Analyzing the Growth of a Statewide Network to Increase Recruitment to and Persistence in STEM

By Marjorie Darrah, Kimberly Cowley, Christopher Wheatley, Leah McJilton, and Roxann Humbert

The First2 Network is a collection of people from K–12, higher education, government, and industry who are coming together to ensure that students of West Virginia, a rural Appalachian state, will be prepared to choose science, technology, engineering, and mathematics (STEM) majors and persist in them. This project—funded by the National Science Foundation—combines many features, including semi-annual conferences, structured working groups, summer immersive experiences for students, a student ambassador program, and network improvement communities. The growth of the First2 Network is vital to make sure that these activities and programs are disseminated and sustained statewide. This article uses social network analysis to examine participation of people around the state during the first three years of the project. Findings indicate that the network is growing in number of people and in strength of connections. Network leadership members are playing key roles in the network, and student participants who persist in their STEM majors have stronger ties to the network. Social network indicators suggest that the network has manifested positive changes in the first three years of the project, which will lead to increased communication and collaboration among state agencies related to STEM persistence within the state.

Introduction

According to "VISION 2025: The West Virginia Science and Technology Strategic Plan," "science, technology, and engineering are West Virginia's leading economic growth drivers attracting investments, creating jobs, and improving our quality of life" (WVHEPC 2015, title page).

Marjorie Darrah *is Professor of Mathematics at West Virginia University.*

Kimberly Cowley *is a senior project manager for child welfare at the consulting firm ICF.*

Christopher Wheatley *is a PhD candidate in Physics at West Virginia University.*

Leah McJilton *is an instructional design and technology expert.*

Roxann Humbert *is the previous statewide director of online learning in West Virginia.*

The state as such needs a highly educated, technically skilled, and entrepreneurial workforce as a foundational element. The report goes on to say that "STEM-based businesses started in West Virginia will diversify and strengthen the State's economy, attract additional population, and grow the tax base to further the development of entrepreneurial opportunities and increase educational attainment in the state. STEM-based education, employment, and entrepreneurial opportunities are critical for diversifying and boosting the State's economy" (WVHEPC 2015, 7).

Although growing a science, technology, engineering, and mathematics (STEM) workforce is critical to West Virginia's economic future, for years it has been known that the United States and many other countries do not produce enough STEM majors to meet the current government and industry employment demand (Peri, Shih, and Sparber 2015; van den Hurk, Meelissen, and van Langen 2019). According to the US Department of Commerce (2017), STEM occupations are growing six times as much (24 percent) as other occupations, which are growing at 4 percent. In addition, STEM degree holders have a higher income even in non-STEM careers and play a key role in the sustained growth and stability of the US economy. Science, technology, engineering, and mathematics education creates critical thinkers, increases science literacy, and enables the next generation of innovators. However, students who enter the STEM pathway often leave by either changing majors or dropping out of postsecondary education altogether (Chen and Soldner 2013). Many authors have referenced this as the "leaky" STEM pipeline, and it has been the subject of many studies (van den Hurk, Meelissen, and van Langen 2019). Despite the rapid growth of enrollment in STEM disciplines in recent years, the number of students graduating with a STEM degree has remained relatively stagnant due to diminishing student retention rates (Eagan, Hurtado, and Chang 2010; Thompson and Bolin 2011).

Nationally, more than half of all college students who declare a STEM major drop out or change their major in the first two years of postsecondary education, and this problem is particularly acute for first-generation college students (Chen and Soldner 2013). While these results indicate the success of elementary and secondary education in cultivating interest in STEM fields, more work is still needed to understand the dwindling retention rates at the postsecondary level (Watkins and Mazur 2013). Recent studies have found that among students who enrolled as a major in a STEM field within their first year of postsecondary education, 37 percent had completed a degree or certification in a STEM field within six years, 7 percent maintained enrollment in a STEM field, and 55 percent had either switched to a non-STEM field or left postsecondary education (Chen and Weko 2009). Improving STEM retention is thus crucial to ensuring a stable

STEM pipeline and underrepresented young people's fair access to STEM educational opportunities.

When pursuing postsecondary education, rural students have additional barriers to overcome. West Virginia lies completely in Appalachia and has 51 percent of the population living in rural areas. Additionally, thirty-four of its fifty-five counties are considered rural, and it is ranked the third most rural state in the United States (World Population Review 2020).

Some of the barriers to starting and completing a college degree for these rural students include familial commitments and lack of resources or support needed to identify educational paths to higher-wage jobs (McCann and Keily 2021). These barriers for rural students have created opportunity and participation gaps in education programs. Rural students also experience access issues, such as limited broadband, which further contribute to the challenges that residents of nonmetropolitan areas face in accessing education and work. The WHO pandemic has exposed and exacerbated these challenges. Addressing these barriers to promote higher educational attainment and better access to education can provide economic opportunity and help to address skill shortages within labor markets, especially in rural communities (McCann and Keily 2021).

According to Showalter et al. (2017), nearly half of rural students are from low-income families, more than one in four is a child of color, and one in nine has changed residence in the previous year. Rural states are also plagued with high transportation costs to get students to schools, and West Virginia has the nation's highest transportation costs for rural schools (Showalter et al. 2017). On average, rural school districts nationally spend about $10.36 on instruction for every dollar spent on transportation, but West Virginia spends only $6.54 on instruction for every transportation dollar spent. West Virginia is also among the seven states ranked in the highest priority quartile—the quartile with the lowest rural National Assessment of Educational Progress (NAEP) scores—on all five NAEP indicators (Showalter et al. 2017).

In spite of the barriers mentioned above in West Virginia, the state's high school students indicate higher levels of interest in STEM than nationally—58 percent versus 48 percent, according to a report by ACT (2016). However, among those ACT-takers indicating interest in pursuing STEM studies in 2017, only 30 percent achieved the mathematics benchmark, with 32 percent reaching the science benchmark (ACT 2017). Even more concerning, only 11 percent achieved the STEM benchmark (a derived score combining mathematics and science scores and correlated with success in STEM courses that STEM students commonly enroll in) (ACT 2017).

The state's postsecondary students are served by thirteen public four-year institutions, nine public community and technical colleges, and eight independent four-year colleges (WVHEPC and West Virginia Community and Technical College System 2018). In terms of persistence and degree completion, the state falls below national and regional averages. The state is in last place among the sixteen Southern Regional Education Board states in overall first-year persistence, with a rate of 77 percent for 2016 compared to 85 percent for the region in 2015 (Southern Regional Education Board 2018). The state's Higher Education Policy Commission reports a 31 percent on-time graduation rate for first-time freshmen pursuing bachelor's degrees, compared to 40.6 percent nationally (WVHEPC and West Virginia Community and Technical College System 2019). In this state, low-income students, many of whom are also first-generation college students, graduated at a rate of just 22.4 percent in 2014, a slight increase from 21.5 percent in 2013 (WVHEPC and West Virginia Community and Technical College System 2018).

West Virginia's policymakers, education leaders, and advocates have taken up the call to improve STEM education and persistence across the state. The state's Department of Education, for instance, has planned a comprehensive statewide approach to improving STEM education, and advocacy organizations such as West Virginia Forward, the Education Alliance, and the West Virginia Public Education Collaborative are implementing initiatives to promote STEM.[1] In addition, young people have access to various STEM enrichment opportunities, including STEM summer camps at state institutions of higher education, the Governor's STEM Institute (GSI), and programs sponsored by the National Aeronautics and Space Administration (NASA) and Green Bank Observatory.

One of the National Science Foundation's (NSF) "10 Big Ideas" for positioning "the United States at the cutting edge of global science and engineering leadership through pioneering research and enabling activities" is the INCLUDES program that supports a variety of projects (IGEN 2022).[2] These projects include design and development launch pilots, alliances, a coordination hub, and planning grants to build collaborative infrastructures and create the foundation for a national network. West Virginia was awarded a two-year design and development launch pilot in 2016 that focused on increasing the STEM persistence of first-generation rural college students in that state (National Science Foundation 2016). Project partners included faculty and students from institutions of higher education, staff from a radio telescope observatory, staff from a nonprofit educational organization, local school system administrators, and officials from the state's Department of Education. In 2018, the state received one of only seven alliance awards funded nationwide. The focus of this five-year effort is to develop a statewide network to pull together many state resources and address the issue of STEM persistence.

The First2 Network strives to improve college enrollment rates and success of undergraduate STEM students, especially first-generation and rural students, during their first two years of college. At the heart of the network's shared vision is the belief that the students themselves should be co-creators of the solutions. Several organizations make up the core of the network, as noted above, but the network has grown continually over the past three years and, at the time of this analysis, included more than 128 individuals from higher education institutions, the state's Department of Education, STEM industries, local education agencies, and other organizations.

The network employs improvement communities to carry out plan-do-study-act improvement cycles within five working groups: capacity building, college readiness, faculty-student engagement, immersive experiences, and student leadership. Other key activities include summer research internships for incoming freshmen and semi-annual conferences, as well as internal research and external evaluation components.

Social network analysis (SNA) provides answers to key questions about how networks are forming and growing. Social networks are comprised of the relationships within and between individuals, groups, or organizations. Social network analysis is a technique for identifying and measuring those relationships and can focus on the whole network and on the ties or connections between individuals (Honeycutt 2009). Wasserman and Faust (1994) explain network analysis as focusing on both the exchange of resources and the patterns of relationships between individuals. In SNA terminology, relationships are viewed in terms of "nodes" and "ties"—nodes being the individual actors within the network, and ties being the relationships between and among individuals.

Noonan et al. (2014) suggest two fundamental relational aspects: intensity/frequency/strength of interaction between pairs of individuals and direction of interactions between pairs. Fredericks and Durland (2005) summarize the main points of consideration for examining results at the network and individual levels. Igarashi, Takai, and Yoshida (2005) expand the concept of individual-level centrality further. They note that it can include outdegree (number of individuals an actor identifies), indegree (number of individuals who identify a particular actor), betweenness (frequency with which an actor falls between pairs of other individuals on the shortest path connecting them), and information (proportion of total information flow in a network controlled by an individual).

When cross-agency networks are formed to increase collaboration and/or coalition building, network analysis can provide "unique information about the structure of connections between agencies" (Cross et al. 2009, 311). Further, Cross et al. (2009) note existing research indicates that network structure and tie strength affect outcomes such as transfer of knowledge, organizational change, and innovation. Social network studies seek to observe, characterize, and draw conclusions from patterns of connections among associates in some setting of interest. Examples range from patterns of play among school children (Guralnick, Connor, and Johnson 2009) to strategic alliance and competition among corporations on the international economic stage (Osborn and Hagedoorn 1997).

Kezar (2014) does a review of research that looks at higher education change and social networks. Numerous researchers have employed SNA in studying STEM communities of practice (Ma et al. 2019; Ma et al. 2018). Baker-Doyle and Yoon (2011) look at network analysis for understanding and facilitating teacher collaboration in STEM professional development. Shadle et al. (2018) researched transformational change in higher education STEM instruction and Knaub, Henderson, and Fisher (2018) compared SNA results with other data collection methods to identify leaders of STEM education reforms.

Social network analysis has been used in evaluation efforts that ranged from assessing capacity building to exploring policy networks (Carman and Fredericks 2018) in fields including public health, social services, and education. Further, SNA can examine the relationship between networks and outcomes as well as track network evolution over time (Honeycutt 2009). Halgin and Borgatti (2012) describe specific longitudinal measures to examine the formation of new ties, the retention of existing ties, and the loss of old ties (collectively termed "tie churn").

Often of interest in SNA research is the degree of collaboration occurring within the network. The National Network for Collaboration created a collaboration framework that identifies five levels of relationships, with the highest being collaboration (accomplishing a shared vision, building an interdependent system) and the lowest being networking (communication) (Cross et al. 2009; Hogue et al. 1995).

The First2 Network is attempting to bring together disparate resources to increase the number of students who choose STEM majors and to support these students to increase retention to graduation and on to a STEM career. Social network analysis was employed specifically to consider whether cross-institution collaborations were being formed throughout the state to share information for facilitating student STEM persistence. We also sought to determine if new members were being added to the network and if collaborations were getting stronger over time. The aim of the study was to answer the following questions: (1) How and in what ways is the network growing over time? and (2) To what extent do the interconnectedness and information flow improve among members? To do this, the research and evaluation teams used SNA. The following sections outline the methods used, the result of the analysis over three years, and the conclusions that were drawn about the growth of the First2 Network.

Methods

The research team collects yearly social network survey data from people who are participating in the First2 Network and uses social network analysis to analyze the first three years of data to try to answer the questions above. The individual years of data are analyzed, and then changes for the three-year time frame are observed.

The study focuses on several social network key concepts that will be explained in more depth later. These concepts include:

- Density: proportion of total ties connecting individuals
- Centrality: degree to which an individual is in a central role in the network
- Social conductivity: combination of all the paths between those two persons based on the application of the series/parallel resistor formula taught in introductory physics
- Robustness: measure of connection after the shortest path between individuals is removed

The last two measures, social conductivity and robustness, were developed by Buch (2019) for this project and are explained in depth in Appendix A.

Participants

In the first three years, 127 individuals either responded to our social network survey or were named by someone who responded to the survey. There were twenty-three organizations represented in the data, ten colleges and universities, five companies, two statewide educational agencies, two county school systems, one state research organization, one state technology consortium, one nonprofit, and one statewide student program. The respondents included administrators, faculty members, researchers, student advisers, graduate students, undergraduate students, and staff members. A breakdown of people in the network each year is available in table 1.

Measures

After reviewing several social network surveys and realizing most are developed specifically for the given situation, the research and evaluation teams developed a survey using standard social network survey questions that were applicable to the First2 Network. One particular example

Table 1: Categories of people represented by nodes in the graphs (with and without isolated nodes)

	2018	2018 (No Isolated Nodes)	2019	2019 (No Isolated Nodes)	2020	2020 (No Isolated Nodes)
Students	6	6	19	13	41	28
College/university faculty	17	17	30	27	37	27
Statewide educational organization	7	7	8	6	11	5
Company staff	2	2	3	3	5	4
K–12 faculty	1	1	2	2	5	4
Academic advisors	1	1	1	1	4	3
Evaluators	1	1	3	3	3	2
Nonprofit	2	2	3	3	3	2
Research organization	1	1	3	3	3	2
Statewide student program	5	5	6	3	6	2
University staff	0	0	0	0	2	2
Statewide technology consortium	1	1	1	1	1	1
College/university administrators	4	4	4	0	4	0
Other	0	0	2	2	2	0

developed by Durant-Law and Milne provided a good basis for the network survey.[3]

After the survey was developed, face validity for the survey was performed using a group from the First2 Network's leadership team and research team to answer the questions and give feedback on whether the survey measured what it was intended to measure. Feedback from this group allowed for revision of the survey before distributing to the full group.

The survey was delivered in an online format and began with an informed consent page that was approved by a university institutional review board. Respondents could give consent or opt out of taking the survey. Survey questions included demographics about the respondent (name, organization, role in organization) and then asked individuals to name up to ten other people in the network with whom they communicate. To classify the strength of the connections between individuals, respondents were asked to use a system of ranking including: 1 (networking), 2 (cooperation), 3 (coordination), 4 (coalition), and 5 (collaboration), developed by Hogue et al. (1995) and Borden and Perkins (1999). These terms were explained in the body of the survey indicating the connection between individuals' increases from ranking 1 to 5.

Procedures

Data were collected from the statewide group for three years using a survey developed and distributed through the Qualtrics survey tool. The data set was downloaded from Qualtrics, then preprocessed and analyzed using R, a free software for statistical computing and graphics. The research team developed code and utilized existing SNA packages in R to compute the metrics and depict the graphs.

To begin the analysis of the First2 Network connections, an adjacency matrix, A, was created, which included the nodes (people) and ties (connections). The adjacency matrix is a square $n \times n$ matrix where n is the number of nodes (people) in the network. Each entry in the matrix, a_{ij}, represents the weight on the tie (connection) from node (person) i to node j. If the persons are not connected, then the entry $a_{ij} = 0$. If person i names person j and gives a number associated with their connection and person j names person i and gives a number associated with their connection, then the two numbers are averaged. If only one person mentions the other person, then 0 is averaged with the one number. The matrix is always symmetric, in other words $a_{ij} = a_{ji}$, for all $1 \leq i \leq n$ and $1 \leq j \leq n$. The graph formed is weighted and undirected. Each year, the same people are in the same positions in the matrix, so that not only can we track the network changes, but we can also compare from year to year how a person's connections change within the network. It should be noted that if people did not fill out the survey each year, we still included their names, since they were part of the network at one time. These individuals might either show up as disconnected nodes or their connectivity may be reduced, but we felt this captured growth and change.

Some of the standard measures were used to quantify the First2 Network. The number of nodes (people) and ties (connections) are reported. The *graph density*, the ratio of ties present in the graph to all possible ties, is given. *Centrality* is used to indicate how central a person is inside the social network by examining the connections attached to that person, as well as the geodesic distances (shortest path lengths) to other persons in the network. The *degree centrality* (*DC*) quantifies how many connections a person has to other people. Mathematically, in our undirected weighted graph, the degree centrality index of each node is the sum of the weights of the ties attached to it. We also consider DC′, which is the standardized index (*DC* divided by the sum of *DC* for all nodes).

We also examine *betweenness centrality* as one of the measures used to characterize the First2 Network because it describes the frequency with which an actor falls between pairs of other actors on the shortest path connecting them. Finally, we consider the *information centrality*, which explains the proportion of total information flow in a social network controlled by a person. The more connections there are, the more observations a person has on specific information, and the smaller the variance of the index.

We also compute *social conductivity* and *robustness* (Buch 2019) for the First2 Network. Social conductivity for a pair of actors is a number that represents a combination of all the paths between those two persons (nodes). This is an application of the series/parallel resistor formula taught in introductory physics. Social conductivity is then averaged over all actors in the network. Robustness looks at all the paths between two persons and deletes any direct connection between them and then combines the numbers from the other paths. Thus, robustness measures the impact on conductivity from indirect paths and is averaged over all actors in the network. The details of the calculations can be found in Appendix A.

Results

In this section, we will outline the findings of the analyses of the data from the social network survey. We divide the results into three sections. First, a look at the First2 Network as it was being built during the first few months of the project; this is considered the baseline. Next, we look at the network in the next two years to see how new connections are changing. Finally, we compare the analysis results from the three years to determine how things have changed and offer possible explanations as to why these changes may have occurred. In this section, the graphs in figures 1, 2, and 3 are drawn without any isolated nodes, where the size and color of the node indicates the number and strength of the connections. The persons with more connections are central to the graph, and the ones with fewer or only one connection are further from the center. The isolated nodes were removed from the graph to make the figures cleaner but are accounted for in the text. It was decided to number members the same each year of the analyses so we could track personal growth for individuals. Therefore, this sometimes resulted in isolated nodes, if an individual did not fill out the survey or was not named by anyone in subsequent years.

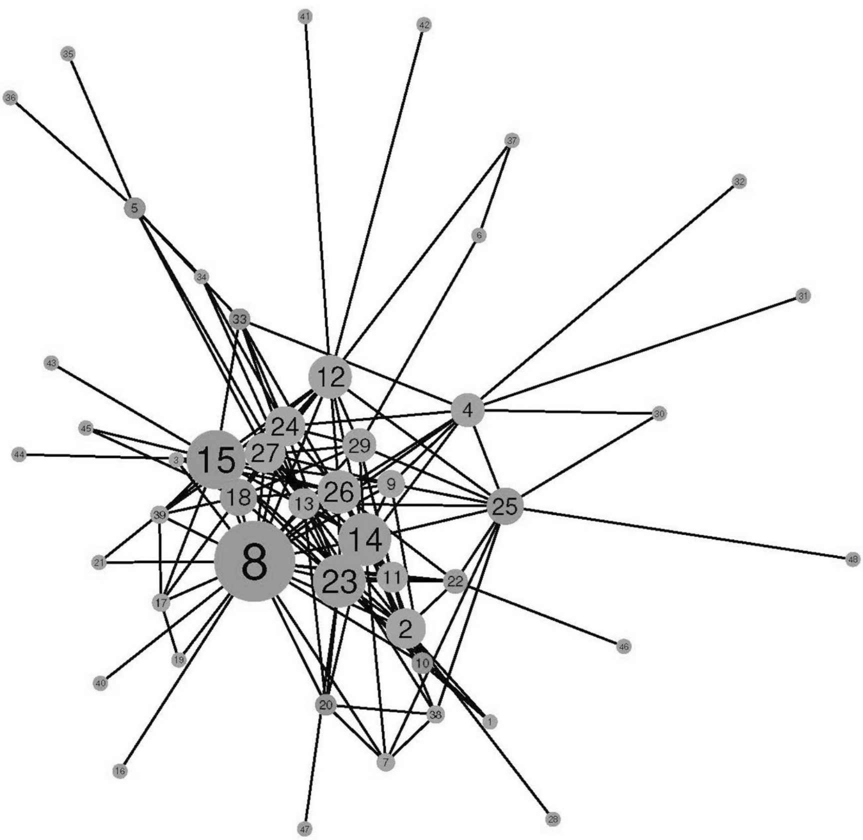
Year 1: Building a Statewide Network

In 2018, there are forty-eight nodes (people) and 146 ties (connections) in the First2 Network. In figure 1, the nodes are sized by degree centrality; the more central nodes are the ones with the highest weighted degrees. From figure 1, we can see that there are several actors who are central to the network at this time. The four largest nodes (8, 14, 15, and 23) are people with strong ties to many others. They represent four different organizations, two universities, one nonprofit, and one statewide research organization. These people are also central to the project in their roles; three are principal investigators (PIs), and the other person is a faculty member engaged in helping to set up the summer research experiences for students. This pattern shows coordination among institutions and agencies within the state. The graph density is 0.13, meaning that approximately 13 percent of all possible ties are present. The average social conductivity is 2.79, and the average robustness in this graph is 2.55.

Network as It Exists in Year 2

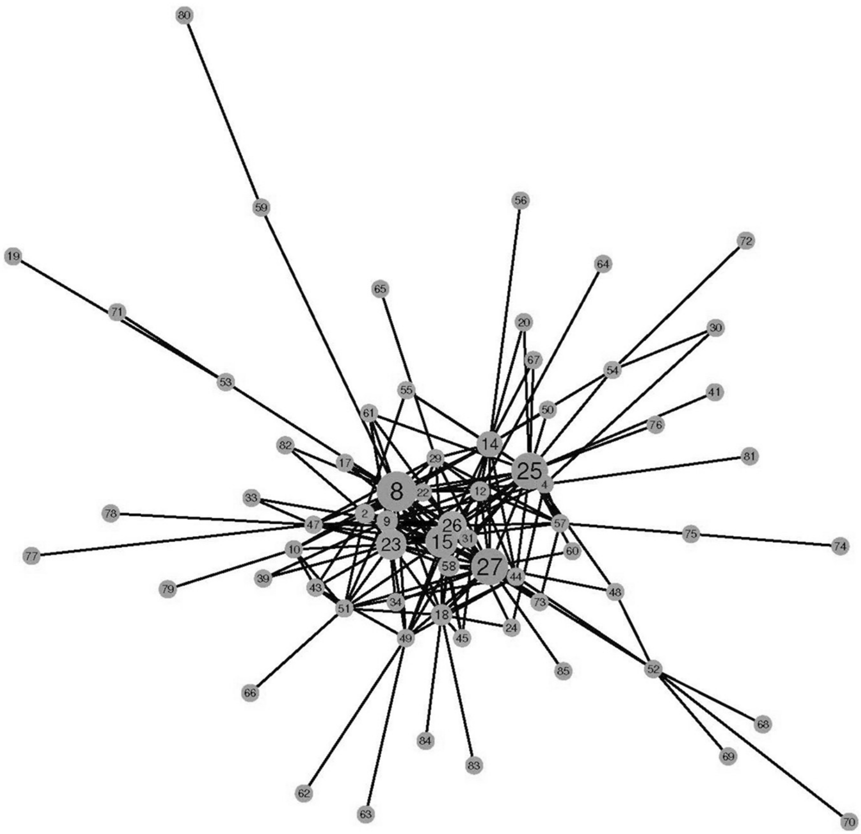
In 2019, there are eighty-five nodes (people) and 183 ties (connections) in the First2 Network overall, but figure 2 shows only the ones who

Figure 1: 2018 First2 Network graph where more central nodes have higher weighted degrees with nodes sized by degree centrality. No isolated nodes.



claimed connections to others. Eighteen isolated nodes were removed, leaving sixty-seven nodes and 183 ties in the graph in figure 2. The isolated nodes were people from 2018 who did not respond to the survey and were not named by anyone who did. However, this means that thirty-seven new members were added to the network. The graph nodes are sized as mentioned before. In figure 2, some of the four largest nodes (8, 15, 25, 27) have changed from the previous year. They still represent four different organizations, two universities, one nonprofit, and one K–12 school district. These people are also central to the project in their roles; two are PIs, one represents a K–12 school district in the state, and the last is a faculty member who chairs one of the network improvement community groups. There are other people emerging as leaders across the state. The graph density is 0.08, meaning that approximately 8 percent of all possible

Figure 2: 2019 First2 Network graph where more central nodes have higher weighted degrees with nodes sized by degree centrality. No isolated nodes.

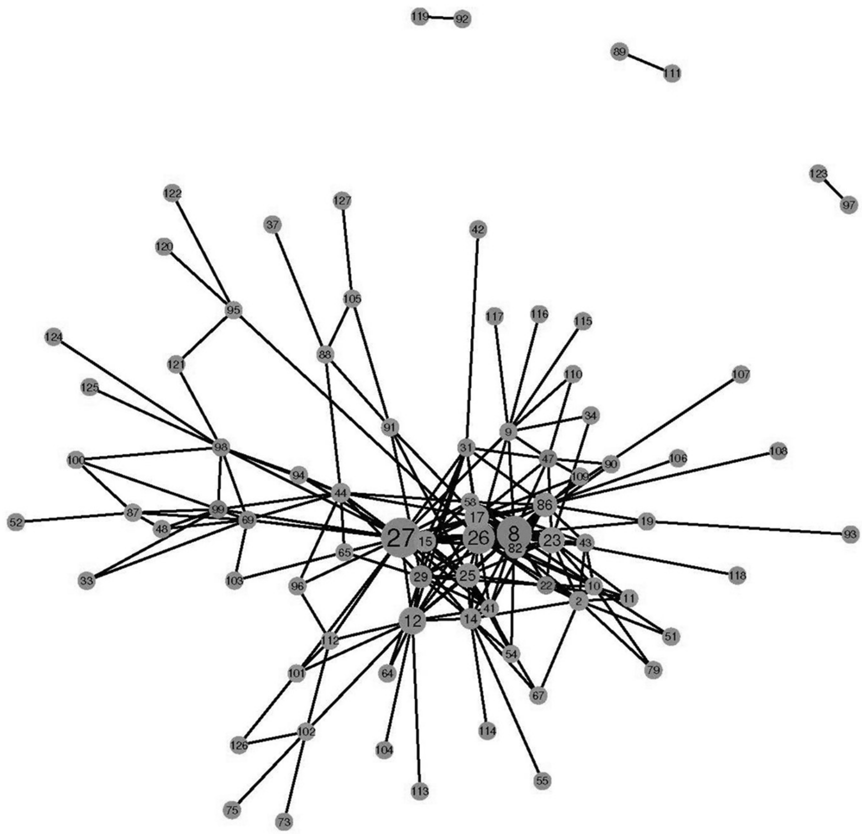


ties are present. The average social conductivity is 2.36, and the average robustness in this graph is 2.20.

Network as It Exists in Year 3

In 2020, there are 127 nodes (people) and 210 ties (connections) in the First2 Network overall, but figure 3 shows only the ones who claimed connections to others. Forty-six isolated nodes were removed, leaving eighty-one nodes and 210 ties in the graph in figure 3. The isolated nodes were people from 2018 and 2019 who did not respond to the survey and were not named by anyone who did. There are many explanations for why people appear as isolated nodes. Some were students who graduated and moved on, others didn't bother to fill out the survey in the given year, and still

Figure 3: 2020 First2 Network graph where more central nodes have higher weighted degrees with nodes sized by degree centrality. No isolated nodes.



others may have changed the people they were working with, thus not naming former associates when asked for connections. On the flip side, it is noteworthy that in this year, forty-two first-time people were added to the network. The graph nodes are sized and colored as mentioned before. In figure 3, some of the four largest nodes (8, 12, 26, 27) have changed again from the previous years. They still represent four different organizations, two universities, one statewide research organization, and a K–12 school district; two are PIs of the project, one represents a school district in the state, and the last is a faculty member who has taken an important role in the summer research experiences for students in the state. While the top four people changed over the three years, there are eight people represented in these largest nodes. We interpret this as leadership being spread out among the members of the network and communications being distributed throughout, so that network members have more than just a few central members from which to receive information. The graph density is 0.07, meaning that approximately 7 percent of all possible ties are present. The average social conductivity is 2.52, and the average robustness in this graph is 2.36.

Longitudinal Look

This section looks at several of the measures mentioned above and compares the measures between the three years.

Degree centrality (DC) quantifies how many connections a person has to other people in the network. Mathematically, in our undirected graphs, the degree centrality index of each node (person) is the sum of the weights on the ties (connections) attached to it (person). The DC metrics were computed on the graph without isolated vertices and isolated subgraphs (there were only a few isolated subgraphs of size two). Table 2 shows a comparison of 2018, 2019, and 2020, related to degree centrality. The DC index for each node, u, is the sum of the weights on the ties associated with that node (person). The DC sum then indicates the sum over all the nodes in the network. For the First2 Network, the DC sum grew from 521 in 2018 to 708

Table 2: Degree centrality comparison (without isolated nodes)

2018	2019	2020
Number of Nodes (N) = 48	Number of Nodes (N) = 67	Number of Nodes (N) = 81
DC Sum = 521	DC Sum = 708	DC Sum = 880
DC Mean = 10.850	DC Mean = 10.567	DC Mean = 11.733
Max DC′ = 0.102 (node 8)	Max DC′ = 0.080 (node 8)	Max DC′ = 0.063 (node 8)
DC′ Mean = 0.021	DC′ Mean = 0.015	DC′ Mean = 0.013

in 2019 and then to 880 in 2020, indicating that more and stronger connections are being made throughout the network. An individual's DC′ is the standardized index, computed by an individual's DC divided by DC sum; this number is always between 0 and 1. The person with the max DC′ in the First2 Network was the same for all three years; this person was one of the PIs and is responsible for working with the other faculty on the project. That person's DC′ number went down and the DC′ mean went down for the network from 2018 to 2020; this could be because the DC sum went up overall and the total network connections were getting stronger. The DC mean stayed consistent for the first two years and then increased in 2020, indicating that average connections were increasing for the overall network. To summarize, in our weighted network, DC quantifies the number and weights (strengths) of the connections that are forming; the fact that the DC sum and the DC mean are increasing indicates that the people who are staying actively involved are forming more and stronger connections.

Betweenness centrality (BC) is used to describe the frequency with which an actor falls between pairs of other actors on the shortest path connecting them. The BC index of a node u is the sum of $\delta(s,t,u)$ for all $s, t \in V$ (set of nodes) where $\delta(s,t,u)$ is a ration of all geodesics between s and t which run through u. BC′ is the standardized index (BC divided by $(N-1)(N-2)/2$ in symmetric nets or $(N-1)(N-2)$ otherwise). Table 3 shows that the BC sum for the First2 Network increased greatly from 2018 to 2020. This indicates that even though new members were being added each year, people were becoming *interconnected*, with more individuals falling between pairs of other people. This interconnectedness would help with dissemination of information. The person with the max BC′ was the same for 2018 and 2019, a state leader known by many around the state, but changed in 2020 to a person who was a member of many of the working groups. The BC′ sum increased from 2018 to 2020, indicating that connectivity became denser and the BC′ mean stayed consistent for the three years, only slightly decreasing.

Information centrality (IC) explains the proportion of total information flow in a social network controlled by a person. The more connections there are, the more observations a person has on specific information, and the

Table 3: Betweenness centrality comparison (without isolated nodes)

2018	2019	2020
BC Sum = 2141.220	BC Sum = 5292.710	BC Sum = 7682.245
Max BC′ = 0.261 (node 8)	Max BC′ = 0.302 (node 8)	Max BC′ = 0.319 (node 41)
BC′ Sum = 1.981	BC′ Sum = 2.467	BC′ Sum = 2.844
BC′ Mean = 0.041	BC′ Mean = 0.037	BC′ Mean = 0.038

smaller the variance of the index. The IC index measures the information flow through all paths between actors weighted by strength of connection and distance; IC′ is the standardized index (IC divided by the sum IC), and all isolated nodes are dropped in the computation. Table 4 shows that IC sum increased each year, indicating that more people shared in controlling the information flow. The max IC′ changes to different people and decreases each year from 2018 to 2020, indicating that more people are sharing in the information flow of the network.

Maximal cliques analysis tells the number of subgroups that are strongly connected within the overall graph. A clique is the largest subgroup of people in the social network who are all directly connected to each other. The word maximal means that for each clique, the group of its members is expanded to include as many actors as possible; no other actors can be added to the clique who would share connections to every other member.

Nodes (people) can belong to more than one clique, and every node belongs to at least one clique (itself—clique of size 1). The Bron and Kerbosch (1973) algorithm can be used to find all maximal cliques in an undirected graph. The output is a census or list of all maximal cliques of each possible size. The clique census report includes the size of the clique and the node membership in the clique.

We considered the maximal cliques in the graph to determine the sizes of subgroups that were forming and looked at the membership of the larger subgroups across the years. Table 5 shows the size of the cliques that formed and the number of these cliques for each year. From the table, we can see that

Table 4: Information centrality comparison (without isolated nodes)

2018	2019	2020
IC Sum = 58.295	IC Sum = 61.28	IC Sum = 72.98
Max IC′ = 0.032 (node 12)	Max IC′ = 0.024 (node 20)	Max IC′ = 0.021 (node 58)
IC′ Mean = 0.021	IC′ Mean = 0.015	IC′ Mean = 0.013

Table 5: Size of cliques with number of cliques for each size

Size of Clique	2018	2019	2020
3	195	242	269
4	152	215	243
5	62	111	140
6	10	29	46
7	0	3	7

as the years progressed, more tightly connected subgroups were forming within the overall structure of the network. All sizes of cliques from three through seven increased over the years. The composition of the largest groups was examined to see if these groups were made up of people from the same organization or from organizations across the state. We found that in every year, the largest groups were composed of people from multiple organizations. Project leaders were evident in these large groups, but there was a range of other people who appeared in the groups. This can be attributed to the project's emphasis on working groups around various topics and the encouragement to share information across institutions.

Discussion

We looked at the results above to answer the two research questions posed in the introduction to this article: (1) How and in what ways is the network growing over time? and (2) To what extent do the interconnectedness and information flow improve among members? These preliminary findings indicate that the First2 Network is growing in number of connections, strength of the connections, interconnectedness, and number of closely knit groups that are forming. Network leadership members are playing key roles in the First2 Network. Comparing figures 1, 2, and 3, one can see from the size of the nodes and the number of edges present that there are more people with stronger connections. A comparison of the figures also shows that the number of small nodes on the perimeter of the network are increasing; these are likely new members. All these things indicate that the network has changed in several ways from 2018 to 2020.

Based on the research of Cross et al. (2009), network structure and tie strength affect outcomes such as transfer of knowledge, organizational change, and innovation. For the First2 Network, the change in the structure from 2018 to 2020 indicates that transfer of knowledge should be increasing. The number of people had a 165 percent increase, and connections had a 44 percent increase from 2018 to 2020, but the density of the graph (percentage of possible ties) decreased from 13 percent to 7 percent. This decrease in density was likely due to the new members who were being added each year who had less connections on average at the beginning. Social conductivity changed from 2.78 in 2018 to 2.52 in 2020, which shows that the average strength of the connections between people stayed very similar, between 2 (cooperation) and 3 (coordination), even though the network almost tripled in size from 48 nodes to 127 nodes. Robustness, which excludes all direct ties, decreased from 2.55 in 2018 to 2.36 in 2020. This implies that even with the increase of the number of nodes, the removal of direct ties between individuals will not disrupt the flow of knowledge through the network.

Another indicator that the flow of knowledge through the First2 Network should be increasing is the degree centrality (DC) sum that also increased from 521 in 2018 to 880 in 2020. Since the DC is the sum of the weights (strength of connection) on the ties associated with that node (person), this change indicates that connections are increased and/or strengthened throughout the network. The betweenness centrality (BC) sum also increased from 2,141 in 2018 to 7,682 in 2020, from fifty-eight to seventy-three nodes/people, indicating that interconnectedness is increasing; more people are on the paths between others to help with information flow. The information centrality (IC) sum also increased from 2018 to 2020, indicating that information flow through all paths between actors is getting better every year.

We were able to track forty-eight of the student participants from the First2 Network. Of these, thirty-nine were retained, and nine had dropped out or changed to a non-STEM major. When we looked at their connections to the network, the students who were retained in STEM had an average connection strength of 2.81, and the ones who had not been retained in STEM had an average connection strength of 2.28. This is by no means the only measure of persistence that we plan to use, but it is an early indicator of how the network may help students connect to resources that will support them in persisting.

This social network analysis (SNA) research serves as a way to track the First2 Network efforts of building connections throughout the state. Each year, the research team presents results of the growth of the network to the leadership of the project to help them assess the growth and determine ways to improve. The network has added features and resources to the website to encourage individuals to sign up. They have increased communication through the website, sending out announcements and weekly updates to keep people informed. The network community working groups have increased efforts to recruit new members, and the statewide gatherings are held twice per year, with working webinars held in between. We should note that the pandemic made these gatherings harder to achieve, and they had to hold them virtually for the last year discussed. All these efforts are adding ways for members to communicate and strengthen connections.

Future Work: Implications for STEM Reform in West Virginia

The main goal of the project is to help recruit and retain students in STEM majors and ultimately increase the STEM workforce in the state. The First2 Network is providing a way to make connections to accomplish this purpose. As academia, state agencies, and state industries come together, they can provide a framework for change in the state. The network can provide a means to exchange ideas and important information. It can also match people with resources and help identify where needs are not being met. For example, the network improvement communities that are forming are made up of people from various organizations who are sharing promising practices in the form of plan-do-study-act (PDSA) cycles. These PDSAs being replicated throughout the state at different organizations are just one example of how information is being spread due to the connections that are growing in the First2 Network.

Limitations and Future Research

As with any survey, the social network survey had limitations. For one, the number of respondents was relatively low. In 2018, 33 percent (22/66) of the people invited responded to the survey; in 2019, 24 percent (31/129) responded; and in 2020, the survey was sent to 621 people with only 44 people or 7 percent (44/621) responding after multiple attempts. The large list of people to receive the survey in 2020 was given to the researchers by one of the partner organizations. In all years, the survey was sent to everyone who was invited to a meeting or who had joined the project website. This was a very large list (especially in 2020) and did not represent people who actually engaged in the network. The survey email list represented people who may have heard of the network or people who were invited to join. We believe that people who actually engaged in the network would either fill out the survey or would be named by people who filled out the survey and would be counted. Thus, we think the data we used as a basis of the network analyses was representative of the active network in the project. However, the increased number of respondents and the increased number of people the respondents named show that the active members were indeed growing. Even though only a small portion of the invited people filled out the survey each year, many other people were named by the respondents, so the information was more complete than it may seem. For example, in 2018, twenty-two people filled out the survey, but twenty-six other people were named, bringing the network nodes to forty-eight. The survey also only allowed for a respondent to name ten other connections. This was due to the fact that the developers of the survey did not want to overburden the respondents since they were asked to give names, affiliations, and connections strengths for each connection named.

It is our intention to continue to administer the social network survey for the remaining two years of the project (2021–2022). This should allow us to determine how the First2 Network is changing and to inform the project team of our results.

Conclusion

From our analyses, it was determined that the First2 Network is growing in number of people, number of connections, and depth of connections. People from K–12, higher education, government, and industry are coming together under this project and are able to coordinate and collaborate around the goal of recruiting and retaining students in STEM majors. The growth of the network, which has been demonstrated through several measures, is vital to ensure that the activities and programs are disseminated and sustained throughout the state.

This analysis attempts to provide evidence to answer two main questions: (1) How and in what ways is the network growing over time? and (2) To what extent do the interconnectedness and information flow improve among members? We have seen the strengthening of connections between people in the entire network. The network improvement working groups have allowed leaders to emerge, and the social network analysis (SNA) has confirmed that many new people emerged as being central to the First2 Network. The number and strength of the ties, based on the research of Cross et al. (2009), would indicate that the transfer of information or knowledge should have increased from 2018 to the latter years. Overall, the First2 Network growth indicates that the project is moving toward the goal of forming a functioning network within a rural Appalachian state to support the recruitment and retention of students in STEM majors and then to assist with the matriculation of these students into STEM jobs in the state and beyond. The next step in this research is to determine how the state is doing as far as recruiting and retaining students in STEM majors.

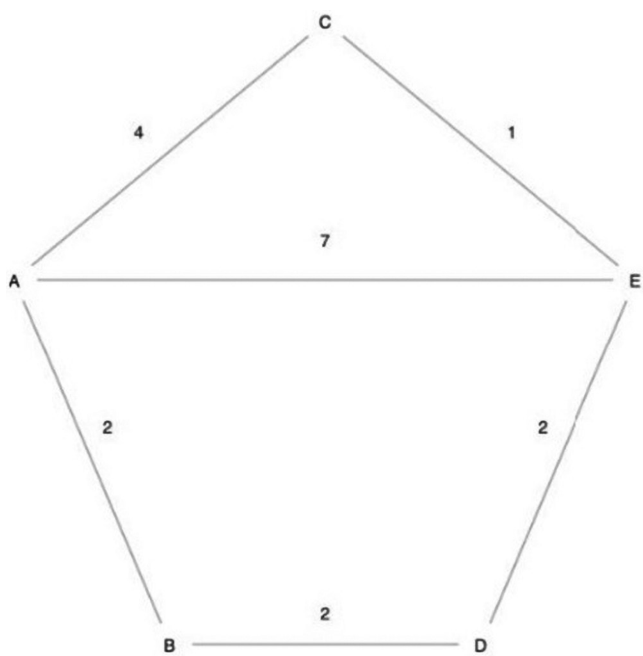
Finally, the significance of this research includes the introduction and application of two new SNA measures (social conductivity and robustness) and the use of SNA for the NSF INCLUDES program, which is focusing on building a national network. The new measures are presented along with existing measures to show the new information that could be gained by using them. Our work will help inform the NSF INCLUDES National Network, of which this project is part. This large national network brings together the many agencies to strengthen STEM equity by connecting individuals, alliances, pilot programs, federal agencies, educational institutions, and other entities across the nation working to shift inequitable systems and broaden participation in STEM education and careers.

Appendix A

Social Conductivity

A frequent object of interest in social network analysis is the extent to which a pair of nodes are connected or disconnected. This question has motivated several measures of path value, which are then used as a scale to identify optimal paths between nodes, and the value of these optimal paths is used as a proxy for the connectedness of a pair (Yang and Knoke 2001). An obvious shortcoming of this strategy is a failure to account for the extent to which two nodes are connected indirectly. If pair A and pair B are close friends, but the members of pair A share numerous other common friends, whereas the members of pair B have no mutual friends, it would be inaccurate to describe both pairs as being equally well-connected. We would like to account for all paths joining pairs of interest. One physical setting in which we have available precise tools for the accounting of multiple paths of transmission is electrical networks. There is a similarity between social ties and conductive elements, and there are well-developed mathematics for describing total conductivity between points in networks of conductors. Accordingly, we propose social conductivity as a useful measure for summarizing all-paths connectedness between nodes.

Figure 4: Tie strength, optimal connections, and distance in social networks.



To see a little of how this works, we consider figure 4. Using Ohm's law and treating ties as conductive elements with conductivity proportional to tie strength, we can find the social conductivity between nodes *A* and *E* as follows: To obtain the effective resistance of a series of resistors like *AC*, *AE*, we add their respective resistances. Since we are treating the tie weights as conductivities, we will use harmonic addition instead, so we see that path *ACE* has conductivity

$$\left(\frac{1}{4}+\frac{1}{1}\right)^{-1}=\frac{4}{5}$$

Similarly, path *ABDE* has conductivity

$$\left(\frac{1}{2}+\frac{1}{2}+\frac{1}{2}\right)^{-1}=\frac{2}{3}$$

Finally, path *AE* simply has conductivity 7.

Next, recall that the effective resistance of parallel resistors is the harmonic sum of the component resistances. Therefore, working with conductivities, we simply take the sum of the component conductivities. Therefore, the conductivity between nodes *A* and *E* is the sum of the conductivity of the three paths joining these nodes,

$$\left(\frac{4}{5}+\frac{2}{3}+7\right)=\frac{127}{15}\approx 8.5$$

The key properties of social conductivity are that no path is more conductive than its least conductive tie, and the total conductivity of two disjoint parallel paths is the sum of their individual conductivities.

Matrix techniques are required to account for all paths in highly complex networks. The calculations for these networks is described in the manual for the R package for electrical engineering resistor array (Hankin 2006). Notice that, in its current definition, social conductivity can only operate on symmetric networks.

Robustness is calculated by using the process above but eliminating the direct ties between two individuals. This measure captures the extent to which strong indirect connections are available between actors in case direct ties are lost or unavailable.

Notes

1. See https://wvforward.wvu.edu/; http://educationalliance.org/; and https://wvpec.wvu.edu/ for more information about these groups.

2. Inclusion across the Nation of Communities of Learners of Underrepresented Discoverers in Engineering and Science (INCLUDES). For more information, see https://www.includesnetwork.org/about-us/who-we-are.

3. The Durant-Law and Milne survey is no longer available online.

References

American College Testing (ACT). 2016. The condition of STEM 2016: West Virginia. http://www.act.org/content/dam/act/unsecured/documents/STEM2016_49_West_Virginia.pdf.

———. 2017. The state of STEM in your state—2017 West Virginia state report. https://www.act.org/content/dam/act/unsecured/documents/STEM/2017/West-Virginia-State-of-STEM-2017.pdf.

Baker-Doyle, Kira J., and Susan A. Yoon. 2011. In search of practitioner-based social capital: A social network analysis tool for understanding and facilitating teacher collaboration in a US-based STEM professional development program. *Professional Development in Education* 37 (1): 75–93. https://doi.org/10.1080/19415257.2010.494450.

Borden, Lynne M., and Daniel F. Perkins. 1999. Assessing your collaboration: A self-evaluation tool. *Journal of Extension* 37 (2): 67–72.

Bron, Coen, and Joep Kerbosch. 1973. Algorithm 457: Finding all cliques of an undirected graph. *Communications of the ACM* 16 (9): 48–50.

Buch, David Anthony. 2019. *Tie strength, optimal connections, and distance in social networks*. MS thesis, West Virginia University.

Carman, Joanne G., and Kimberly A. Fredericks. 2018. Applications of social network analysis in evaluation: Challenges, suggestions, and opportunities for the future. *Canadian Journal of Program Evaluation* 33 (2): 175–87.

Chen, Xianglei, and Matthew Soldner. 2013. STEM attrition: College students' paths into and out of STEM fields. Statistical analysis report. National Center for Education Statistics. https://nces.ed.gov/pubs2014/2014001rev.pdf.

Chen, Xianglei, and Thomas Weko. 2009. Students who study science, technology, engineering, and mathematics (STEM) in postsecondary education. *Stats in Brief*. NCES 2009-161. https://nces.ed.gov/pubs2009/2009161.pdf.

Cross, Jennifer Eileen, Ellyn Dickmann, Rebecca Newman-Gonchar, and Jesse Michael Fagan. 2009. Using mixed-method design and network analysis to measure development of interagency collaboration. *American Journal of Evaluation* 30 (3): 310–29.

Eagan, M. Kevin, Jr., Sylvia Hurtado, and Mitchell J. Chang. 2010. What matters in STEM: Institutional contexts that influence STEM bachelor's degree completion rates. Annual meeting of the Association for the Study of Higher Education, Indianapolis, IN. https://www.heri.ucla.edu/nih/downloads/ASHE%202010%20-%20Eagan%20Hurtado%20Chang%20-%20STEM%20Completion.pdf.

Fredericks, Kimberly A., and Maryann M. Durland. 2005. The historical evolution and basic concepts of social network analysis. *New Directions for Evaluation* 107: 15–23.

Guralnick, Michael J., Robert T. Connor, and L. Clark Johnson. 2009. Home-based peer social networks of young children with Down Syndrome: A developmental perspective. *American Journal on Intellectual and Developmental Disabilities* 114 (5): 340–55.

Halgin, Daniel S., and Stephen P. Borgatti. 2012. An introduction to personal network analysis and tie churn statistics using E-NET. *Connections* 32 (1): 37–48.

Hankin, Robin K. S. 2006. Resistor networks R: Introducing the ResistorArray package. *Newsletter of the R Project* 6/2: 52–53. http://www.stat.uchicago.edu/~pmcc/pubs/Rnews_2006-2.pdf.

Hogue, Todd, Daniel Perkins, Richard Clark, Arno Bergstrum, Margaret Slinski, and Associates. 1995. *Collaboration framework: Addressing community capacity*. Columbus, OH: National Network for Collaboration. https://www.uvm.edu/sites/default/files/media/Collaboration_Framework_pub.pdf.

Honeycutt, Todd. 2009. Making connections: Using social network analysis for program evaluation. Issue Brief. Number 1. Mathematica Policy Research. https://files.eric.ed.gov/fulltext/ED507482.pdf.

Igarashi, Tasuku, Jiro Takai, and Toshikazu Yoshida. 2005. Gender differences in social network development via mobile phone text messages: A longitudinal study. *Journal of Social and Personal Relationships* 22 (5): 691–713.

IGEN 2022: IGEN. 2022. National Science Foundation. http://igenetwork.org/org/national-science-foundation (accessed October 7, 2022).

Kezar, Adrianna. 2014. Higher education change and social networks: A review of research. *Journal of Higher Education* 85 (1): 91–125. https://doi.org/10.1353/jhe.2014.0003.

Knaub, Alexis V., Charles Henderson, and Kathleen Quardokus Fisher. 2018. Finding the leaders: An examination of social network analysis and leadership identification in STEM education change. *International Journal of STEM Education* 5 (1): 1–14. https://doi.org/10.1186/s40594-018-0124-5.

Ma, Shufeng, Geoffrey L. Herman, Jonathan H. Tomkin, Jose P. Mestre, and Matthew West. 2018. Spreading teaching innovations in social networks: The bridging role of mentors. *Journal for STEM Education Research* 1 (1): 60–84.

Ma, Shufeng, Geoffrey L. Herman, Matthew West, Jonathan Tomkin, and Jose Mestre. 2019. Studying STEM faculty communities of practice through social network analysis. *Journal of Higher Education* 90 (5): 773–99. https://doi.org/10.1080/00221546.2018.1557100.

McCann, Meghan, and Tom Keily. 2021. Perceptions of postsecondary education and training in rural areas. Education Commission of the States. https://www.ecs.org/perceptions-of-postsecondary-education-and-training-in-rural-areas/ (accessed May 28, 2021).

National Science Foundation. 2016. First NSF INCLUDES awards issued to 37 projects. https://www.nsf.gov/news/news_summ.jsp?cntn_id=189706.

Noonan, Patricia, Amy Gaumer Erickson, Zach McCall, Bruce B. Frey, and Chunmei Zheng. 2014. Evaluating change in interagency collaboration of a state-level interagency education team: A social network approach within a utilization-focused framework. *Educational Assessment, Evaluation, and Accountability* 26 (3): 301–16.

Osborn, Richard N., and John Hagedoorn. 1997. The institutionalization and evolutionary dynamics of interorganizational alliances and networks. *Academy of Management Journal* 40 (2): 261–78.

Peri, Giovanni, Kevin Shih, and Chad Sparber. 2015. STEM workers, H-1B visas, and productivity in US cities. *Journal of Labor Economics* 33 (S1): S225–S255.

Shadle, Susan E., Yujuan Liu, Jennifer E. Lewis, and Vicky Minderhout. 2018. Building a community of transformation and a social network analysis of the POGIL project. *Innovative Higher Education* 43 (6): 475–90.

Showalter, Daniel, Robert Klein, Jerry Johnson, and Sara L. Hartman. 2017. Why rural matters 2015–2016: Understanding the changing landscape. Rural School and Community Trust. https://www.ruraledu.org/user_uploads/file/WRM-2015-16.pdf.

Southern Regional Education Board. 2018. West Virginia state progress report: Looking closer. https://www.sreb.org/sites/main/files/file-attachments/2018progress_wv.pdf?1529594302.

Thompson, Ruthanne, and Greta Bolin. 2011. Indicators of success in STEM majors: A cohort study. *Journal of College Admission* 212: 18–24.

van den Hurk, Anniek, Martina Meelissen, and Annemarie van Langen. 2019. Interventions in education to prevent STEM pipeline leakage. *International Journal of Science Education* 41 (2): 150–64.

US Department of Commerce. 2017. STEM jobs: 2017 update. https://www.commerce.gov/data-and-reports/reports/2017/03/stem-jobs-2017-update.

Wasserman, Stanley, and Katherine Faust. 1994. *Social network analysis: Methods and applications*. Vol. 8. Cambridge, UK: Cambridge University Press.

Watkins, Jessica, and Eric Mazur. 2013. Retaining students in science, technology, engineering, and mathematics (STEM) majors. *Journal of College Science Teaching* 42 (5): 36–41.

West Virginia Higher Education Policy Commission (WVHEPC). 2015. VISION 2025: The West Virginia science and technology strategic plan. https://wvresearch.org/wp-content/uploads/2015/08/2025-Strategic-Report.pdf.

West Virginia Higher Education Policy Commission (WVHEPC) and West Virginia Community and Technical College System. 2018. West Virginia higher education report card. http://www.wvhepc.edu/wp-content/uploads/2019/01/ReportCard_FINAL_8Jan2019.pdf.

———. 2019. Academic readiness report. https://www.wvhepc.edu/wp-content/uploads/2019/06/AcademicReadinessReport_2018_AMENDED_15Jan2019.pdf.

World Population Review. 2020. West Virginia population 2020. https://worldpopulationreview.com/states/west-virginia-population/ (accessed May 28, 2021).

Yang, Song, and David Knoke. 2001. Optimal connections: Strength and distance in valued graphs. *Social Networks* 23 (4): 285–95.