

Social Network Analysis and STEM Education

A Brief Overview & Demo

Dr. Joshua Rosenberg, University of Tennessee, Knoxville
2021/08/11 (updated: 2021-08-11)



Agenda: Two parts

1. Brief overview of network analysis (20 min.)
 - Introducing five ideas
2. Walkthrough of a network analysis learning lab (20 min.)
 - You running code using RStudio Cloud for a social network analysis

1. Short presentation

1. Overview

1. Why use network analysis?
2. What is network analysis?
3. What are common techniques used for network analysis?
4. What are some of the technical and ethical considerations?
5. Summary

1.1. Why use network analysis?

Networks matter

Personally

- *Choosing* to interact with certain departmental colleagues expanded my network
- My friend Ryan Estrellado sharing about Ted Lasso *influenced* me

For young people

- Peers *impact* young people (and students) directly and indirectly
- Older students *selecting* particular students to reply to in a discussion forum

For educators

- More experienced colleagues *shape* what teachers perceive to be effective
- New ideas about teaching *diffusing* through social media

What are some aspects of networks that may matter to your own life—and research and teaching?

But, we often focus on individuals

- This is for good reason in STEM education research—as individual educator and student characteristics matter
- But, there is often a great deal of unexplained variation in the things we care about, even when we account for *dependencies* at the multiple levels at which educators and students are impacted

	Week 4 (Ave & Std. Dev.)	Week 13 (Ave & Std. Dev.)	Matched- Pair <i>t</i> (<i>df</i> = 14)	<i>p</i> -Value	Cohen's <i>d</i>
Time and Effort Questions					
T1 – Overall, I have been working very hard in this course	5.80 (1.27)	2.07 (1.16)	7.30	< .001*	3.16
T2 – I have spent a lot of my time and effort doing the readings in this course	5.53 (1.06)	2.47 (1.06)	6.63	< .001*	1.86
T3 – I have spent a lot of time and effort in the online discussions for the course	5.33 (1.23)	2.80 (1.15)	4.60	< .001*	1.94
T4 – I have spent most of my time and effort working alone and independently	5.00 (1.07)	3.80 (1.86)	2.00	= .067	0.37
T5 – I have spent most of my time and effort working in groups	4.53 (1.25)	2.87 (1.19)	3.25	< .01*	1.25

Koehler and Mishra (2005)

Relationships matter

- One place to look is at the **interdependencies** among educators, students, and others
- In other words, relationships are another potential source of variation
 - They may also be a phenomenon of interest independent of how well they explain variation
- There is an extensive tradition of qualitative lines of inquiry on how interactions and relations matter and social network analyses may complement these analyses

Idea #1: Though we often focus on individuals (or the dependencies between individuals), social network analysis allows us to focus on relationships between individuals

Frank (1998), Schwarz et al. (2009)

Zoom poll question

Question 1: *Which of the following has to do with a social relationship?*

1.2. What is social network analysis?

All about relationships

- We can think of network analysis as a collection of techniques for examining relationships which can take many different forms:
 - Face-to-face relationships between educators
 - Online relationships among students
- The data we use as input, too, can take different forms:
 - Self-reported relationships (e.g., [Coburn et al., 2012](#), [Horn et al., 2020](#), and [Spillane et al. \(2012\)](#))
 - Relationships inferred from digital sources of data (e.g., [Chen and Huang, 2020](#), [Hu et al., 2018](#), [Kellogg et al., 2014](#), and [Staudt Willet & Carpenter, 2020](#))

Idea #2: The kinds of relationships we investigate depend on our research questions—prior research and theory can guide us

Not one method

- Moreover, the technique we can use to study these relationships take different forms:
 - *Visualizing networks* through network visualizations, or sociograms
 - *Describing networks* through the use of individual or network statistics
 - *Using network models* for key processes, such as selection and influence
- These methods go by many names:
 - Social network analysis (SNA)
 - Network analysis
 - Network science

Idea #3: Social network analysis is a collection of methods for visualizing, describing, and modeling social network data— which technique(s) you choose depends on the claims you want to make

What are some ways you could gather data on relationships?

1.3. What are common techniques used for social network analysis

Collecting and Creating Data

- **Self-reported relationships:** Asking people about who they turn to for advice about teaching mathematics ([Horn et al., 2020](#))
 - The *nominator* refers to the person asked
 - The *nominees* refer to the people the nominator indicated they turn to for advice
 - What can we say about those who are the nominees of many people?

Collecting and Creating Data

- **Inferred relationships from digital data:** Using digital social media data as records of who interacted with whom ([Rosenberg et al., 2020](#))
 - The *nominator* refers to the person who initiated—or *sent*—an interaction
 - The *nominee* refers to the person who was the recipient of the interaction
 - What can we say about those who are the nominees of many people?

Idea #4: Relationships can be measured using self-reported and—increasingly—digital data sources

Analyzing data

- For analyzing social network data (self-reported or digital), R is open-source, free, and feature-rich, especially with respect to network analysis
 - tidygraph for creating graph "objects"
 - ggraph for visualizing networks
 - the statnet collection of packages for preparing and modeling data, including ergm

What are your goals for analyzing social network data?

1.4. What are some of the technical and ethical considerations for social network analysis?

Technical considerations

- In most analyses, preparing data takes time; this is (perhaps especially) true of network for social network analyses
- Regardless of whether the data is self-reported or digital, a useful way to structure data is in terms of an **edgelist**
 - An edgelist contains two columns, one for the *nominator* and one for the *nominee*
 - In most cases, both nominators and nominees appear many times in each of the two columns
 - Additionally, there may be other columns that represent information on the particular relationship

Idea #5: An edgelist is a key data structure that can position us to visualize, describe, and model social network data

Ethical considerations

- When collecting self-report data, a high response rate is often important
 - To achieve this, building trust with experts working within schools and districts is often essential
- Privacy can matter a great deal for social network data (self-reported *and* digital)
 - Consider anonymizing individuals' names or do not include their names
- Follow Institutional Review Boards guidelines, but consider ethics beyond what IRBs mandate
 - Ask, *should I collect this data?* What are the *risks* and *potential benefits*?
 - Balance open science and privacy is a challenge

How might ethical challenges intersect with your social network analysis research?

1.5. Summary of ideas

- Though we often focus on individuals (or the dependencies between individuals), **social network analysis allows us to focus on relationships between individuals**
- The kinds of relationships we investigate depend on our research questions—prior research and theory can guide us
- Social network analysis is a collection of methods for visualizing, describing, and modeling social network data—**which technique(s) you choose depends on the claims you want to make**
- Relationships can be measured using self-reported and—increasingly—digital data sources
- **An edgelist is a key data structure** that can position us to visualize, describe, and model social network data

Walkthrough

Let's head over to [the AERA-ICPSR LASER Workshop](#) space in RStudio Cloud to start the walkthrough.

Click on `sna-demo.Rmd` *within the Learning Lab Demos project* (the same one you have been using) to begin!

Conclusion

Thank you for attending our workshop today!

Resources

- [Learning Analytics Goes to School](#)
- [Learning Analytics 101](#)
- [Data Science in Education Using R](#)

The LASER Institute

- [Homepage and interest form](#)
- [Resources for this workshop](#)
- [Resources for the LASER Institute](#)

Contact

Homepage and interest form

- Shaun B. Kellogg, North Carolina State University
- Rob Moore, University of Florida
- Shiyang Jiang, North Carolina State University
- Joshua M. Rosenberg, University of Tennessee, Knoxville
- Jennifer Houchins, North Carolina State University