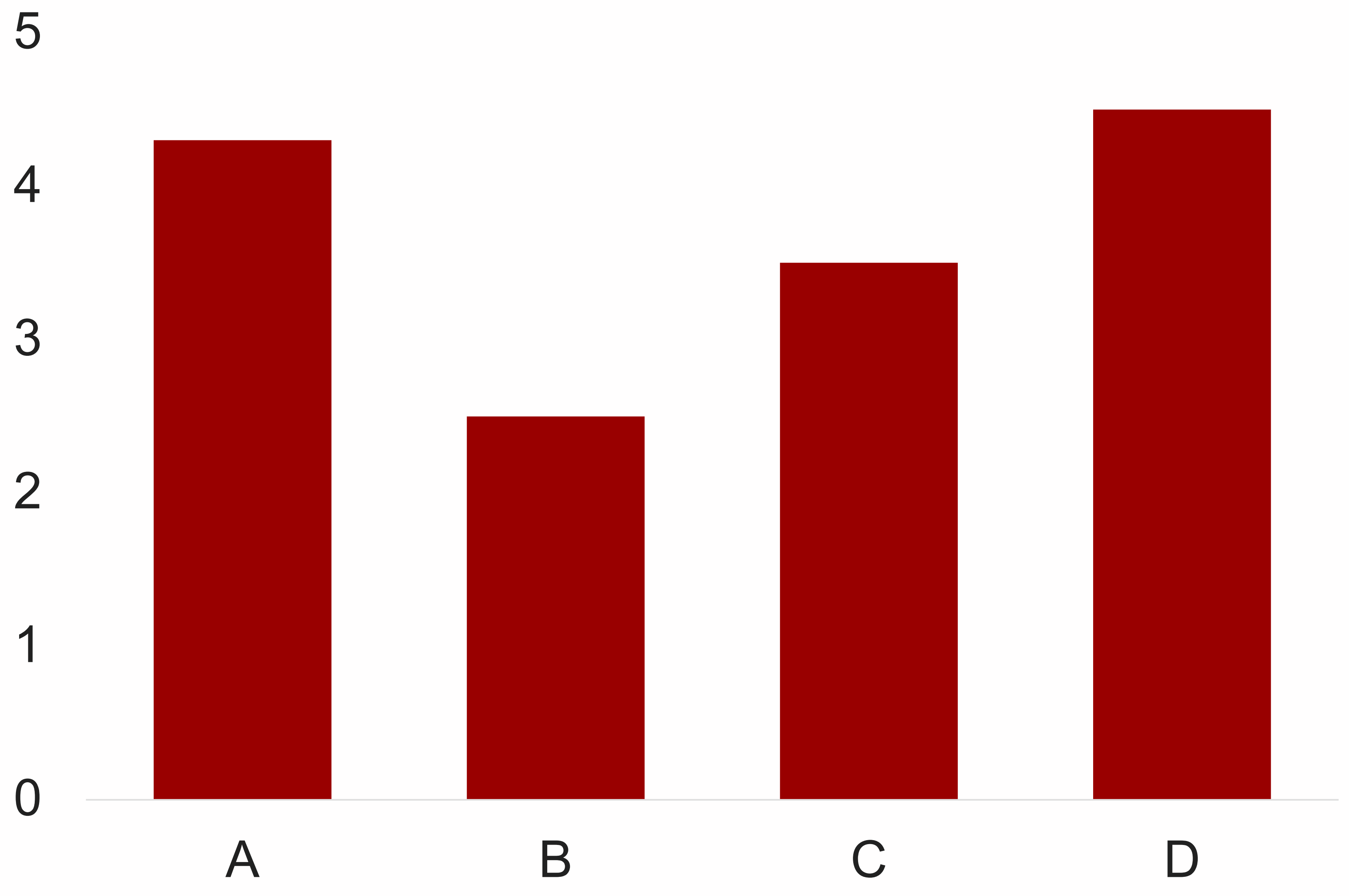
Network Visualization Literacy

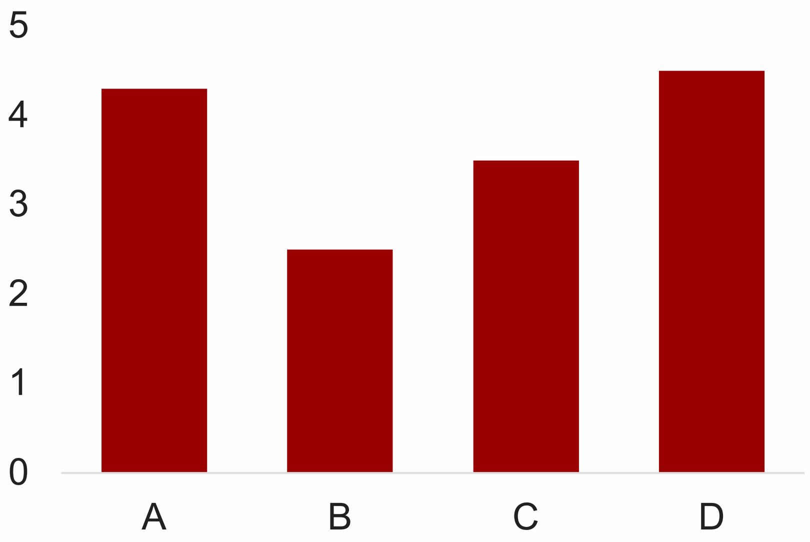
Angela Zoss

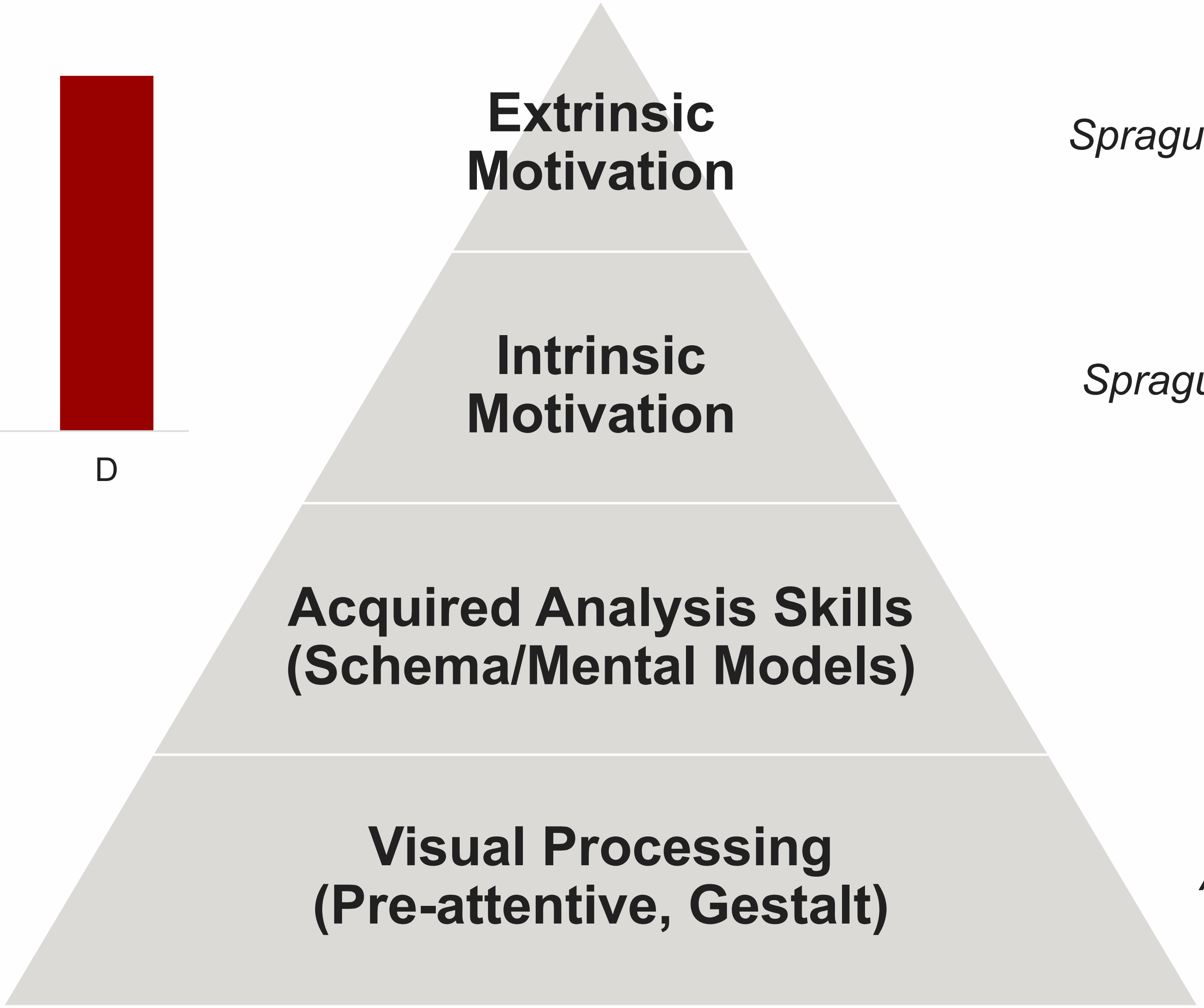
Friday, March 23, 2018

Duke Visualization Friday Forum

netvislit.org


5
4
3
2
1
0
A
B
C
D



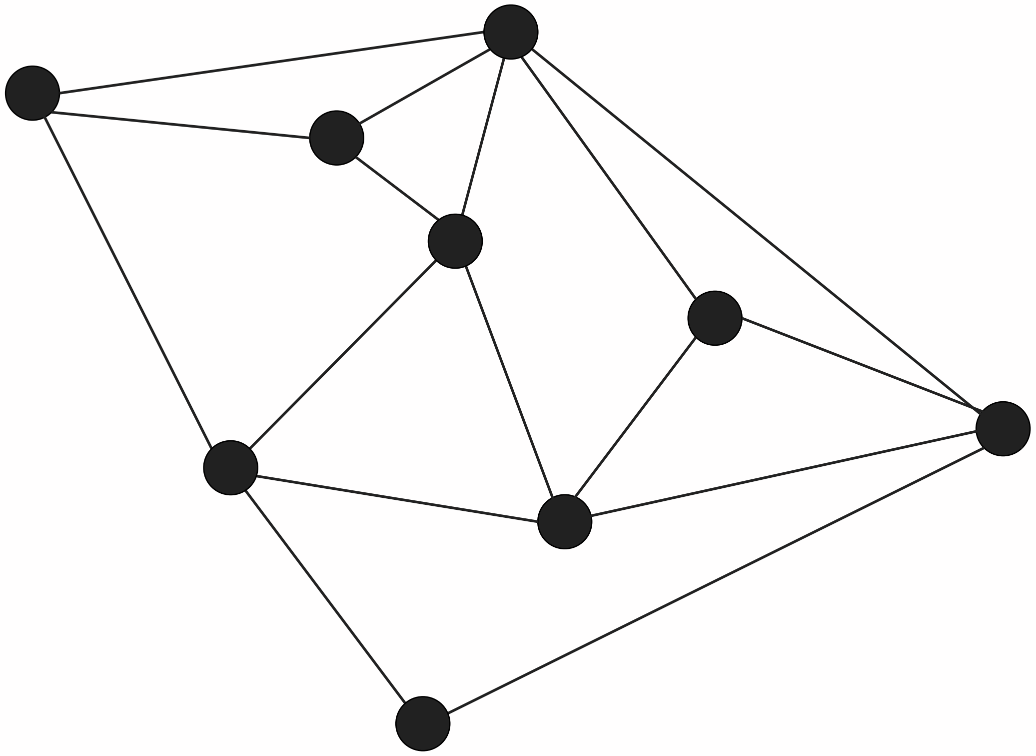


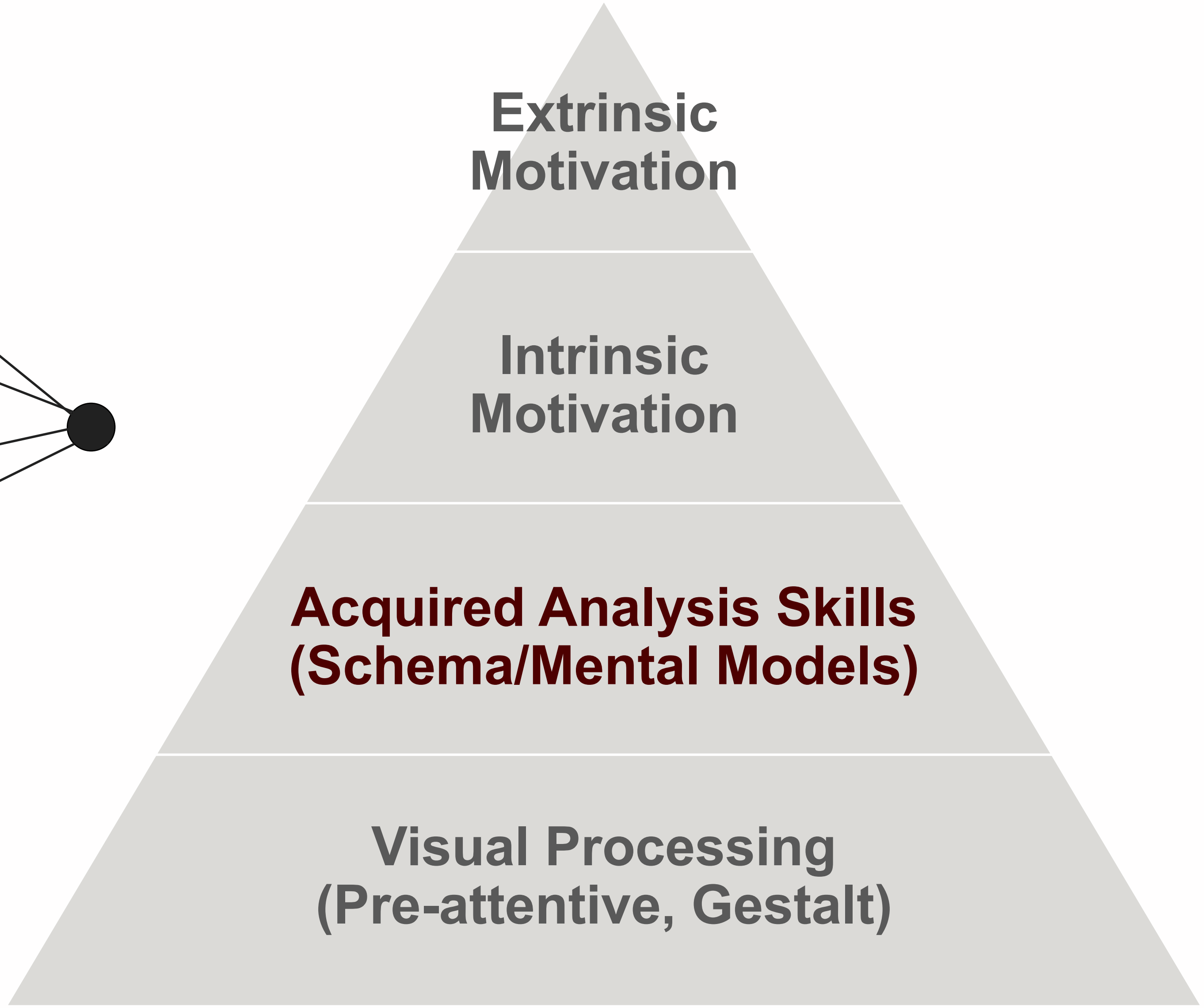
Sprague & Tory, 2012

Sprague & Tory, 2012
Du et al., 2017

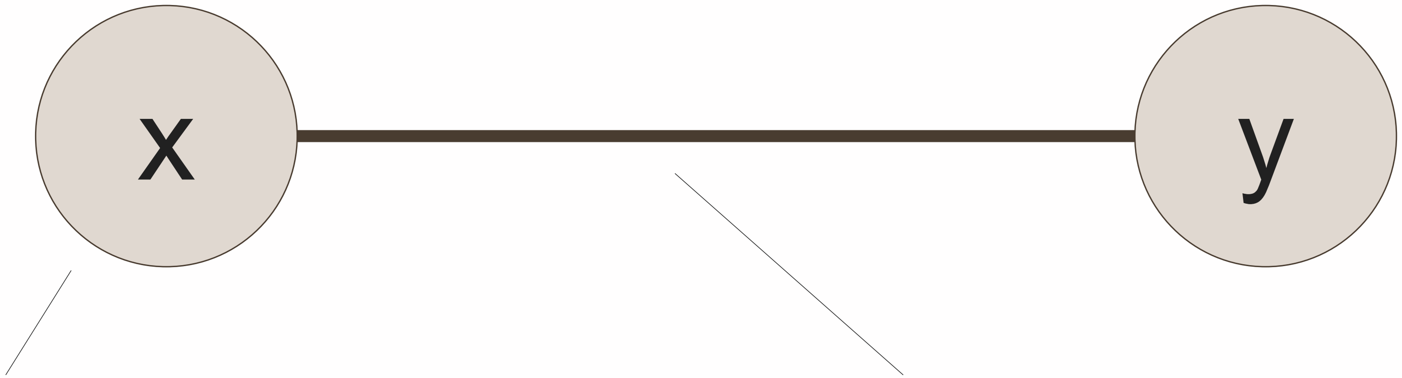
Zoss, 2014

Ware, 2012
Arnheim, 1969
Healy, 2012




Extrinsic Motivation
Intrinsic Motivation
Acquired Analysis Skills (Schema/Mental Models)
Visual Processing (Pre-attentive, Gestalt)

Network terminology



node or vertex

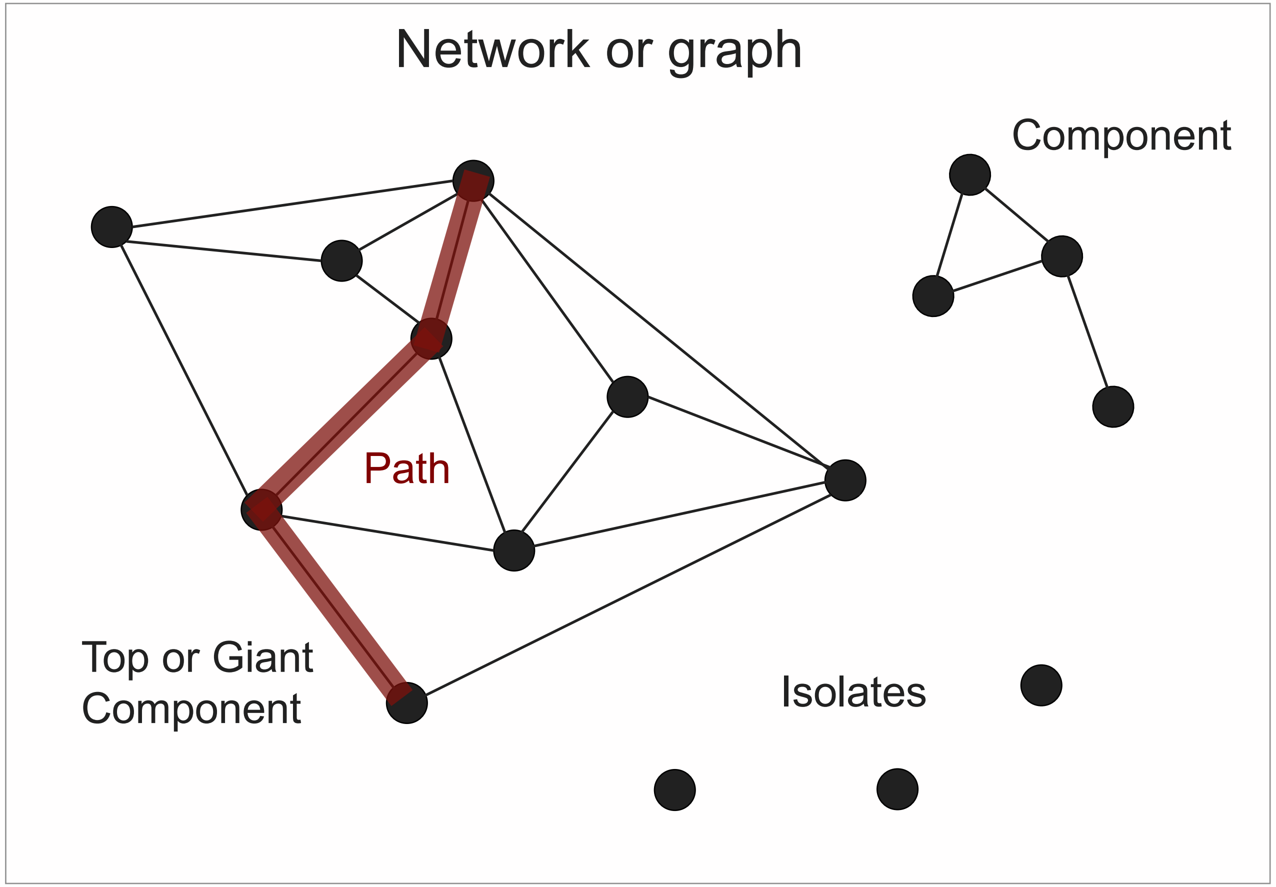
can have attributes, like:
- Gender
- Location
- Start/end time
- etc.

edge or link

can have attributes, like:
- Weight
- Type (directedness)
- Start/end time
- etc.

http://clementlevallois.net/download/networkvisualization.pdf

Network terminology

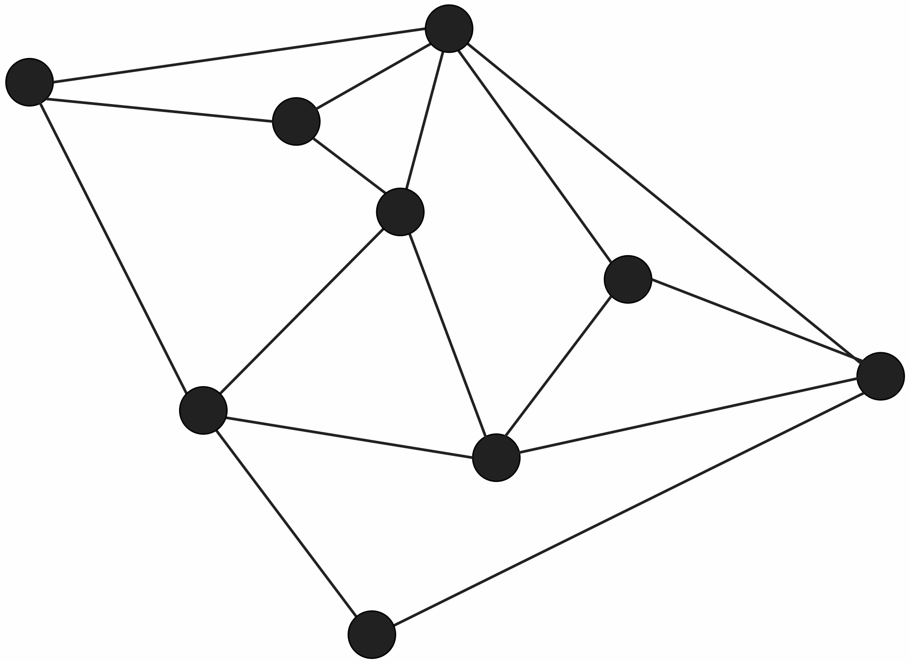


node-link
diagram
(NLD)

https://www.dropbox.com/s/6fngvky6qr0sz8o/Class2_2017.pdf?dl=0

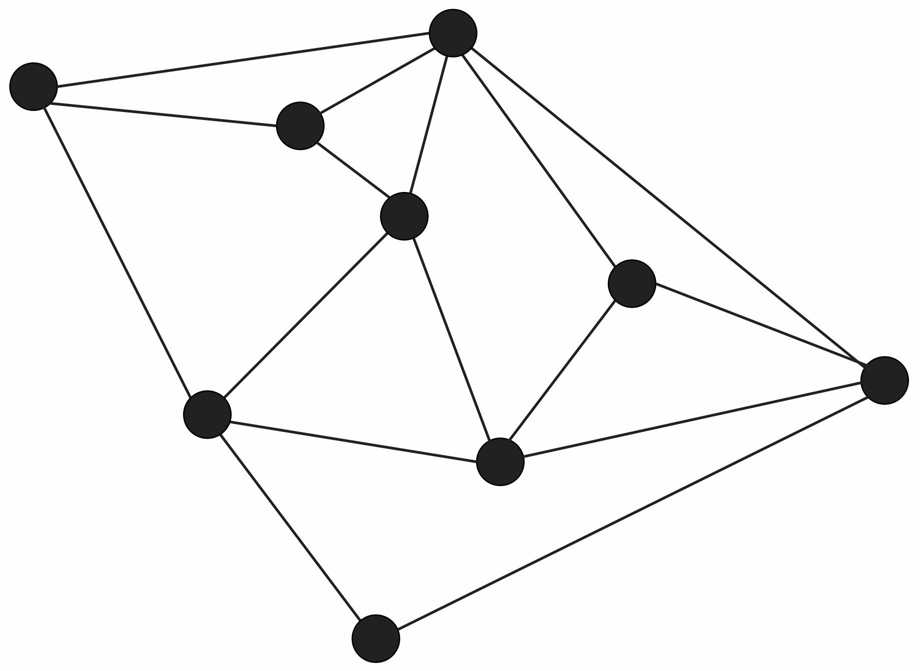
What does it mean to “read” a netvis?

- Understand elements
- Make accurate assessments of data
- Generate insights



What does it mean to “read” a netvis?

- Understand **elements**
- Make **accurate** assessments of data
- Generate insights



Network Visualization Literacy

1. **Tasks**
   What do people need to do with network visualizations?

2. **Design**
   What visualization features might influence usage of netvis?

3. **Training**
   Does training in network science help?

Study 1: Tasks for Network Visualizations

An Opinion Survey of Network Science Researchers

Measurable network properties

Level	Candidate task
Element (node)	1. Closeness Centrality 2. Eigenvector Centrality 3. Node Betweenness Centrality 4. Node Degree
Element (link)	5. Link Betweenness Centrality 6. Loops
Small groups	7. Component Size 8. Modularity 9. Number of Components
Full network	10. Average Degree 11. Average Path Length 12. Average Shortest Path 13. Clustering Coefficient 14. Density 15. Diameter 16. Number of Links 17. Number of Nodes

Measurable network properties

But which are **most important**?

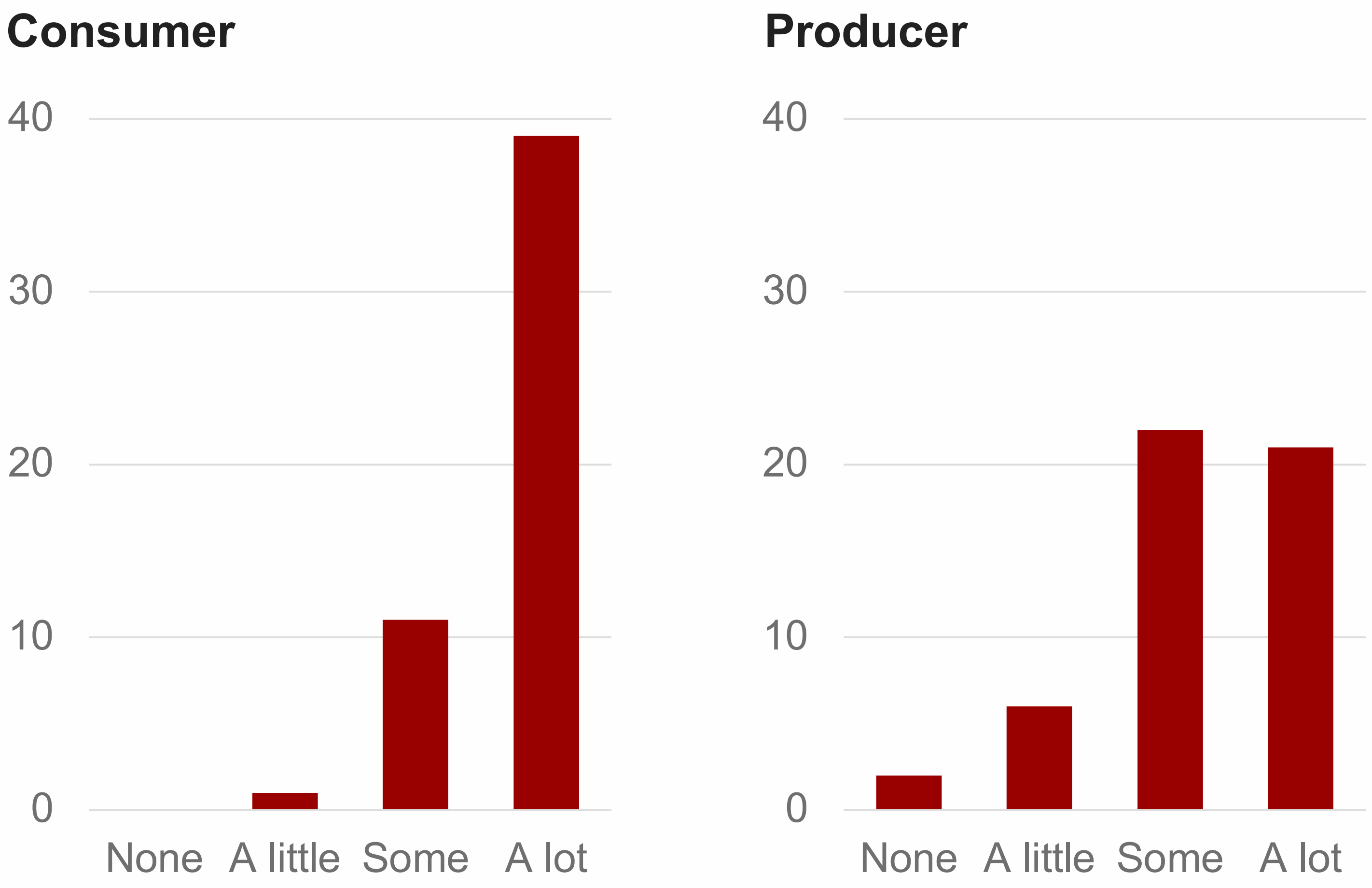
Most likely to be **estimable** from a visualization?

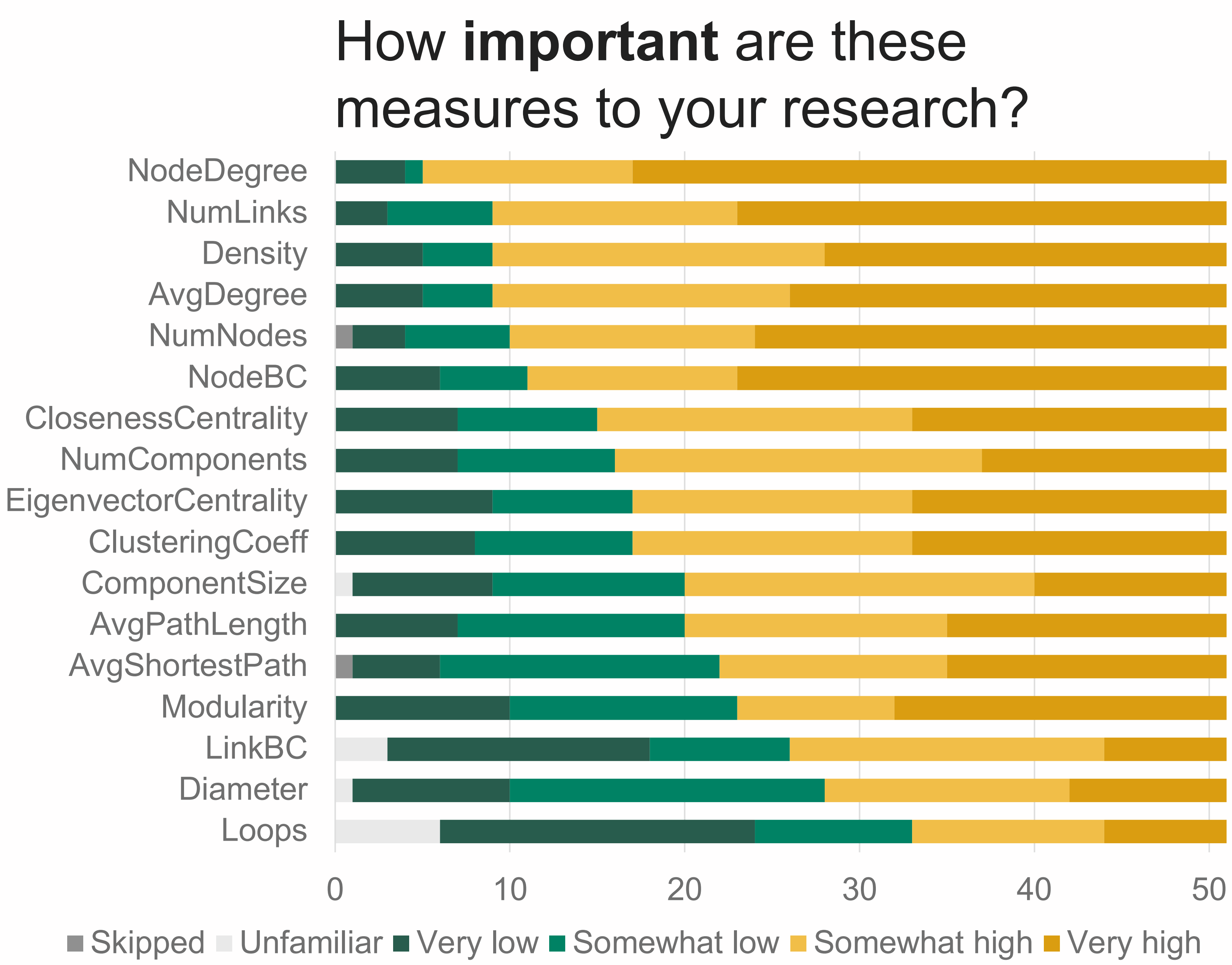
Level	Candidate task
Element (node)	1. Closeness Centrality 2. Eigenvector Centrality 3. Node Betweenness Centrality 4. Node Degree
Element (link)	5. Link Betweenness Centrality 6. Loops
Small groups	7. Component Size 8. Modularity 9. Number of Components
Full network	10. Average Degree 11. Average Path Length 12. Average Shortest Path 13. Clustering Coefficient 14. Density 15. Diameter 16. Number of Links 17. Number of Nodes

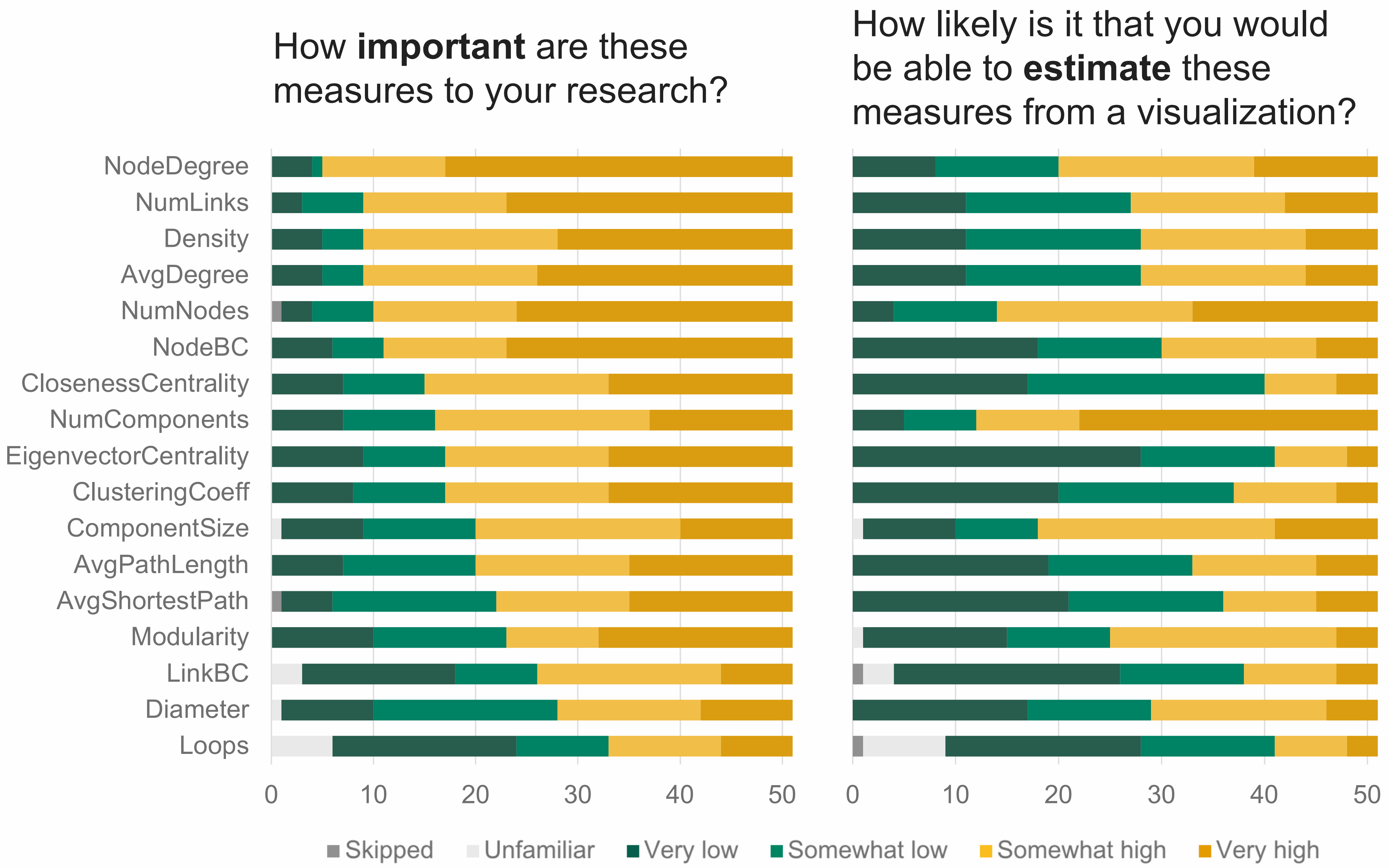
Survey Participants

Open invitation to SOCNET listserv (n=51)

Experience as **consumer** and **producer** of network science research?



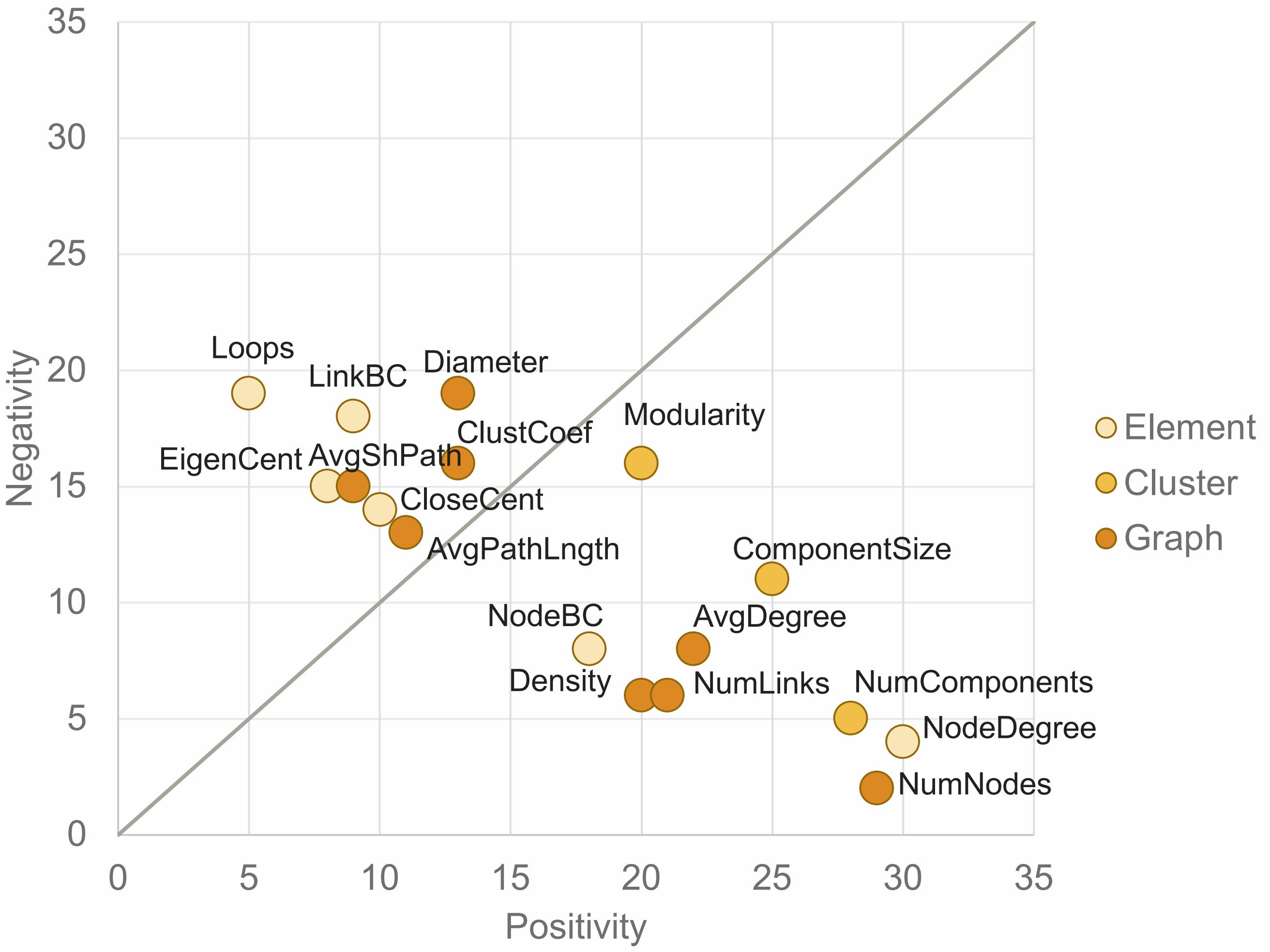

How **important** are these measures to your research?
NodeDegree
NumLinks
Density
AvgDegree
NumNodes
NodeBC
ClosenessCentrality
NumComponents
EigenvectorCentrality
ClusteringCoeff
ComponentSize
AvgPathLength
AvgShortestPath
Modularity
LinkBC
Diameter
Loops
0
10
20
30
40
50
Skipped
Unfamiliar
Very low
Somewhat low
Somewhat high
Very high


How **important** are these measures to your research?
How likely is it that you would be able to **estimate** these measures from a visualization?
NodeDegree
NumLinks
Density
AvgDegree
NumNodes
NodeBC
ClosenessCentrality
NumComponents
EigenvectorCentrality
ClusteringCoeff
ComponentSize
AvgPathLength
AvgShortestPath
Modularity
LinkBC
Diameter
Loops
0
10
20
30
40
50
0
10
20
30
40
50
Skipped
Unfamiliar
Very low
Somewhat low
Somewhat high
Very high

Combining Importance and Estimability

Positivity
number of people rating measure **somewhat or very high** on both importance and estimability

Negativity
number of people rating measure **somewhat or very low** on both importance and estimability



Final tasks

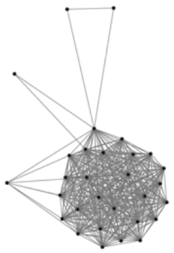
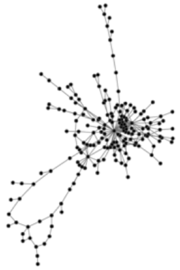
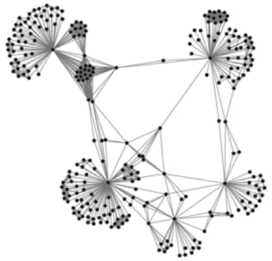
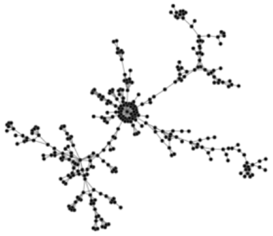
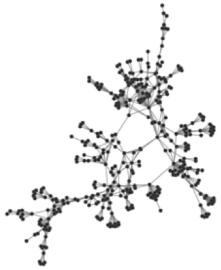
Level	Task
Element (node)	1. Node Degree 2. Node Betweenness Centrality
Small group	3. Number of Components 4. Component Size
Full network	5. Number of Nodes 6. Number of Links 7. Average Degree 8. Density

Study 2a: Impact of Graphic Design, Context

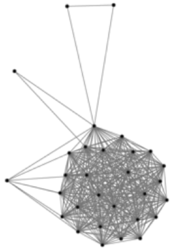
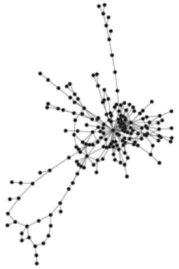
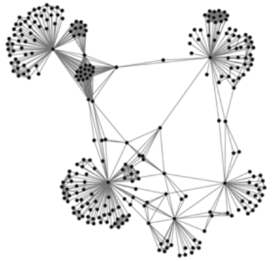
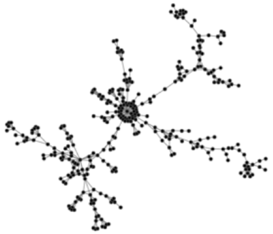
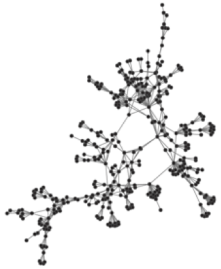
A performance analysis of workers on Amazon's Mechanical Turk

Network datasets

7 real-world datasets

	0	1	3	5	7	8	9
network							
nodes	8	30	67	184	270	321	379
edges	14	337	143	246	932	583	914
density	0.5	0.775	0.065	0.015	0.026	0.011	0.013

A note on density

	0	1	3	5	7	8	9
network							
nodes	8	30	67	184	270	321	379
edges	14	337	143	246	932	583	914
density	**0.5**	**0.775**	**0.065**	**0.015**	**0.026**	**0.011**	**0.013**

Final tasks

Level	Task
Element (node)	1. Node Degree 2. Node Betweenness Centrality
Small group	3. Number of Components 4. Component Size
Full network	5. Number of Nodes 6. Number of Links 7. Average Degree ~~8. Density~~

Final tasks

Numerical Response

- Degree of highest degree node
- Number of clusters (and confidence)
- Number of nodes
- Number of links
- Average degree

Click Response

- Highest degree node
- Highest betweenness centrality node

Percentage Response

- Size of largest cluster

Task phrasing

Degree of Highest Degree Node

Formal

Find the node with the most links.

About how many links does it have?

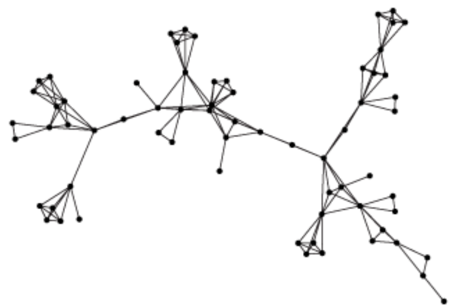
Informal

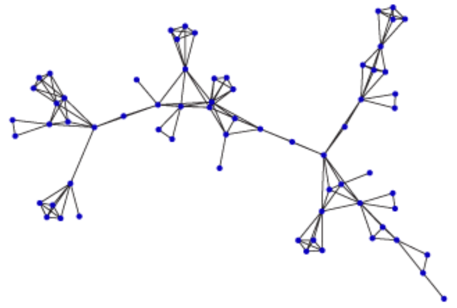
Find the most popular person.

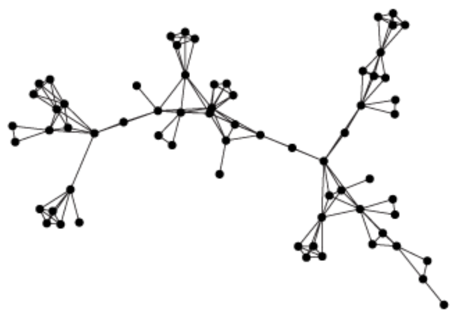
About how many friends does he or she have?

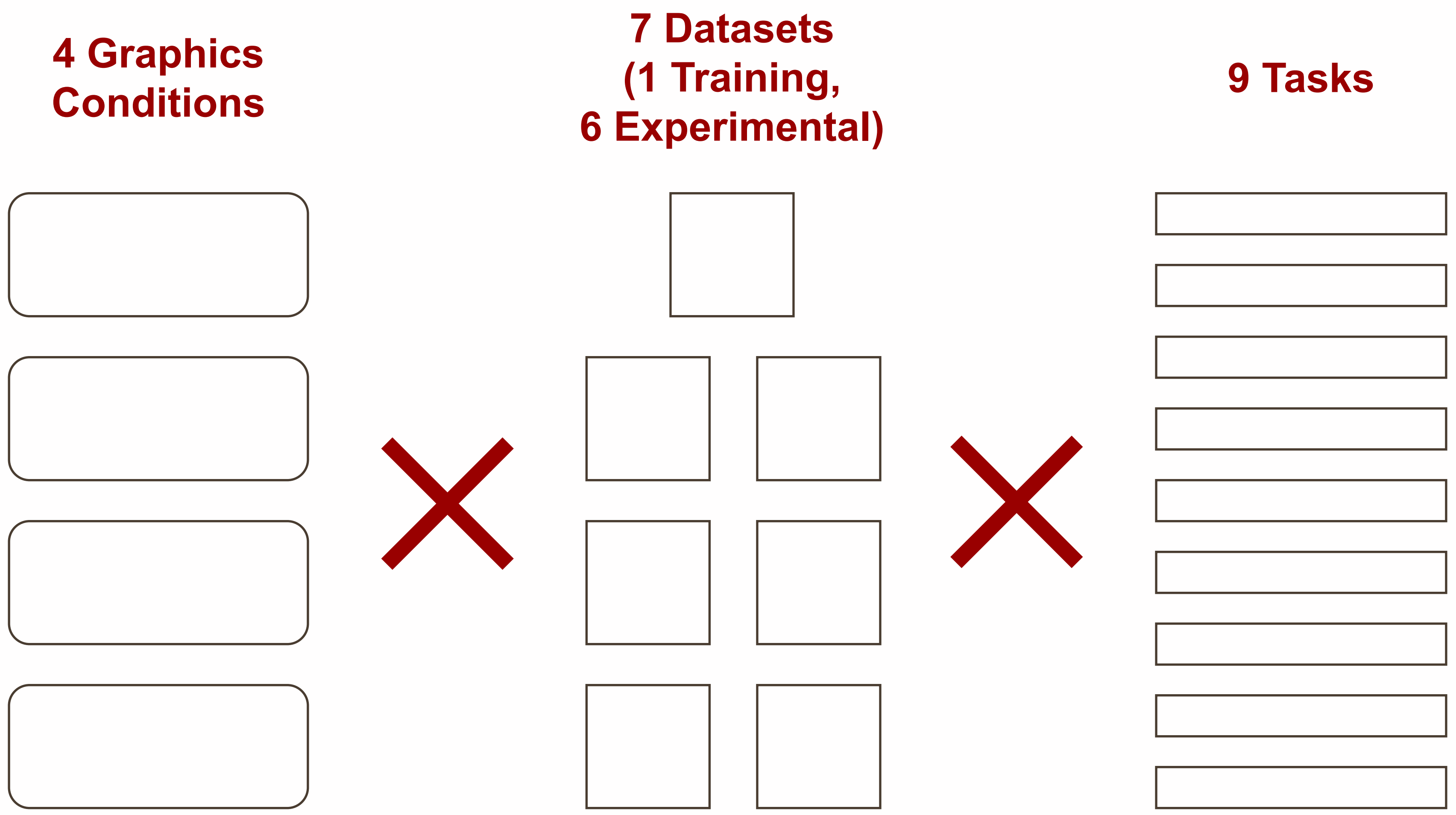
Graphic Conditions (between subjects)

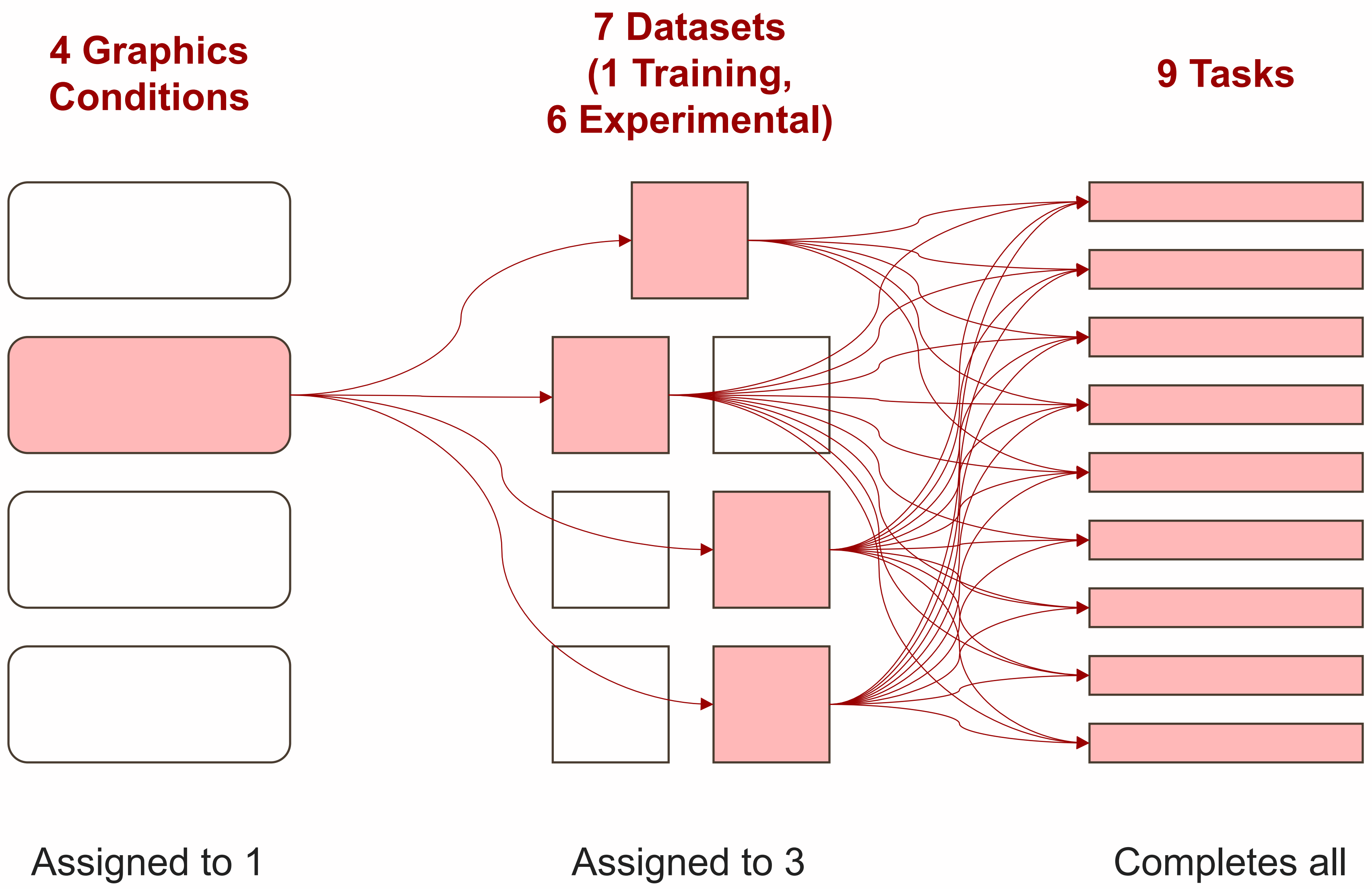
- **Baseline**
  NLD with GEM layout
- **Concrete phrasing**
  Using "person" and "relationship" rather than "node" and "link"
- **Color**
  change all nodes to a different color
- **Size**
  make all nodes slightly larger








4 Graphics Conditions
7 Datasets (1 Training, 6 Experimental)
9 Tasks


4 Graphics Conditions
7 Datasets (1 Training, 6 Experimental)
9 Tasks
Assigned to 1
Assigned to 3
Completes all

Amazon Mechanical Turk

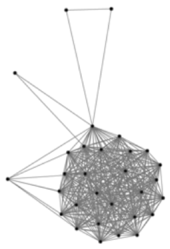
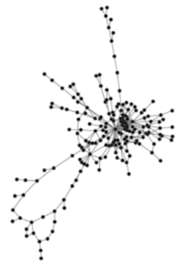
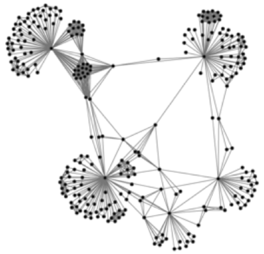
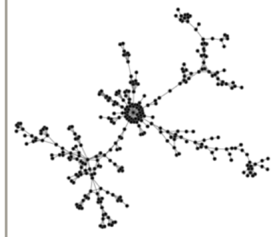
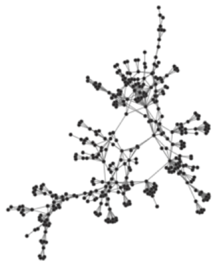
Selection criteria:

- located in the United States
- approval rate for the worker is at least 95%
- number of approved tasks is at least 100

Compensation:

- $3.50 for a 25-30 minute study

AMT Participants

	1	3	5	7	8	9
network						
control	46	49	49	44	50	47
phrasing	54	51	51	52	52	52
color	52	51	52	50	51	54
size	52	55	52	51	49	48

On Measuring Accuracy

On task accuracy

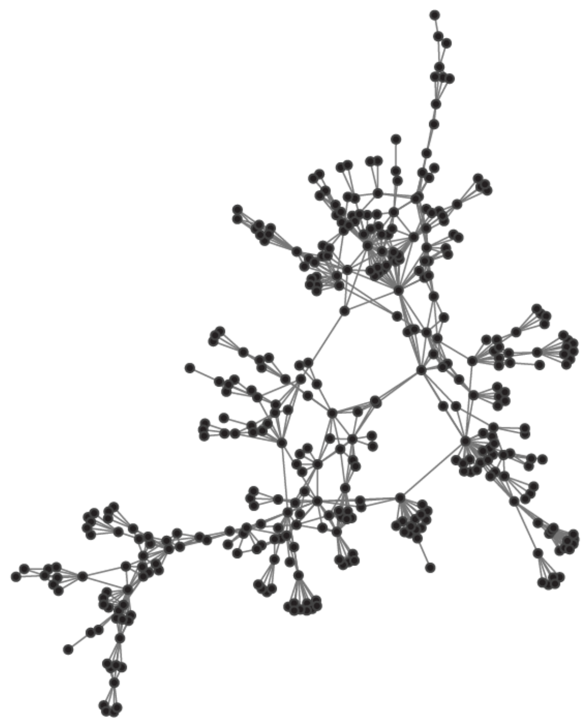
Numerical Responses

Example: Number of Nodes

Correct Answer:	67
Example Response:	50
Absolute Error:	|67-50| = 17
Error Percentage:	17/67 = 25.4%
Log Absolute Error:	log10(|67-50| + 1) = **1.25**

On task accuracy

Numerical Responses

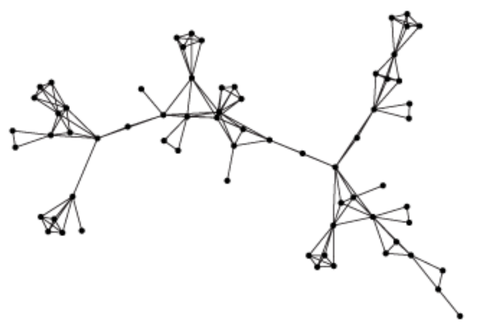


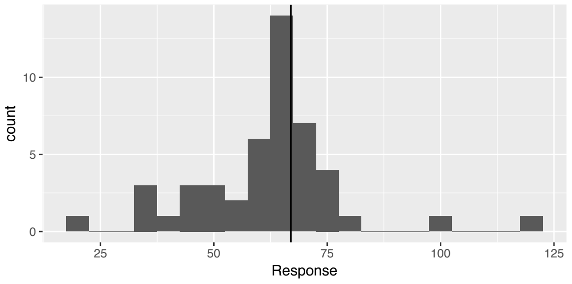
Example: Number of Nodes

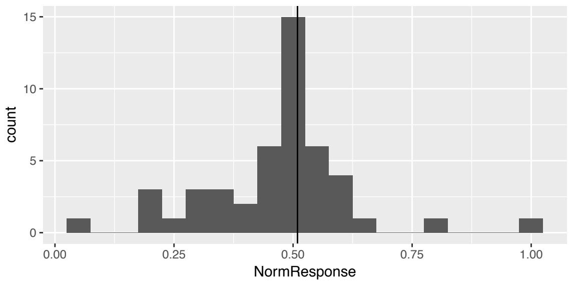
Correct Answer:	379
Example Response:	475
Absolute Error:	|475-379| = 96
Error Percentage:	96/379 = 25.3%
Log Absolute Error:	log10(|475-379| + 1) = **1.97**

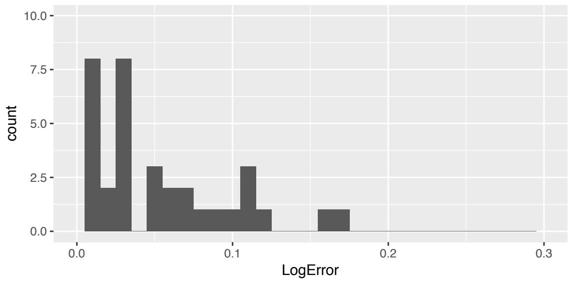
Log Absolute (Normalized) Error

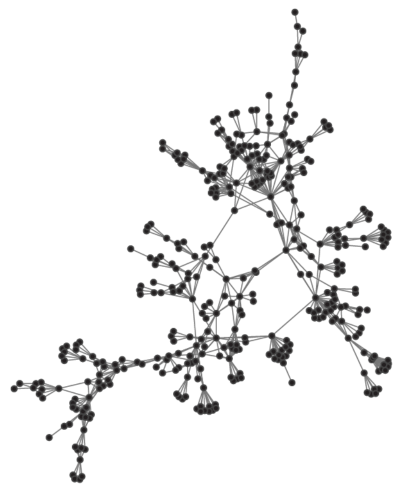
Responses | Normalized Responses | Log Abs Norm Error

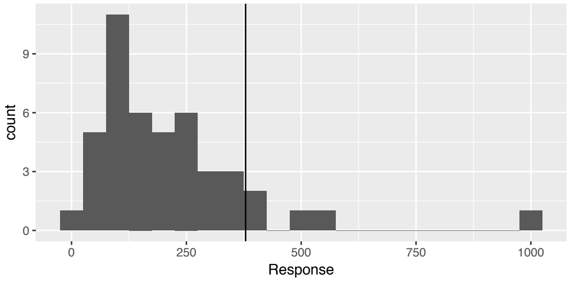


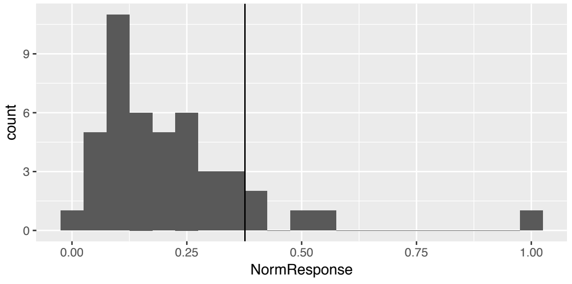


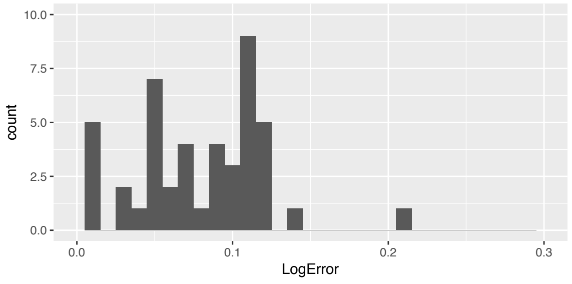








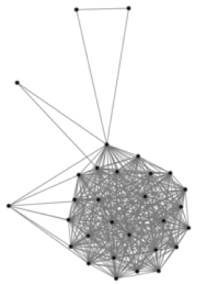
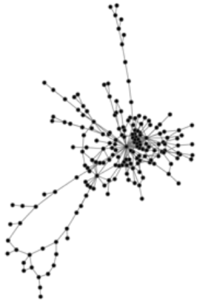
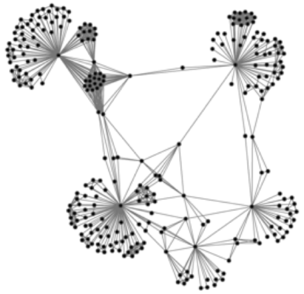
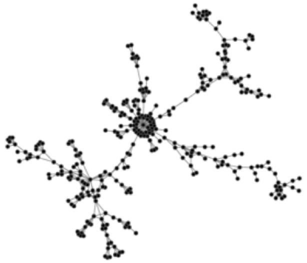
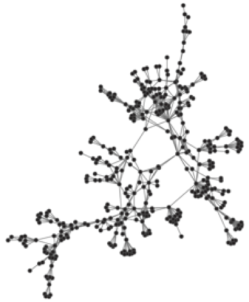




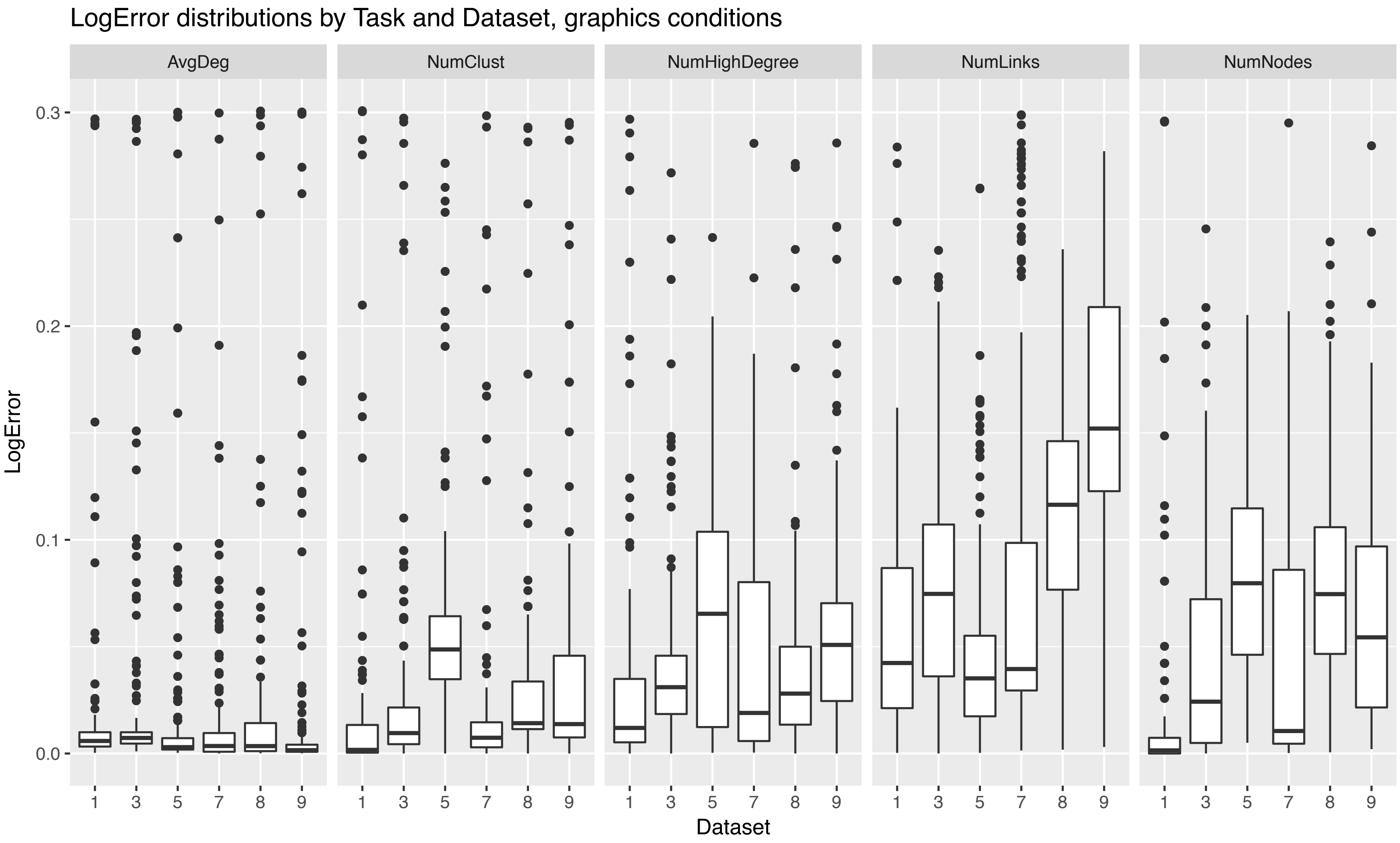
Hypothesis 1:
Larger networks will be harder

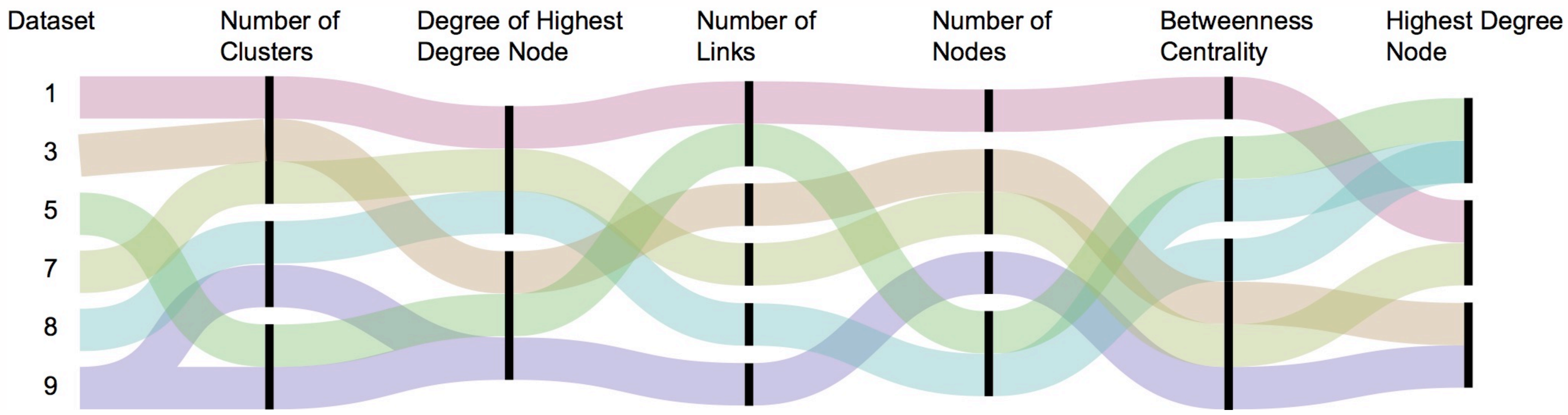
Performance will decrease as networks increase in node size, link size, and density.

Network datasets

	1	3	5	7	8	9
network						
nodes	30	67	184	270	321	379
edges	337	143	246	932	583	914
density	0.775	0.065	0.015	0.026	0.011	0.013

Task difficulty by dataset



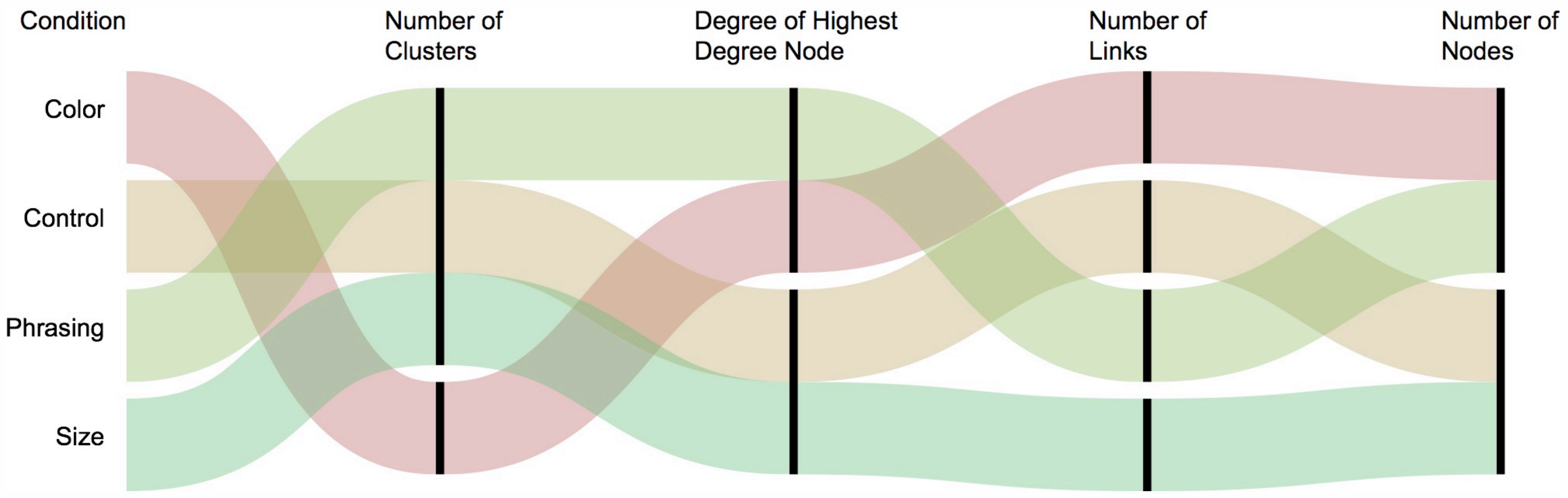


Dataset NS for Average Degree

Hypothesis 2: Informal phrasing will be easier

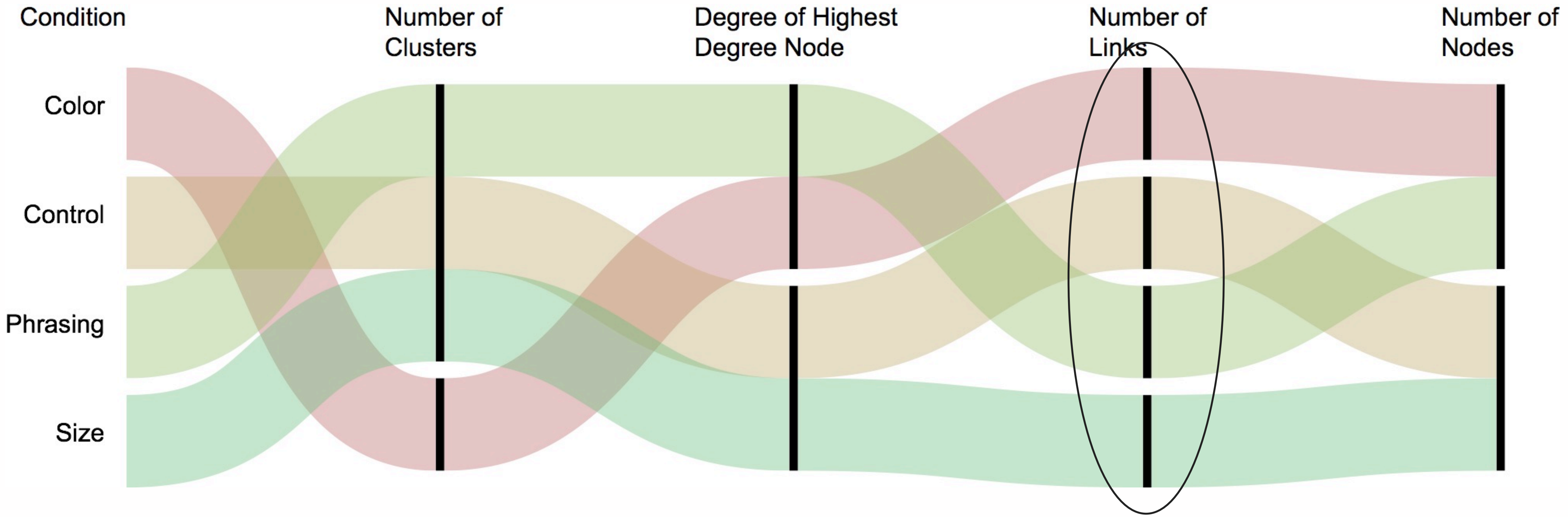
Performance will improve when questions are phrased using informal terminology.

Phrasing does perform best, except for Number of Links task.



Condition NS for Average Degree, Betweenness Centrality, Highest Degree Node

Phrasing does perform best, except for Number of Links task.



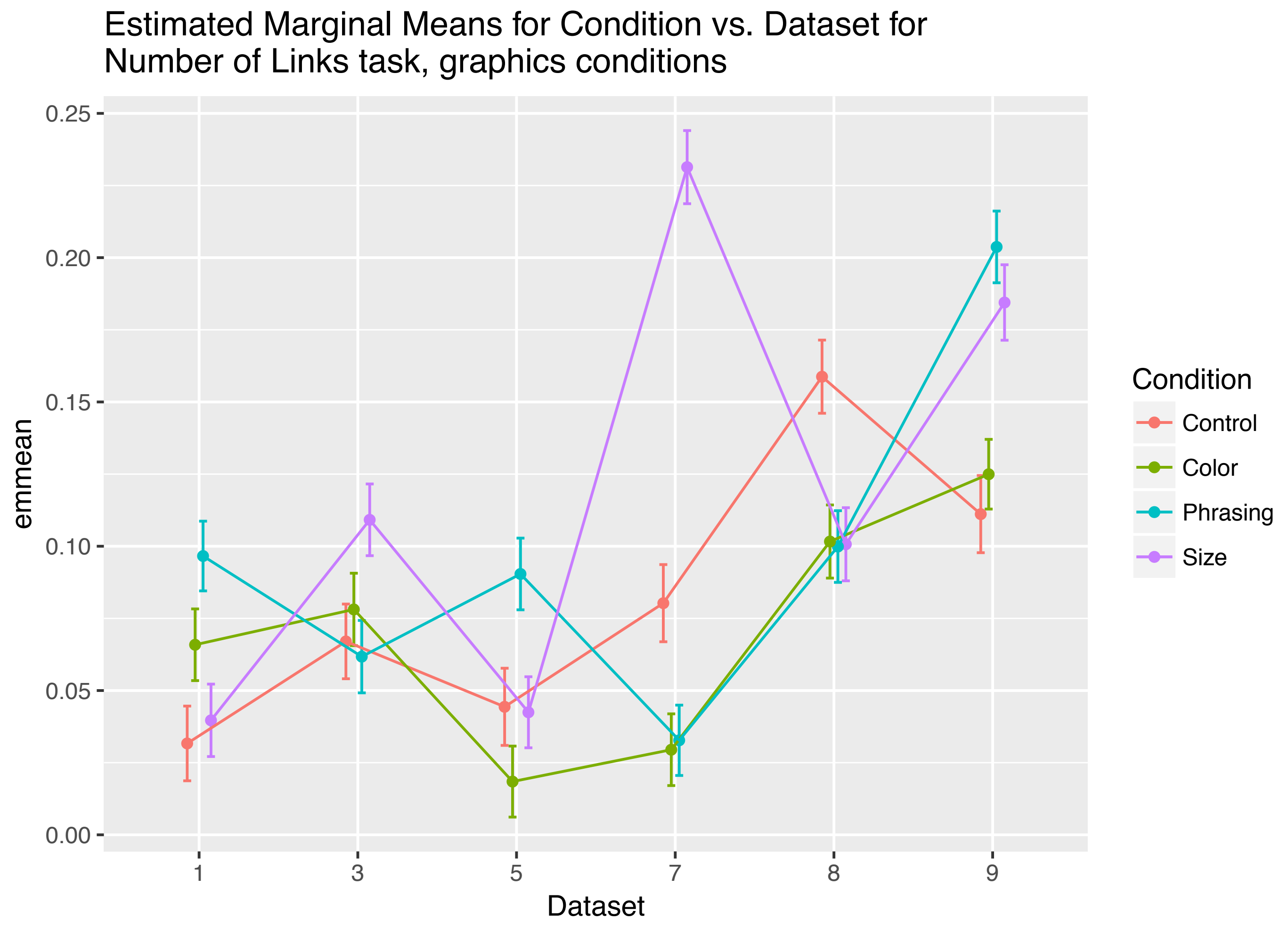
Condition NS for Average Degree, Betweenness Centrality, Highest Degree Node

Number of links, fixed effects

Condition vs. Dataset

1		3		5	
Cond	.group	Cond	.group	Cond	.group
Ctrl	1	**Phr**	1	Col	1
Siz	1	Ctrl	1	Siz	2
Col	2	Col	1	Ctrl	2
Phr	3	Siz	2	**Phr**	3

7		8		9	
Cond	.group	Cond	.group	Cond	.group
Col	1	**Phr**	1	Ctrl	1
Phr	1	Siz	1	Col	1
Ctrl	2	Col	1	Siz	2
Siz	3	Ctrl	2	**Phr**	2

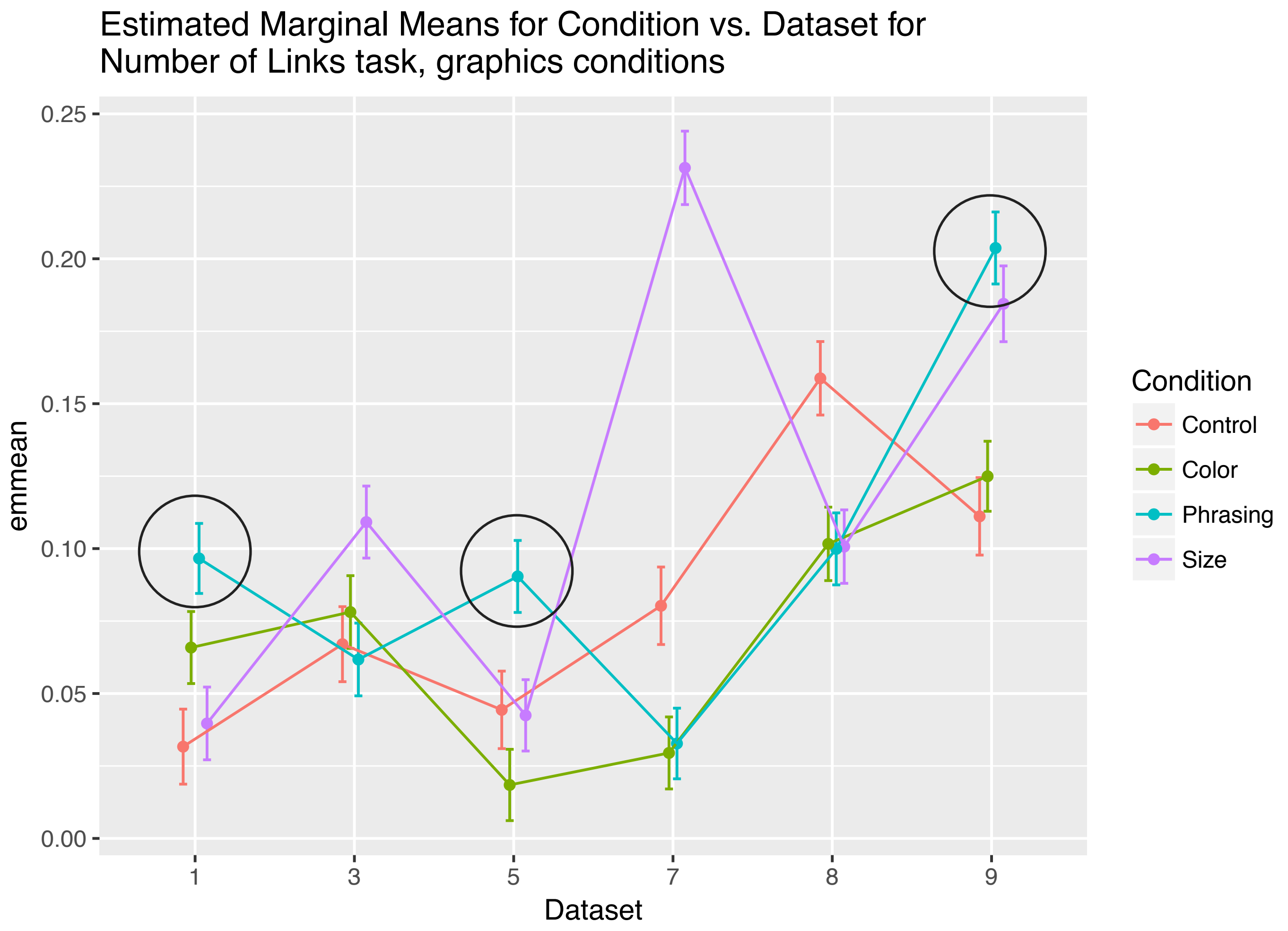


Number of links, fixed effects

Condition vs. Dataset

1		3		5	
Cond	.group	Cond	.group	Cond	.group
Ctrl	1	**Phr**	1	Col	1
Siz	1	Ctrl	1	Siz	2
Col	2	Col	1	Ctrl	2
Phr	3	Siz	2	**Phr**	3

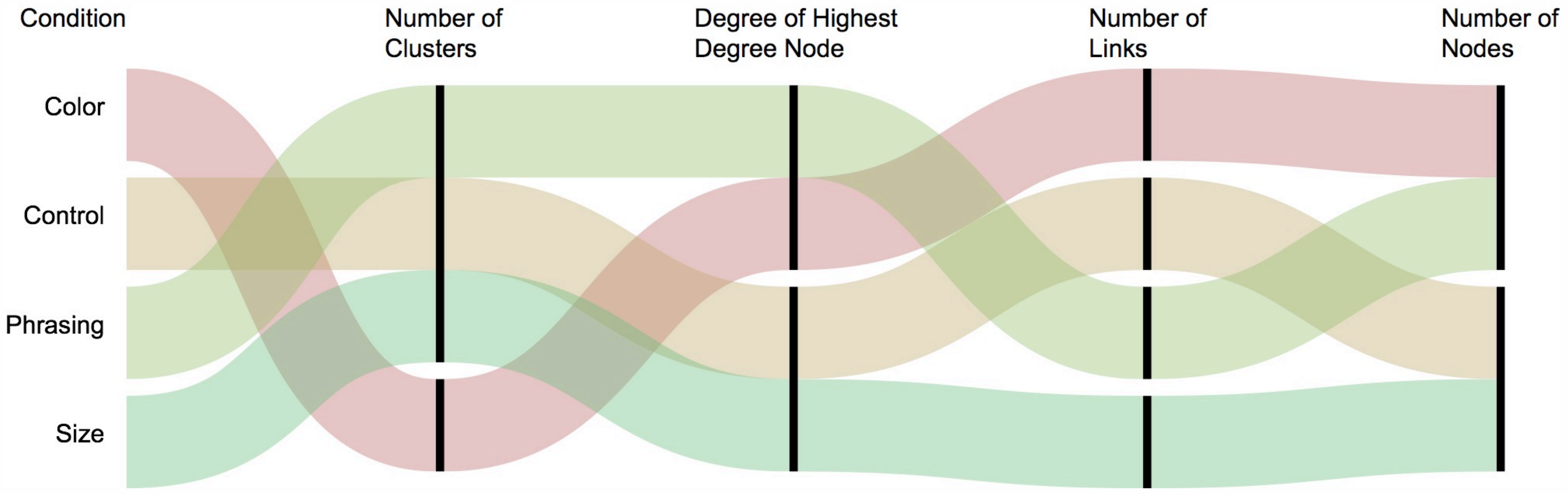
7		8		9	
Cond	.group	Cond	.group	Cond	.group
Col	1	**Phr**	1	Ctrl	1
Phr	1	Siz	1	Col	1
Ctrl	2	Col	1	Siz	2
Siz	3	Ctrl	2	**Phr**	2



Hypothesis 3: Changing node color will make no difference

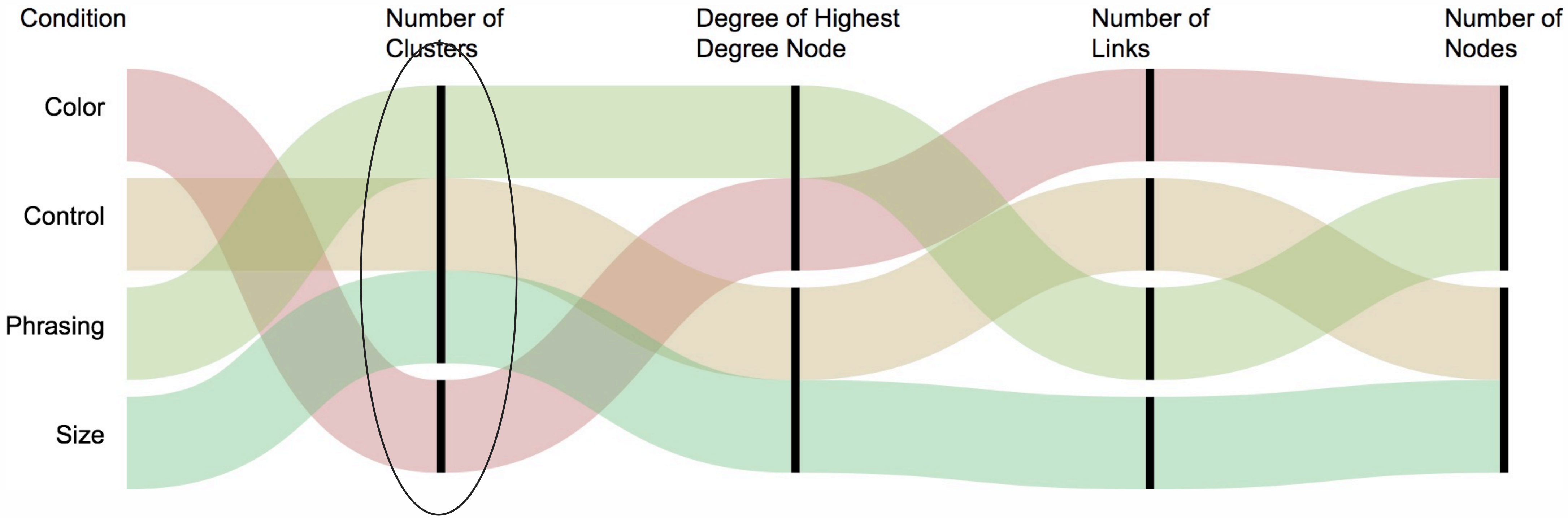
Performance when nodes are changed to an alternate color will be the same as performance on the control condition.

Color almost always better than control.



Condition NS for Average Degree, Betweenness Centrality, Highest Degree Node

Color almost always better than control.



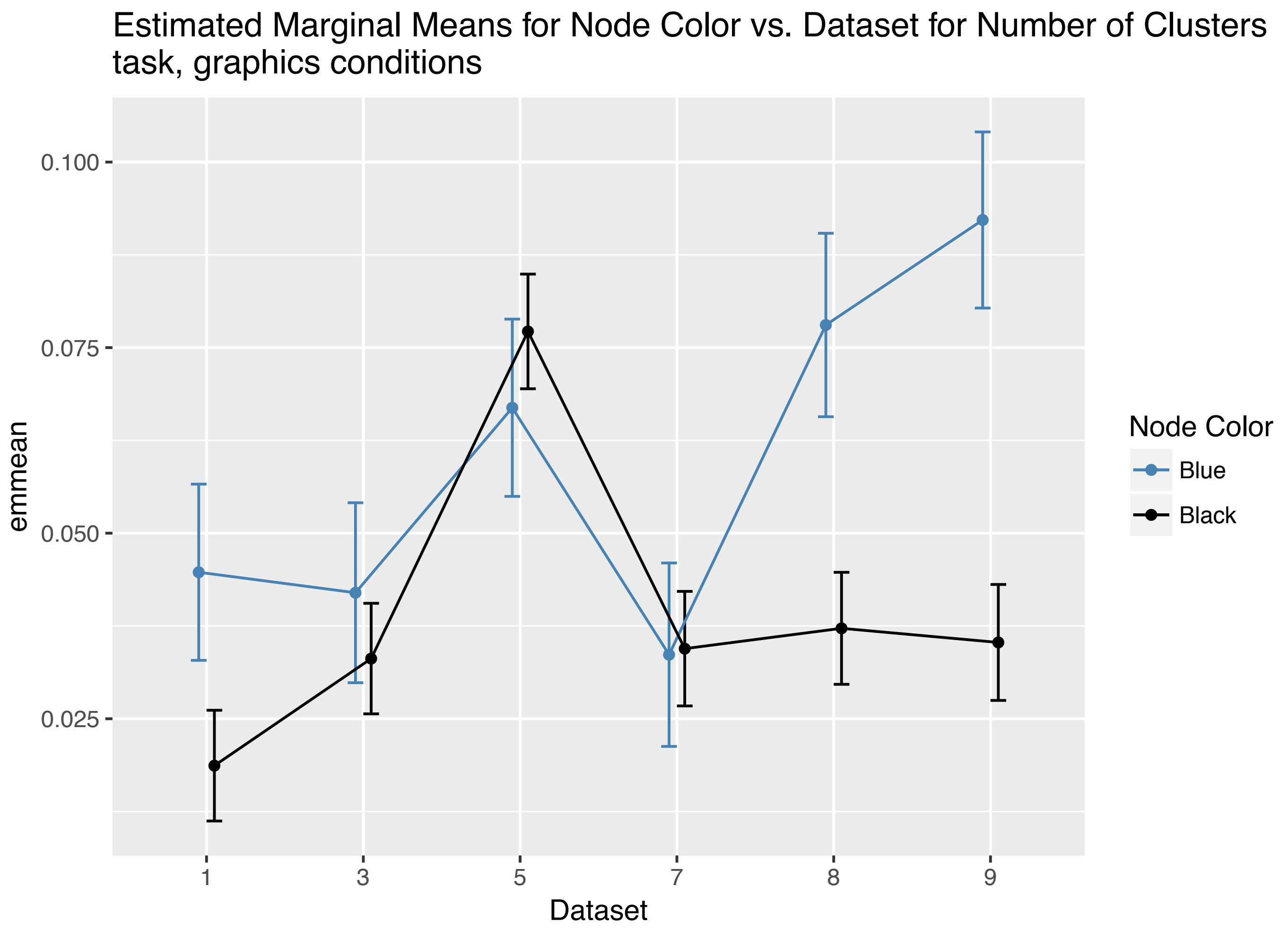
Condition NS for Average Degree, Betweenness Centrality, Highest Degree Node

Number of clusters, fixed effects

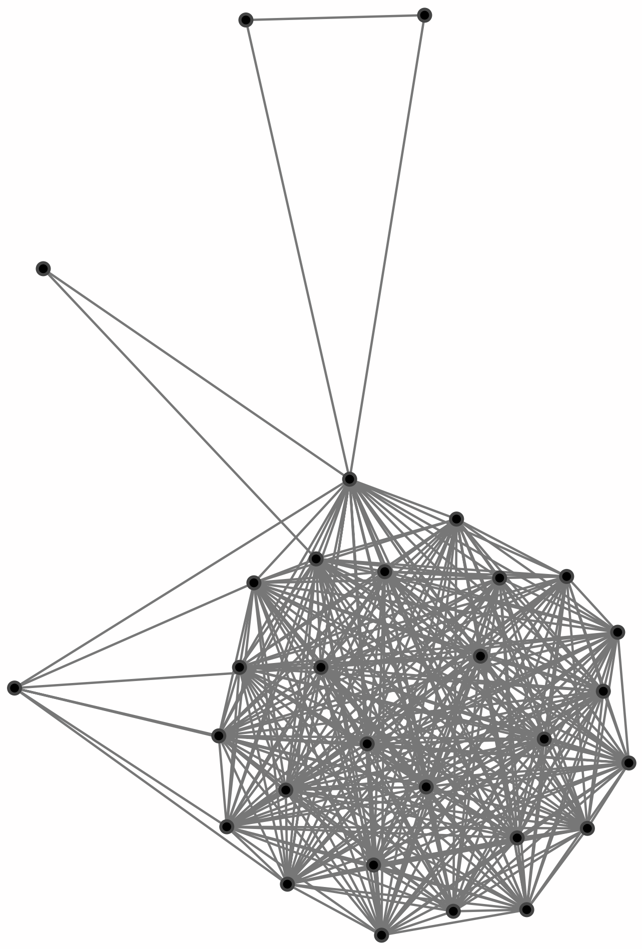
Color of nodes vs. Dataset

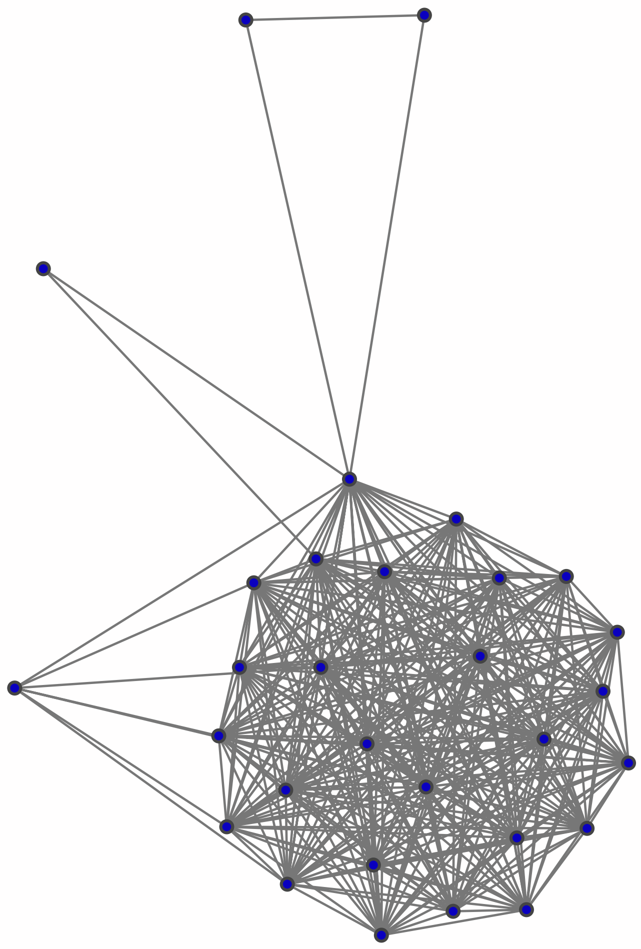
1		3		5	
Cond	.group	Cond	.group	Cond	.group
Black	1	Black	1	**Blue**	1
Blue	2	**Blue**	1	Black	1

7		8		9	
Cond	.group	Cond	.group	Cond	.group
Blue	1	Black	1	Black	1
Black	1	**Blue**	2	**Blue**	2

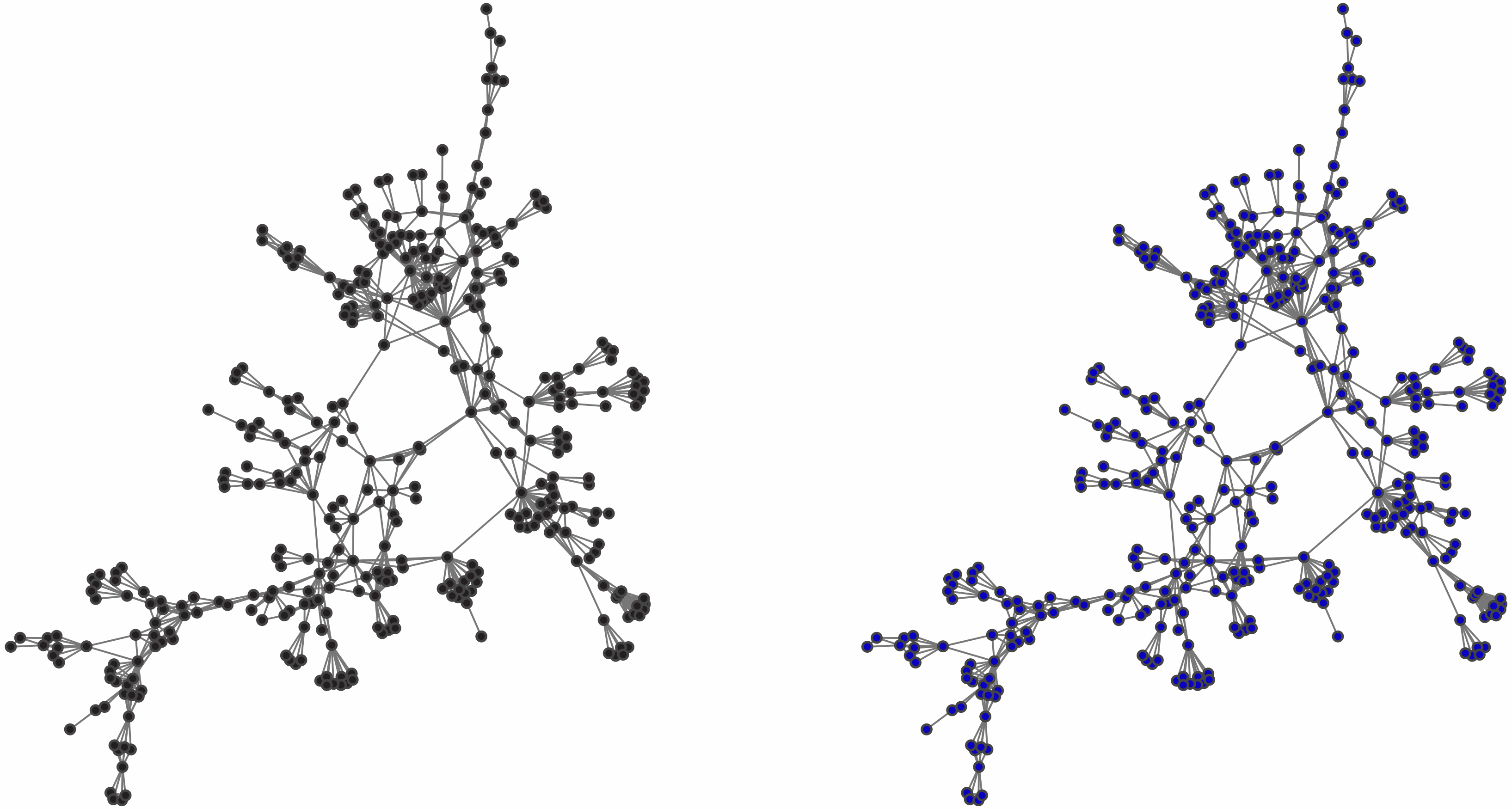


Contrast problem?





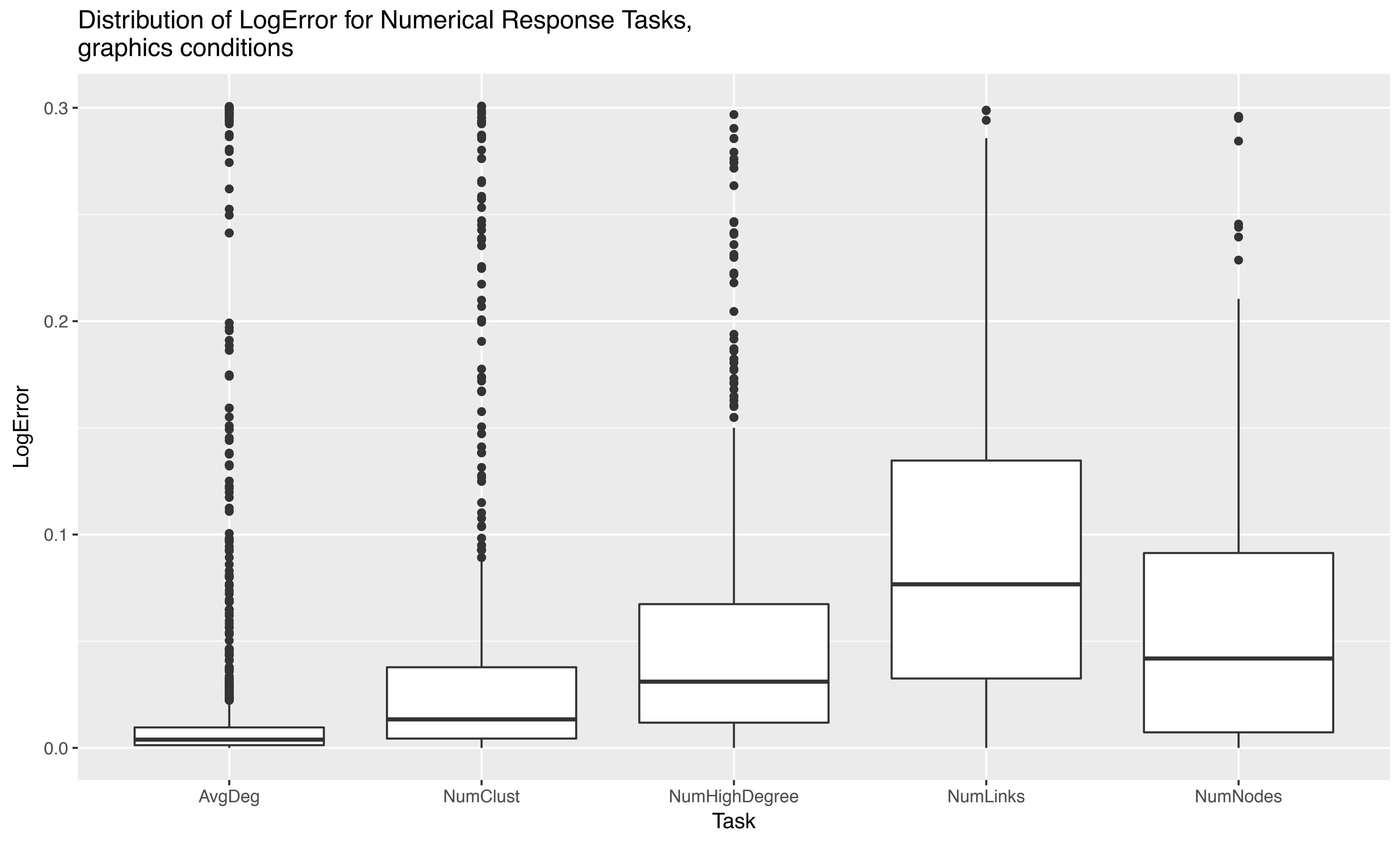
...?



Hypothesis 4:
Tasks involving clusters will be harder

Performance will decline on the number of clusters and size of clusters tasks.

Task difficulty

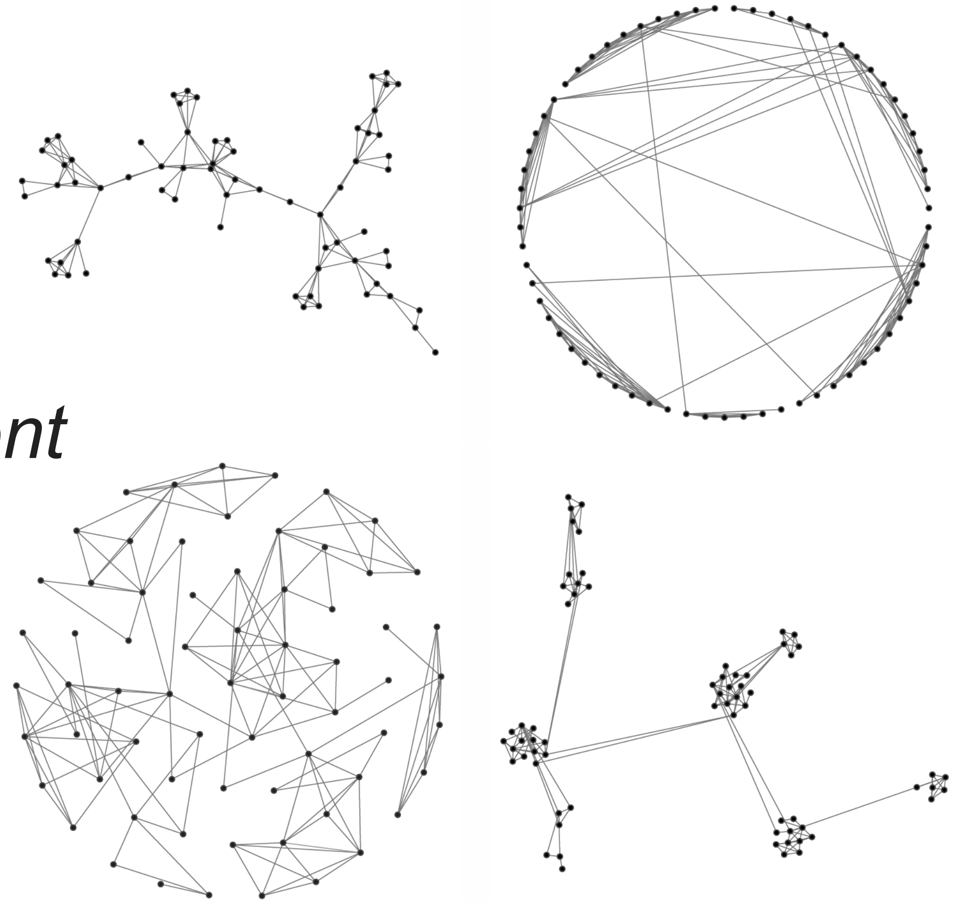


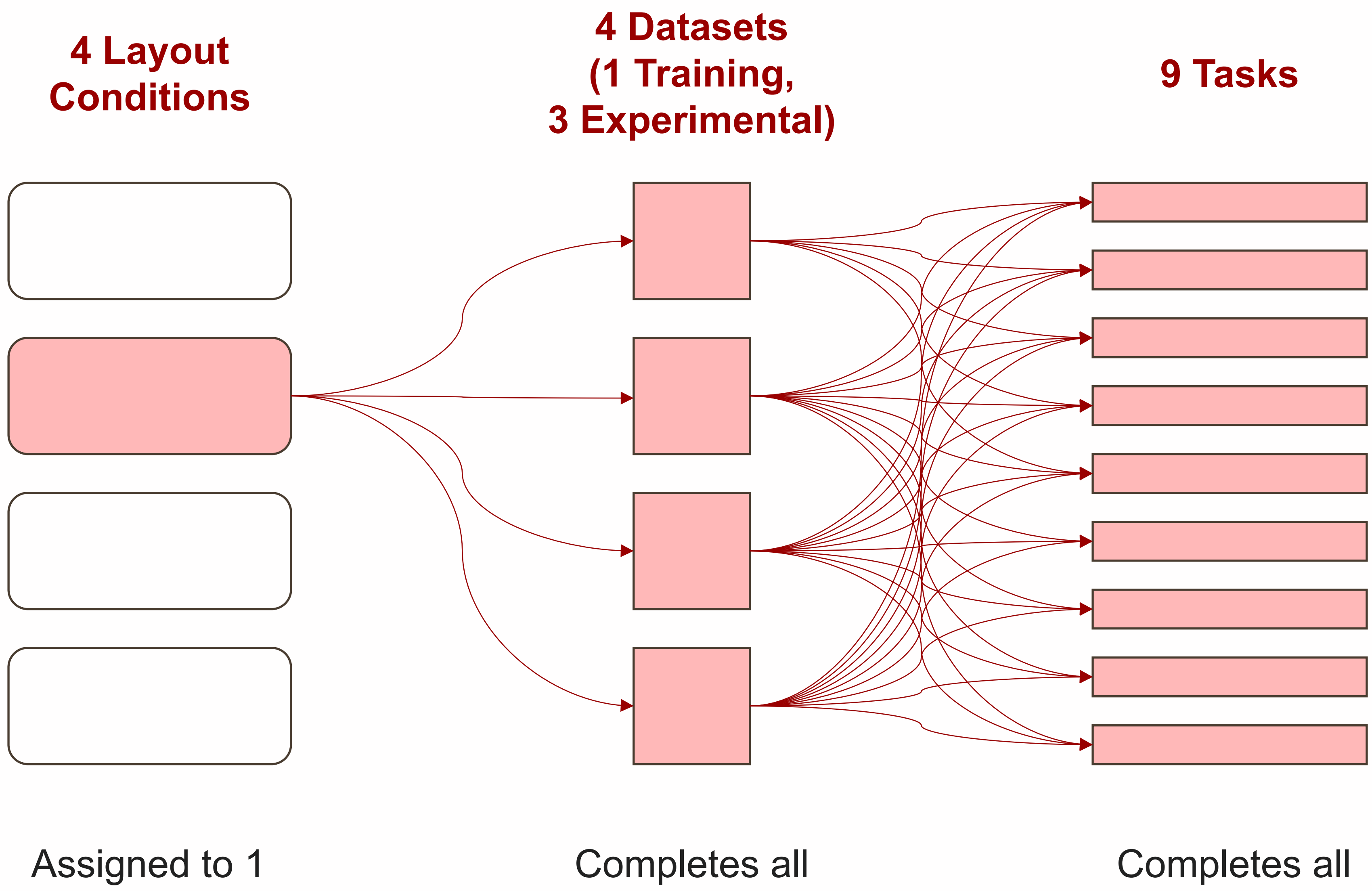
Study 2b: Impact of Layout, Expertise

A comparison between Turkers and IUNI Network Science researchers

Layout Conditions (between subjects)

- **GEM layout**
  force-directed layout
- **Circular layout**
  nodes positioned by cluster assignment
- **Fruchterman-Reingold**
  nodes evenly distributed
- **OpenOrd**
  emphasizes clusters




4 Layout Conditions
4 Datasets (1 Training, 3 Experimental)
9 Tasks
Assigned to 1
Completes all
Completes all

IU Network Science community

Selection criteria:

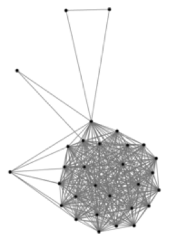
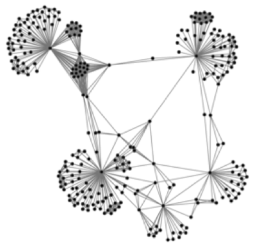
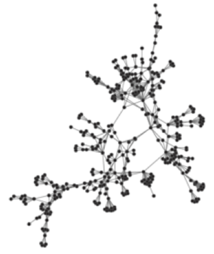
- Affiliated with IUNI, CNS program, or other network science training

Compensation:

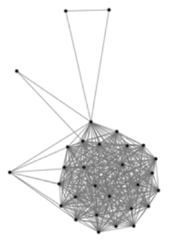
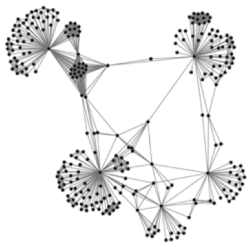
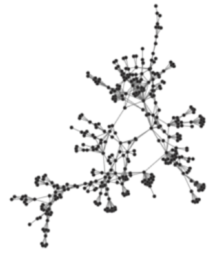
- Pilot: drawing for two $50 Amazon Gift Cards
- Graduate Students: $10 Amazon Gift Cards, pizza
- Faculty/Staff: randomly assigned to two conditions: $10 Amazon Gift Card, $10 donation to IU diversity scholarship

Participants

IU NetSci

network			
control	23	23	22
circle	19	17	19
frucht	17	17	16
openord	22	23	23

MTurk

network			
control	46	44	47
circle	46	44	44
frucht	49	51	52
openord	49	48	49

H5: Network experts will perform better.

Network science expertise will improve performance, even across layout changes.

~~H5: Network experts will perform better.~~

Network science expertise will improve performance, even across layout changes.

H5 not supported.

None of the task performance models retained **network science training** as a significant predictor of accuracy, even when collapsing across condition increases the sample size.

H5 not supported.

None of the task performance models retained **network science training** as a significant predictor of accuracy, even when collapsing across condition increases the sample size.

☺ Users don’t need much training to perform as well as experts!

H5 not supported.

None of the task performance models retained **network science training** as a significant predictor of accuracy, even when collapsing across condition increases the sample size.

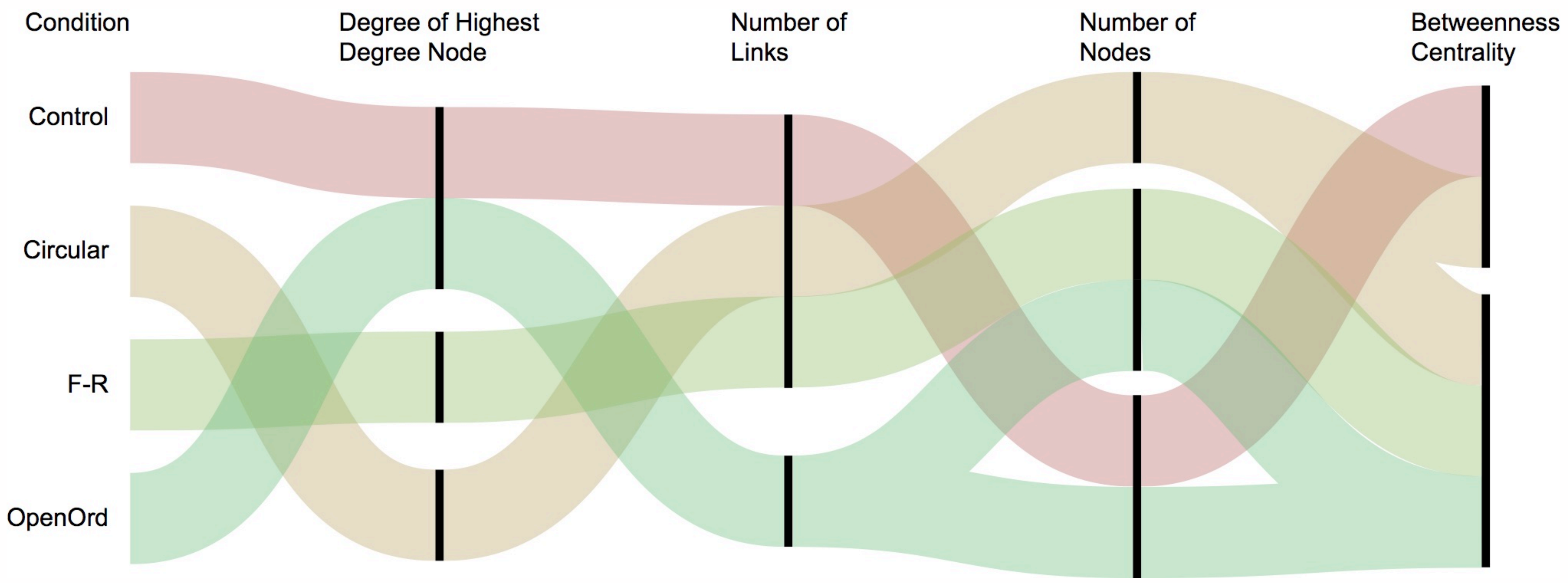
☺ Users don’t need much training to perform as well as experts!

☹ Not even experts get especially high accuracy scores.

H6: Different layouts will improve performance on different tasks.

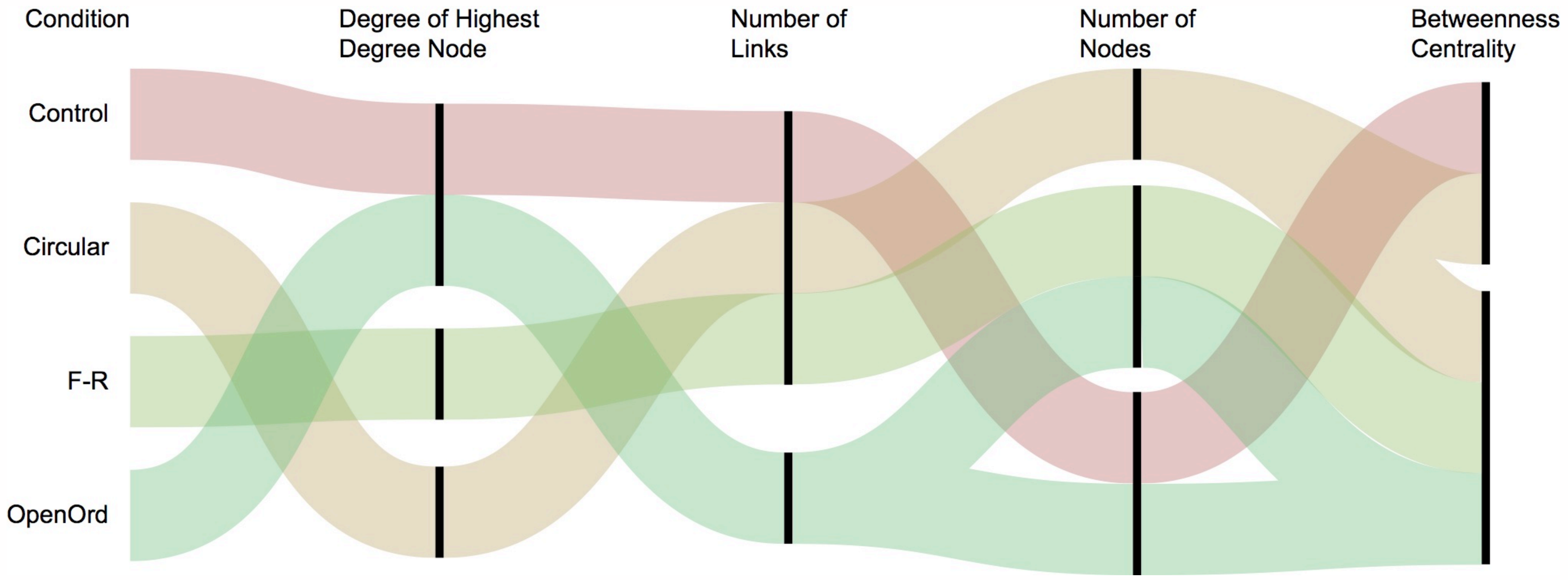
a) OpenOrd will improve performance on clustering tasks.
b) Fruchterman-Reingold will improve performance on node tasks.
c) Circular layout will impair performance on all tasks.

Layout vs. Task



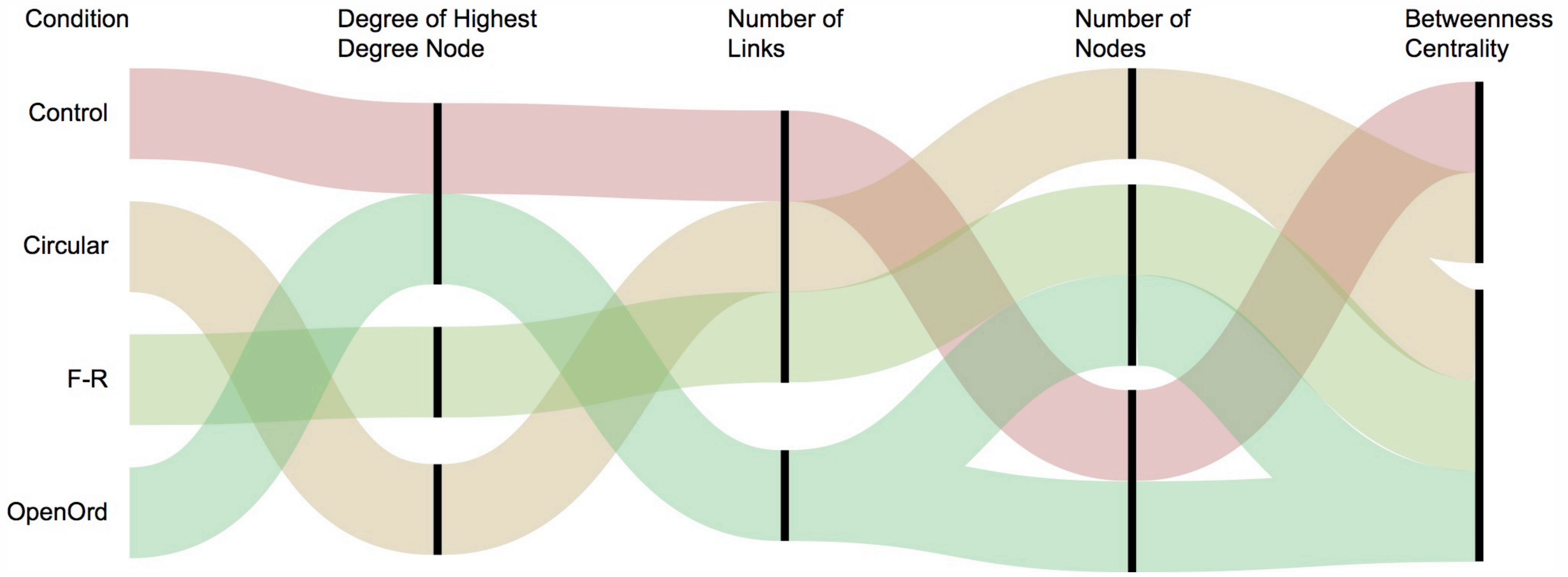
Condition NS for Average Degree, Number of Clusters. Result unknown for Highest Degree Node.

~~OpenOrd~~: Condition NS for cluster tasks



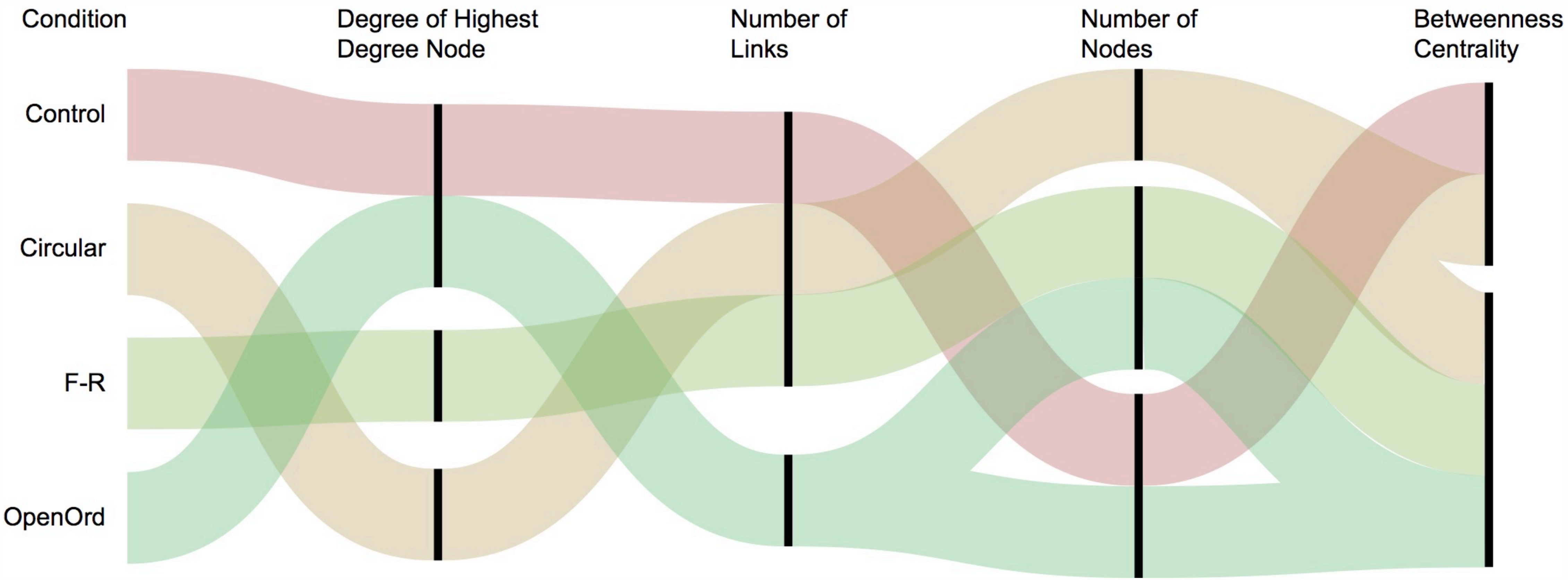
Condition NS for Average Degree, Number of Clusters. Result unknown for Highest Degree Node.

~~FR~~: FR not especially good for node tasks



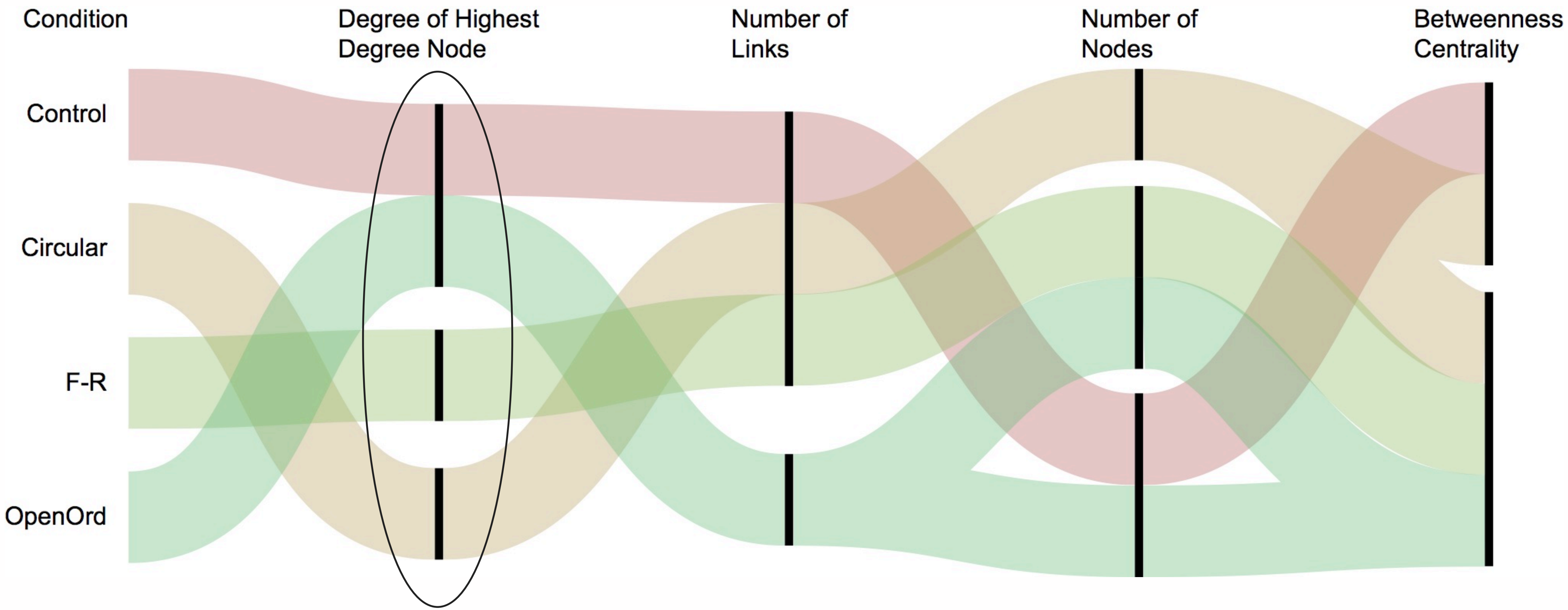
Condition NS for Average Degree, Number of Clusters. Result unknown for Highest Degree Node.

~~Circular~~: Circular is generally fine!



Condition NS for Average Degree, Number of Clusters. Result unknown for Highest Degree Node.

~~Circular~~: Circular is generally fine!

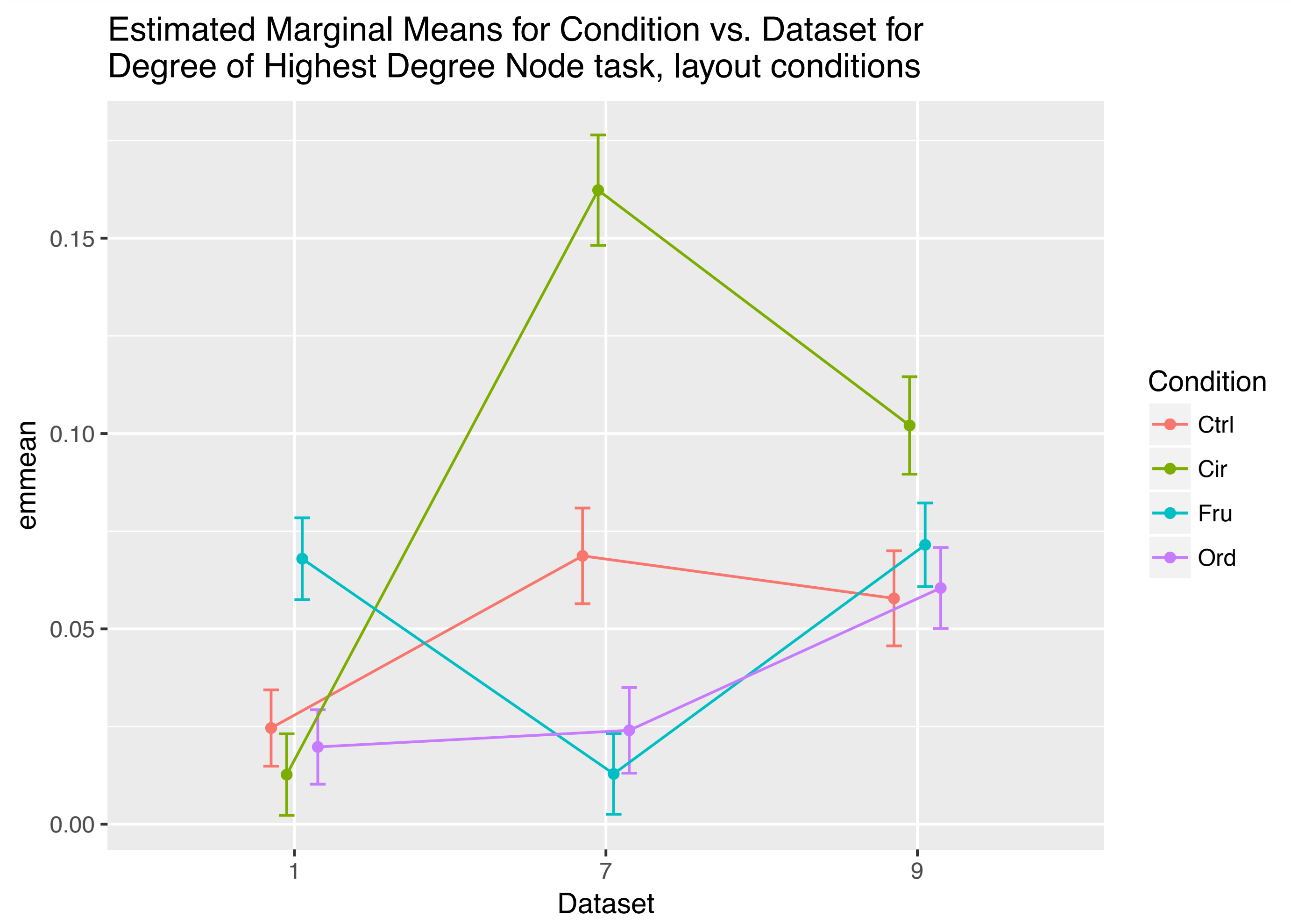


Condition NS for Average Degree, Number of Clusters. Result unknown for Highest Degree Node.

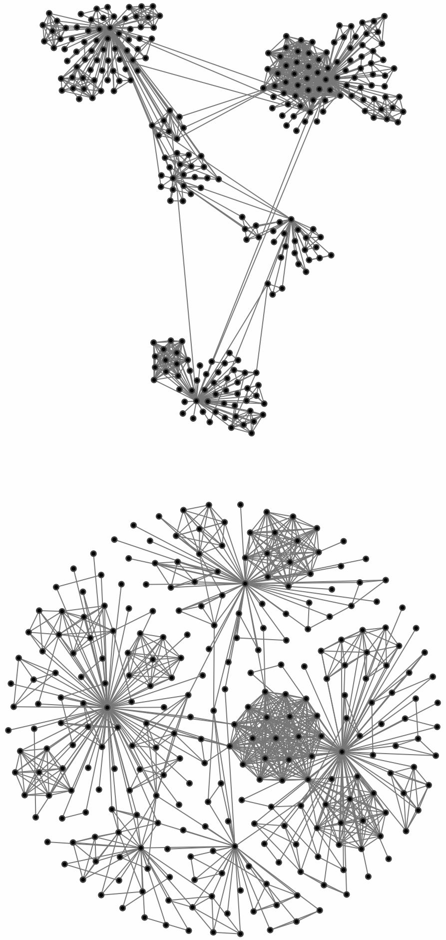
Degree of Highest Degree Node, fixed effects

Condition vs. Dataset

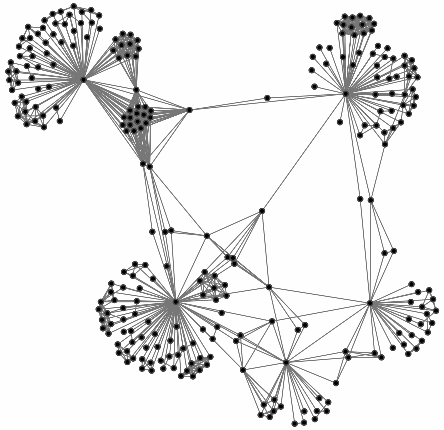
1		7		9	
Cond	.group	Cond	.group	Cond	.group
Cir	1	Fru	1	Ctrl	1
Ord	1	Ord	1	Ord	1
Ctrl	1	Ctrl	2	Fru	1
Fru	2	**Cir**	3	**Cir**	2



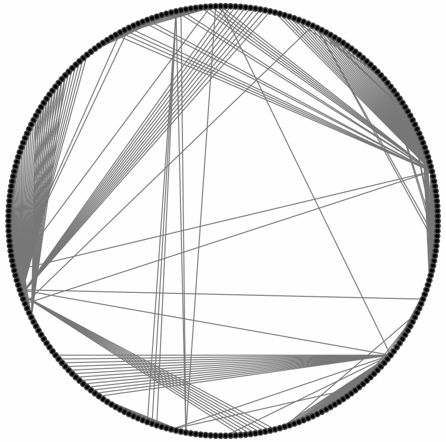
Dataset 7 especially problematic



>



>



Conclusions

Recommendations

- Including brief training should allow novices to read network visualizations well (enough)
- A brighter node color may indeed improve performance
- Layout algorithm may influence task performance, but the relationship is complex.
  - Force-directed is fine overall
  - Circular is fine for tasks that don't require estimating degree

Challenges

- Selecting and operationalizing task
- Operationalizing error
- Recruiting experts

Possible extensions

- Different tasks
- Different populations
- Data overlays
- Interactivity
- Qualitative data on interpretation strategies

References

- Arnheim, R. (1969). *Visual thinking.* Berkeley: University of California Press.
- Du, F., Plaisant, C., Spring, N., & Shneiderman, B. (2017). Finding similar people to guide life choices: Challenge, design, and evaluation. In S. Fussell & G. Mark (Eds.), *CHI '17: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 5498-5544). New York, NY: ACM.
- Healey, C. G., & Enns, J. T. (2012). Attention and Visual Memory in Visualization and Computer Graphics. *IEEE Transactions on Visualization and Computer Graphics, 18*(7), 1170-1188.
- Sprague, D., & Tory, M. (2012). Exploring how and why people use visualizations in casual contexts: Modeling user goals and regulated motivations. *Information Visualization, 11*(2), 106-123.
- Ware, C. (2012). *Information Visualization: Perception for Design*. (3rd Edition). Morgan Kaufman.
- Zoss, A. M. (2014). Cognitive processes and traits related to graphic comprehension. In M. L. Huang, W. Huang (Eds.), *Innovative Approaches of Data Visualization and Visual Analytics*, pp. 94-110, IGI Global. DOI:10.4018/978-1-4666-4309-3.ch005

Questions?

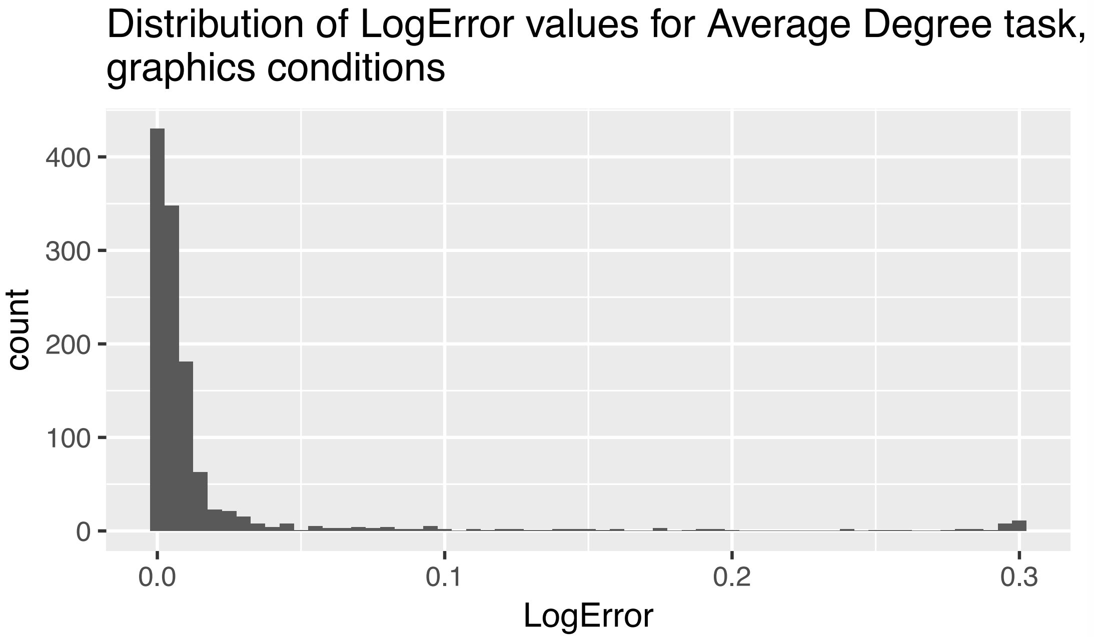
angela.zoss@duke.edu

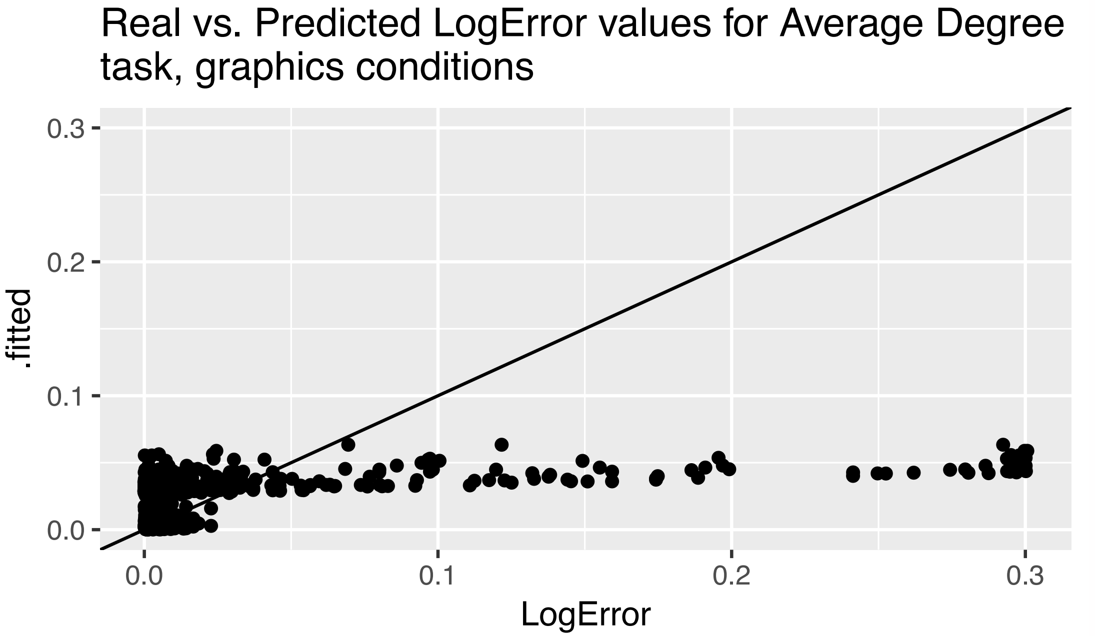
netvislit.org

Graphics Results by Task: Numerical Responses

Average Degree

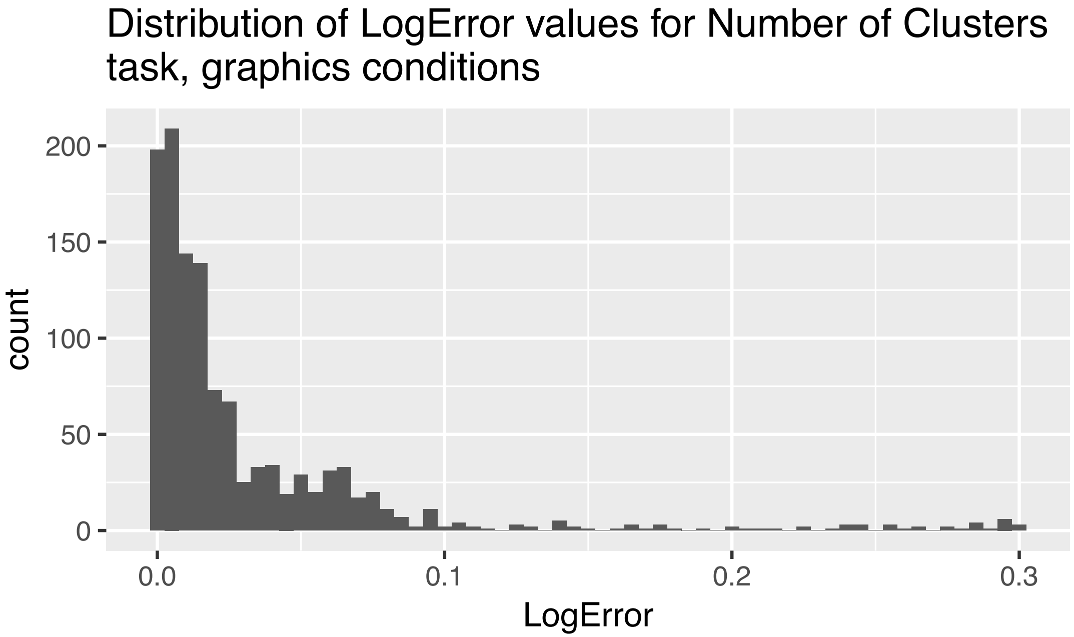
Best-fitting model has limited explanatory value (R^2 = 0.084)



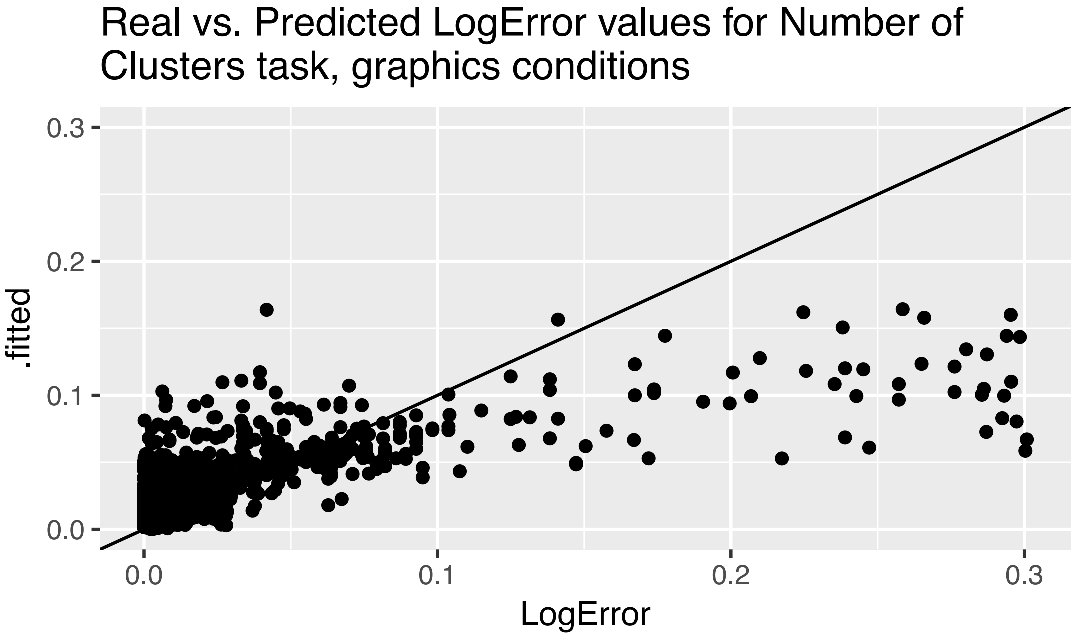


Number of Clusters

Best-fitting model has limited explanatory value ($R^2 = 0.278$)



LogError~ ConditionColor +
Dataset +
Overestimated +
Stats.OperatingSystemNumClust +
ConditionColor:Dataset +
Overestimated:OperatingSystem +
(1|Demo.ResponseID)

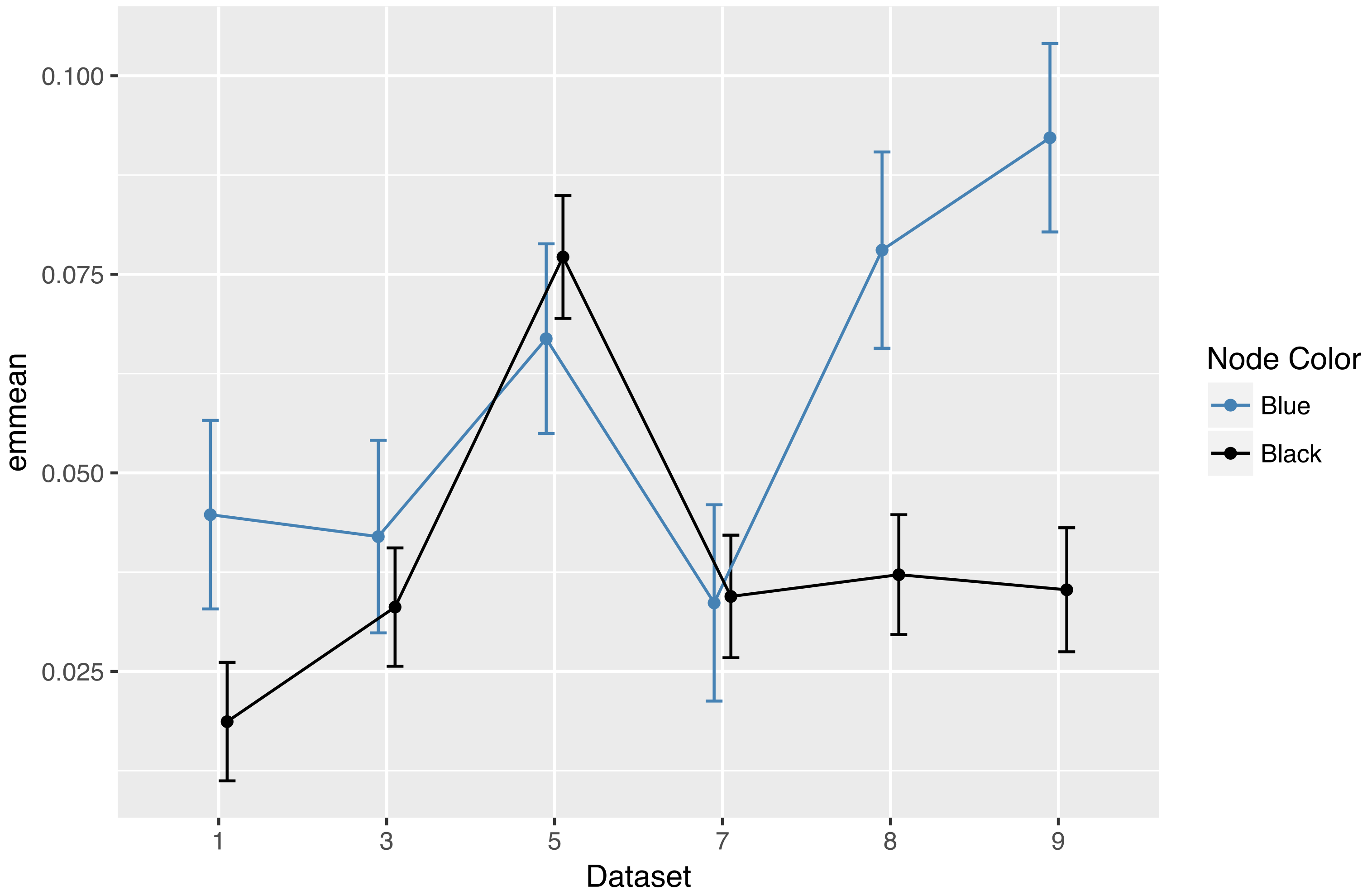


Number of clusters, fixed effects

Color of nodes vs. Dataset

Blue Nodes		Black Nodes	
Dataset	.group	Dataset	.group
7	1	1	1
3	1	3	2
1	12	7	2
5	23	9	2
8	34	8	2
9	4	5	3

Estimated Marginal Means for Node Color vs. Dataset for Number of Clusters task, graphics conditions



Layout Results by Task

Datasets 7 and 9 are less different for layout conditions than graphics.