient can be mapped to how evenly a scholar distributes collaboration across clusters, as defined in the LBC. Here, weighted ties such as multiple publications with the same individual can be represented as w_k, and as $P_{weighted}$ increases, that is interpreted as evenly distributed across clusters, exhibiting classic 'bridging' behavior, and as $P_{weighted}$ decreases, it signals concentrated collaboration in one community, exhibiting 'cohesive core' structure.

$$P_{weighted} = 1 - \sum_k \left(\frac{w_k}{\sum w_k}\right)^2$$

Finally, entropy of collaborator communities can be parsed as a measure of diversity of collaborator affiliations, here understood as distinct departments or fields, wherein a Shannon entropy measure that normalizes for K clusters could be computed, as shown below.

$$H_{norm} = \frac{\log(K)}{\log(N)}$$

4. Results

Descriptively, the Center trainees' collaboration networks show three canonical profiles: (1) cohesive cores; (2) bridges; (3) exploratory constellations. Cohesive cores feature a low cluster count, as well as comparatively higher link strength, indicating a higher degree of embeddedness and specialization. They also feature a more limited count of category coverage with a correspondingly lower structural span. This profile aligns well with advisor-led labs with strong internal cohesion but limited outward reach. Bridges feature multiple clusters with moderate-to-high link strength, indicating some degree of research program momentum. Additionally, they display higher institution and category counts than cohesive cores, with a correspondingly higher structural span relative to their publication count. In alignment with Barley and others' prior work, Bridge-profiling scholars are likely also to be institute-affiliated faculty, translational collaborators, and those likely to spearhead 'working groups' at their future institutions. Finally, exploratory constellations can be defined by higher cluster counts relative to publication counts, lower link strength, and higher category breadth. These are wide but shallow networks, and earmark a career trajectory still emerging. Exploratory constellations are also likely methods-crossers, in that these scholars may use similar tools and approaches across varied terrain. Table 2 shows relative indicator strength across profiles. These profiles reflect distinct collaboration approaches observed among early career scholars and are not intended as performance rankings. Table 3 summarizes the canonical network measures as applied to the network structural profiles, described above.

Table 2: Relative network and bibliometric features across canonical profiles

Feature	Cohesive Cores	Bridges	Exploratory Constellations
Number of Clusters	–	+	++
Number of Links	–	++	+
Total Link Strength	++	+	–
Publication Count	+	+	–
Institution Count	–	+	+
Category Count	–	+	++
Structural Span (Inst × Cat)	–	+	+
Collaboration Breadth	–	+	++
Collaboration Depth	++	+	–
Career Trajectory Signal	Specialization	Translation	Emergence

Table 3: Network measures across canonical profiles

Feature	Cohesive Cores	Bridges	Exploratory Constellations
LBC / Independence	–	++	++
Participation Coefficient	–	++	++
Entropy	–	+	++

Graphically, Figures 1-3, display scholar exemplars from each profile type, as generated with VOSViewer.

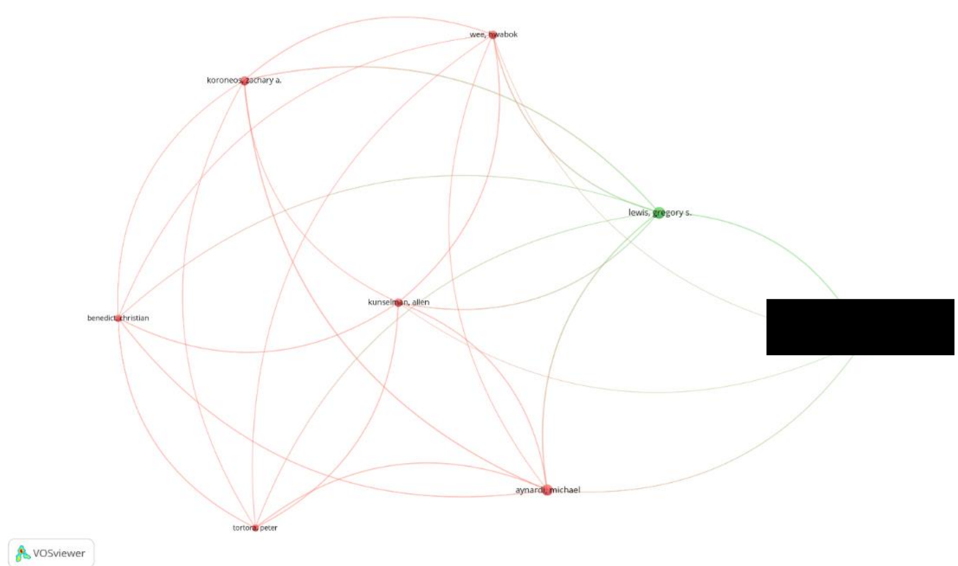

Figure 1: Cohesive core network profile exemplar

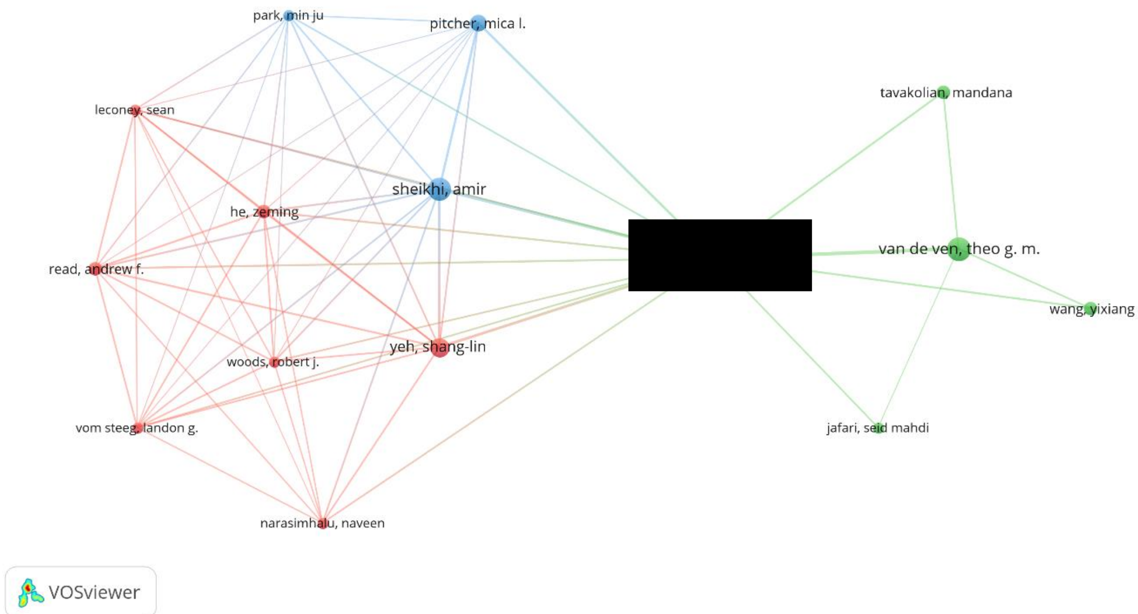

Figure 2: Bridge network profile exemplar

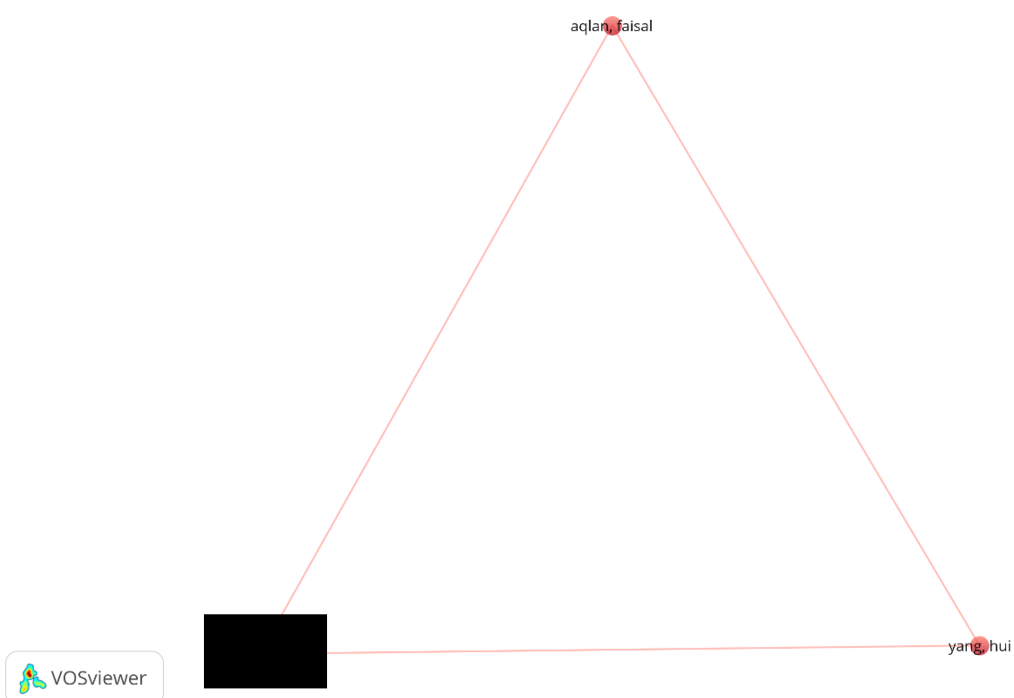

Figure 3: Exploratory network profile exemplar

Applying these heuristics, cohesive cores were defined in the dataset as having 1-2 clusters; bridges were identified by having ≥ 3 clusters, high link strength, and structural span driven by both institution and category counts; and while exploratory profiles had 1-3 clusters and low publication counts, low link strength, and structural spans driven mostly by category counts. As shown in Table 3, Bridge profiles were the most common collaboration pattern among trainees, representing 43% of the sample, with Cohesive Core and Exploratory profiles representing 28.6%, respectively.

Table 3: Distribution of Trainees Across Profiles

Network Profile	**Number of Scholars**		**Percentage of Sample**
Cohesive Cores	4	2 Eng / 2 Med	28.6%
Bridges	6	2 Eng / 4 Med	42.9%
Exploratory Constellations	4	1 Eng / 3 Med	28.6%
Total	14	5 Eng / 9 Med	100%

5. Discussion

Rethinking Early-Career Evaluation

The network profiles identified here underscore the importance of interpreting collaboration metrics as trajectory signals rather than performance scores. Network openness and entropy reflect early positioning and opportunity structures, not scholarly quality. Dense, advisor-led collaboration patterns are developmentally normal for many early-career scholars and should not be interpreted as negatively. Instead, these metrics can be used provide just-in-time supports and connections to scholars who would most benefit, depending on their career stage, role, and the mission of the institution.

Notably, bibliometric collaboration networks may not capture external relationships required for research translation and commercialization. Trainees who are well-connected academically may not necessarily be the same individuals who develop relationships with industry partners, investors, or economic development organizations. These entrepreneurial networks represent a distinct form of social capital that can influence translational outcomes. Together, these distinctions reinforce the value of using network signals to tailor mentorship, training, and external connectivity that align with both academic advancement and translational goals.

Implications for Mentoring and Program Design

The presence of distinct collaboration profiles suggests that uniform mentoring approaches are unlikely to be effective. Scholars embedded in cohesive cores may benefit from selective, low-risk exposure to cross-boundary collaboration, while those exhibiting bridge or exploratory profiles may require support in consolidating breadth into sustainable research programs. Because collaboration networks are path dependent, early interventions are likely to be more effective than later efforts, particularly when framed as optional developmental opportunities rather than as requirements.

Institutional Use Cases and Ethical Considerations

Network analytics offer promise for informing mentoring and program design, but their institutional use requires clear safeguards. These measures should be applied transparently, interpreted in context, and framed developmentally rather than evaluatively. High-stakes use in hiring or promotion without complementary qualitative assessment risks reinforcing existing accessibility differences and obscuring meaningful variation in scholarly trajectories. Of note, proprietary and executive-level platforms like Academic Analytics or machine-learning powered platforms like Dimensions, have the potential afford other kinds of networked overlays for a scholar, to include institutional recognition and award efforts, internal pilot funding support, and working group memberships. Network analytics beyond publication activity can provide a fuller footprint of a scholar's impact signal.

6. Conclusion

This study demonstrates the utility of network-based descriptors for characterizing early-career collaboration patterns within an entrepreneurial training context. By identifying cohesive cores, bridges, and exploratory constellations as non-hierarchical profiles, the analysis shifts attention from static 'counts' toward questions of trajectory and development. Rather than seeking a single optimal network structure, institutions should align collaboration patterns with stated missions and program values. As institutions increasingly turn to analytics to understand and support STEM and engineering talent pathways, this work contributes a theory-informed approach that prioritizes developmental opportunities.

Future research with larger samples and longitudinal designs will examine how these profiles evolve over time and how institutional interventions can support different collaboration trajectories. Larger samples will be needed to identify right-fit network measure thresholds, supporting continued tailoring of individualized entrepreneurial learning experiences for Innovation Fellows with the Penn State Center for Medical Innovation.

References

[1] C. A. Canaria, L. Portilla, and M. Weingarten, "I-Corps at NIH: Entrepreneurial training program creating successful small businesses," *Clinical and Translational Science*, vol. 12, no. 4, p. 324, 2019.

[2] M. S. Ahmed, "Lessons learned from NSF I-Corps boot camp," *Journal of Education and Practice*, vol. 8, no. 26, pp. 1–10, 2017.

[3] K. A. Smith, A. F. McKenna, R. C. C. Guerra, R. Korte, and C. Swan, "Innovation Corps for Learning (I-Corps™ L): Assessing the potential for sustainable scalability of educational innovations," in *Proc. ASEE Annual Conf. & Exposition*, New Orleans, LA, June 26–29, 2016.

[4] J. L. Mondisa, B. W. Packard, and B. L. Montgomery, "Understanding what STEM mentoring ecosystems need to thrive: A STEM-ME framework," *Mentoring & Tutoring: Partnership in Learning*, vol. 29, no. 1, pp. 110–135, 2021.

[5] N. Alter, M. Daiem, M. E. Pontell, I. Galdyn, M. Golinko, G. Perdikis, and W. Lineaweaver, "Limitations of academic bibliometric indices: The need for more

comprehensive metrics," *Annals of Plastic Surgery*, vol. 95, no. 6, pp. 603–606, 2025. doi:10.1097/SAP.0000000000004484.

[6] A. Eggleston and C. Kantor, "Pace and pitch: Predictive factors for seed funding and development," *Journal of Clinical and Translational Science*, vol. 8, suppl. 1, pp. 53–54, 2024. doi:10.1017/cts.2024.169.

[7] A. Eggleston and J. Petrie, "Predicting success: A mixed model of KL2 trainee profiles and outcomes," *Journal of Clinical and Translational Science*, vol. 8, suppl. 1, pp. 49–50, 2024. doi:10.1017/cts.2024.158.

[8] V. Koltun and D. Hafner, "The h-index is no longer an effective correlate of scientific reputation," *PLOS ONE*, vol. 16, no. 6, e0253397, 2021.

[9] M. K. Akhtar, "The h-index is an unreliable research metric for evaluating the publication impact of experimental scientists," *Frontiers in Research Metrics and Analytics*, vol. 9, 1385080, 2024.

[10] M. R. Dougherty and Z. Horne, "Citation counts and journal impact factors do not capture some indicators of research quality in the behavioural and brain sciences," *Royal Society Open Science*, vol. 9, no. 8, 220334, 2022. doi:10.1098/rsos.220334.

[11] C. W. Belter, "Bibliometric indicators: Opportunities and limits," *Journal of the Medical Library Association*, vol. 103, no. 4, pp. 219–221, 2015. doi:10.3163/1536-5050.103.4.014.

[12] B. Uzzi, S. Mukherjee, M. Stringer, and B. Jones, "Atypical combinations and scientific impact," *Science*, vol. 342, no. 6157, pp. 468–472, 2013.

[13] B. F. Jones, "The burden of knowledge and the 'death of the renaissance man': Is innovation getting harder?," *Review of Economic Studies*, vol. 76, no. 1, pp. 283–317, 2009.

[14] W. C. Barley, "Anticipatory work: How the need to represent knowledge across boundaries shapes work practices within them," *Organization Science*, vol. 26, no. 6, pp. 1612–1628, 2015.

[15] W. C. Barley, L. Dinh, L. P. Johnson, and B. F. Allan, "Membership in team science institute enhances diversity of researchers' collaboration networks," *PLOS ONE*, vol. 20, no. 5, e0322943, 2025. doi:10.1371/journal.pone.0322943.

[16] VOSviewer, "VOSviewer," Available: https://www.vosviewer.com/

[17] Web of Science, "Web of Science," Available: https://clarivate.com/academia-government/scientific-and-academic-research/research-discovery-and-referencing/web-of-science/

[18] P. Thoiron, “Diversity index and entropy as measures of lexical richness,” *Computers and the Humanities*, vol. 20, no. 3, pp. 197–202, 1986.