University of Essex

Universidad Católica del Uruguay
Departamento de Ciencias Sociales y Políticas

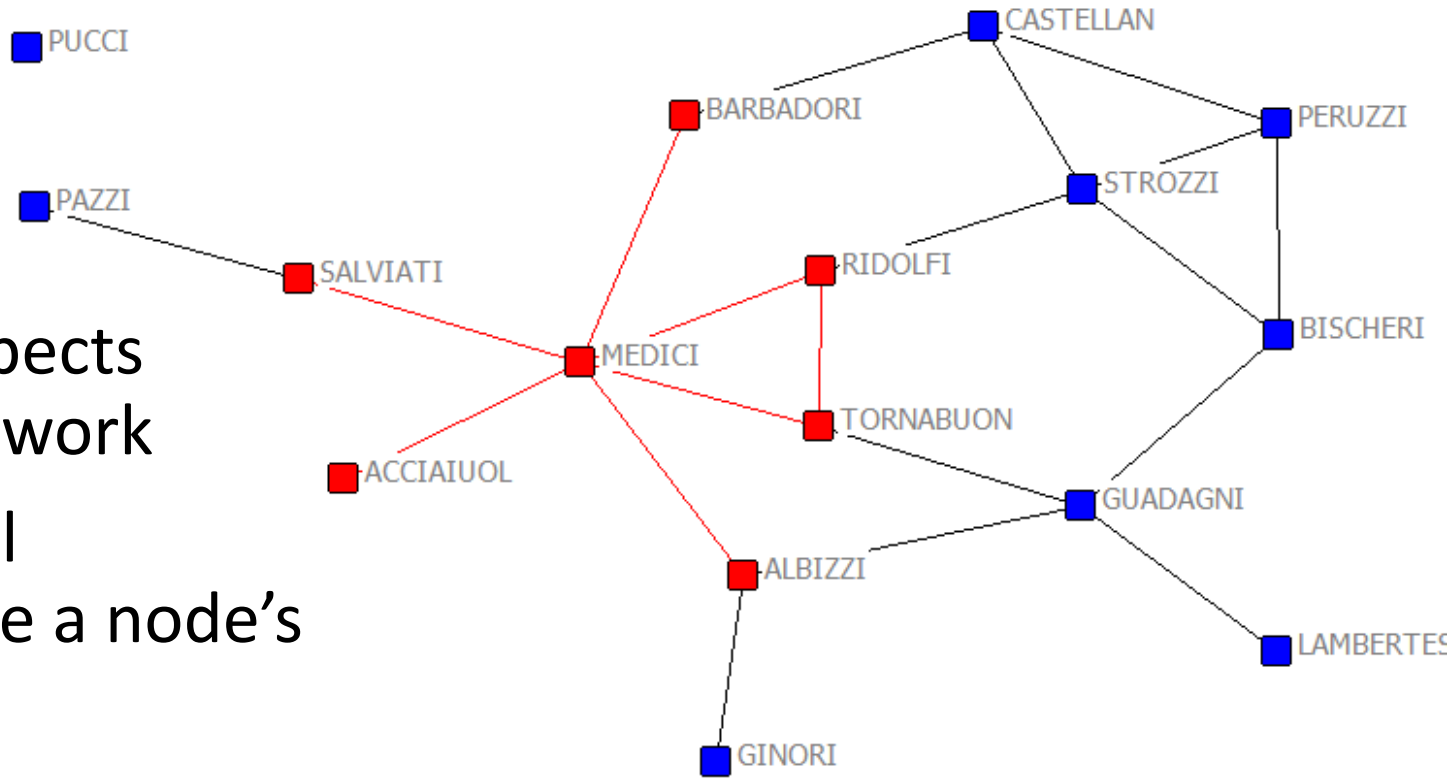
Social Network Analysis

Day 4 - Social Capital & Ego-Networks

ESSEX SUMMER SCHOOL in Social Science Data Analysis

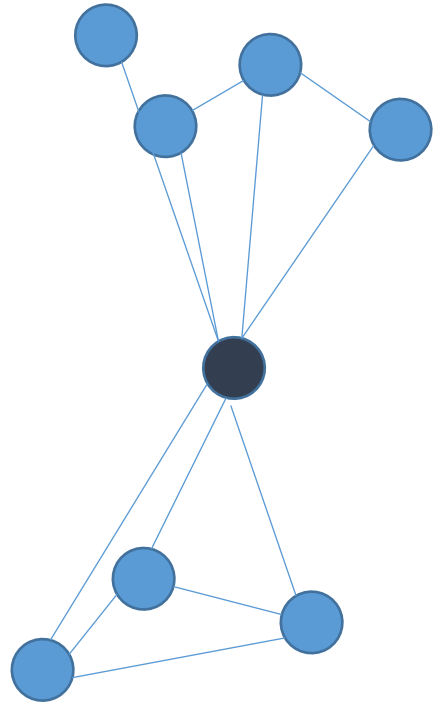
Orientation

- We have been talking about centrality
- Centrality is node-level measure that captures aspects of a node's position in network
- Now we look at node level measures that characterize a node's <u>local</u> environment



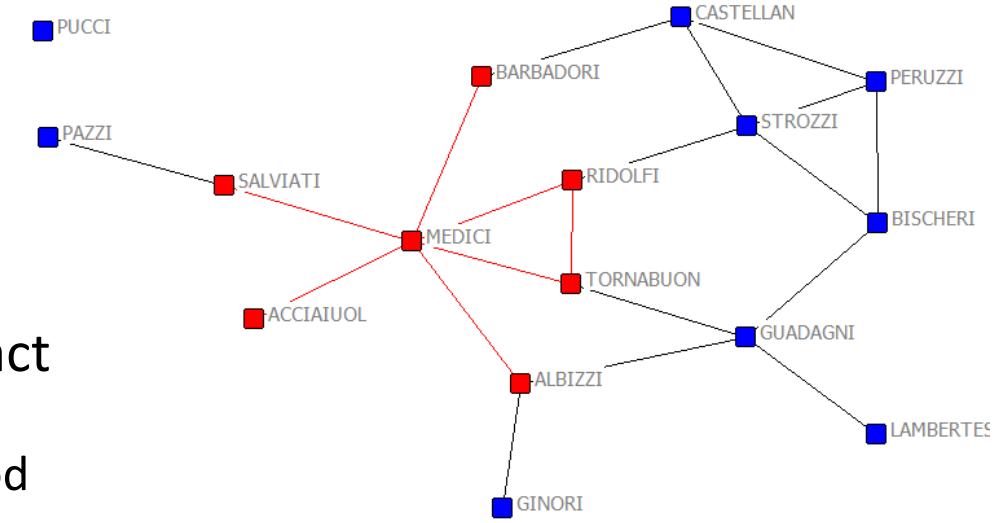
The ego network (network neighbourhood)

- An ego network consists of
  - A focal node (ego)
  - The nodes ego is connected to (the alters)
  - Ties among the alters
  - *Attributes of Alters*
- We characterize the ego network in terms of
  - Size
  - Structure
  - Composition



Why do we care?

- Influence
  - People are influenced by the people they interact with
    - The social environment; the network neighborhood
  - Behaviors, customs, ideas, ways of talking
    - In short: STYLE
  - Both egos and alters can be the drivers
- Selection
  - It is revealing who – what kind of people -- a person associates with
- Social capital
  - What kinds of resources is a person able to access because of their connections?

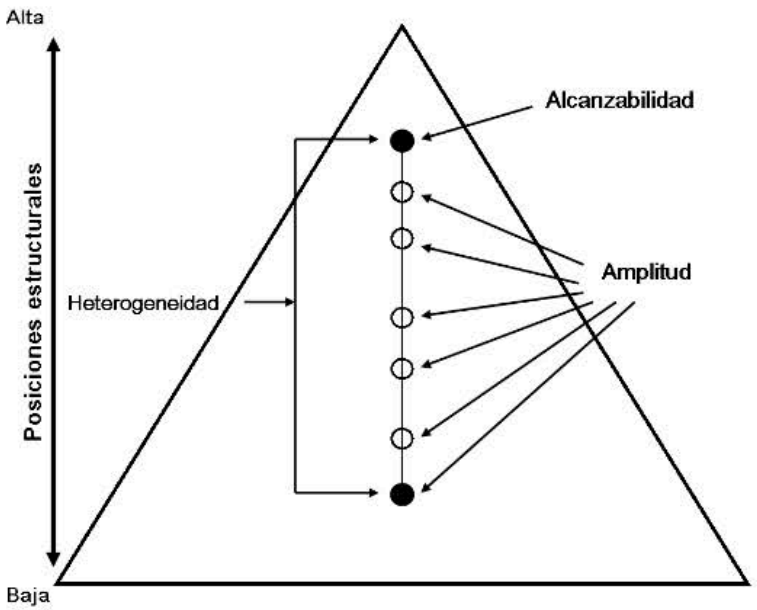


Impetus for influence

Ego \ Alters	Active	Passive
Active	School	Imitation
Passive	Coercion & conformity	Osmosis

Approaches to Social Capital

- Topological (shape-based)
  - Burt (structural holes)
  - Coleman, Putnam (connectivity/embeddedness)
- Connectionist (attribute-based)
  - Lin
- Combination of shape-based and attribute-based
  - Gould & Fernandez



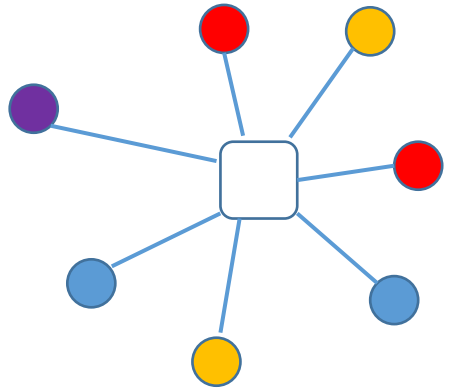
3. **Medidas de Capital Social.** Adaptado de Lin (2001)

Tabla 1
Medidas de capital social a partir del método del Generador de Posiciones

Denominación	Cálculo	Tipo de medición	Antecedentes
Máximo prestigio accedido	Puntuación de la ocupación más alta accedida.	Presencia de elementos específicos en la red social.	Lin 2001;
Rango de prestigio accedido	Diferencia entre el prestigio más alto y más bajo accedidos	Diversidad del CS	Granovetter 1973; Flap 1991; Burt 1992; Erickson 1996b; Lin 2001
Número de posiciones diferentes accedidas	Número total de ocupaciones en que el informante conoce a alguien	Diversidad del CS	
Prestigio promedio accedido	Media del prestigio de todas las ocupaciones en las que el informante conoce a alguien	Volumen o extensión del CS	Campbell, Mardsen and Hurlbert 1986
Prestigio total accedido	Suma del prestigio acumulado de todas las posiciones accedidas	Volumen o extensión del CS	Boxman, Flap, and Weesie 1992; Hsung y Hwang 1992

Fuente: Van der Gaag, Snijders, & Flap, 2008

Social capital theory



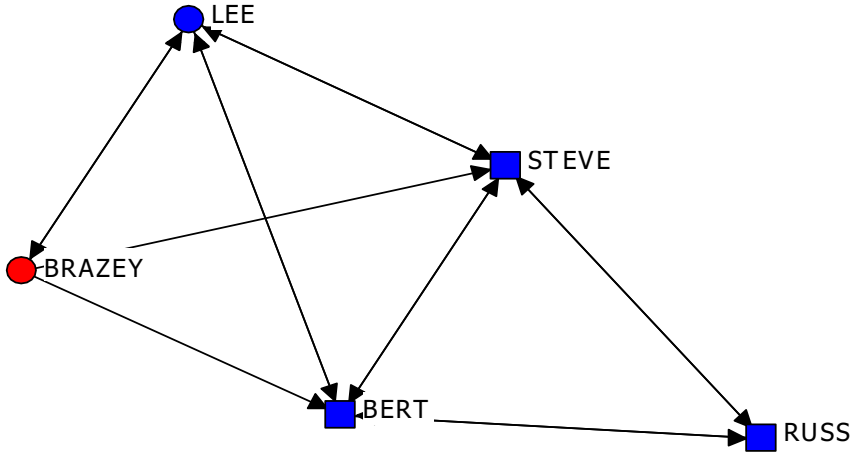
- Two flavors
  - Social resource theory, associated with Nan Lin
  - Structural holes theory, associated with Ron Burt
- Social resource theory
  - You don't personally have to have every skill, every resource needed for achievement
  - You can borrow them, control them via social ties
  - Who you are connected to, and what can they do for you?
- Structural holes theory
  - Advantages of being connected to many others who are unconnected to each

Positional & Attribute-based approaches to social capital

Focus	Measurements	Indicators
Embedded resources	Network resources	Range of resources, best resources, variety of resources, composition (average resources), contact resources
	Contact statuses	Contacts' occupation, authority, sector
Network locations	Bridge to access to bridge	Structural hole, structural constraint
	Strength of tie	Network bridge, or intimacy, intensity, interaction & reciprocity

Nan Lin (Social Resource Theory)

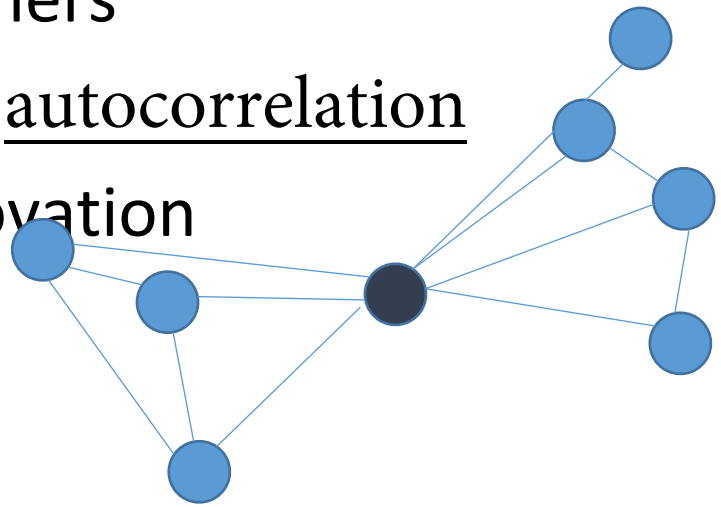
- Lin's view
  - Valued resources in societies represented by wealth, power and status;
  - Social capital is analysed by the amount or variety of such characteristics of others with whom an individual has ties to;
  - In short, it is the **attributes** of those you are connected to that matters.



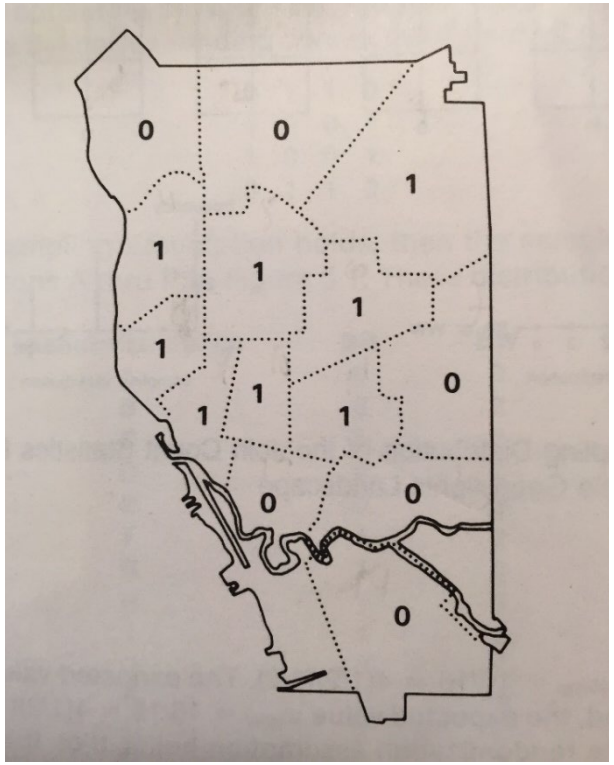
•We can look at the composition of an ego-net in terms of **heterogeneity** in **attributes** of the alters.

Social influence theory

- Your style -- attitudes, beliefs, behaviors, goals etc -- is a function of many things, including the attitudes, beliefs behaviors & goals of your alters
- Contagion: diseases are caught by interacting with infected others
- Mathematically: autocorrelation
- Diffusion of innovation



Autocorrelation pattern (characteristic of contagion)
Buffalo, NY



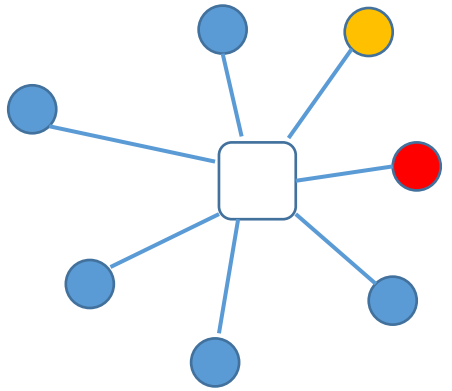
Presence of prostitution, by county

Key premise is that ‘all politics are local’

- Things may diffuse from a great distance away, but ultimately they reach you through one of your **direct contacts**
- You aren’t influenced by strangers except through their influence on your friends
  - A’s state →B’s state →C’s state
  - Unless A is directly connected to C, her only influence on C is through influencing B
- But is the premise true?

Egonet composition

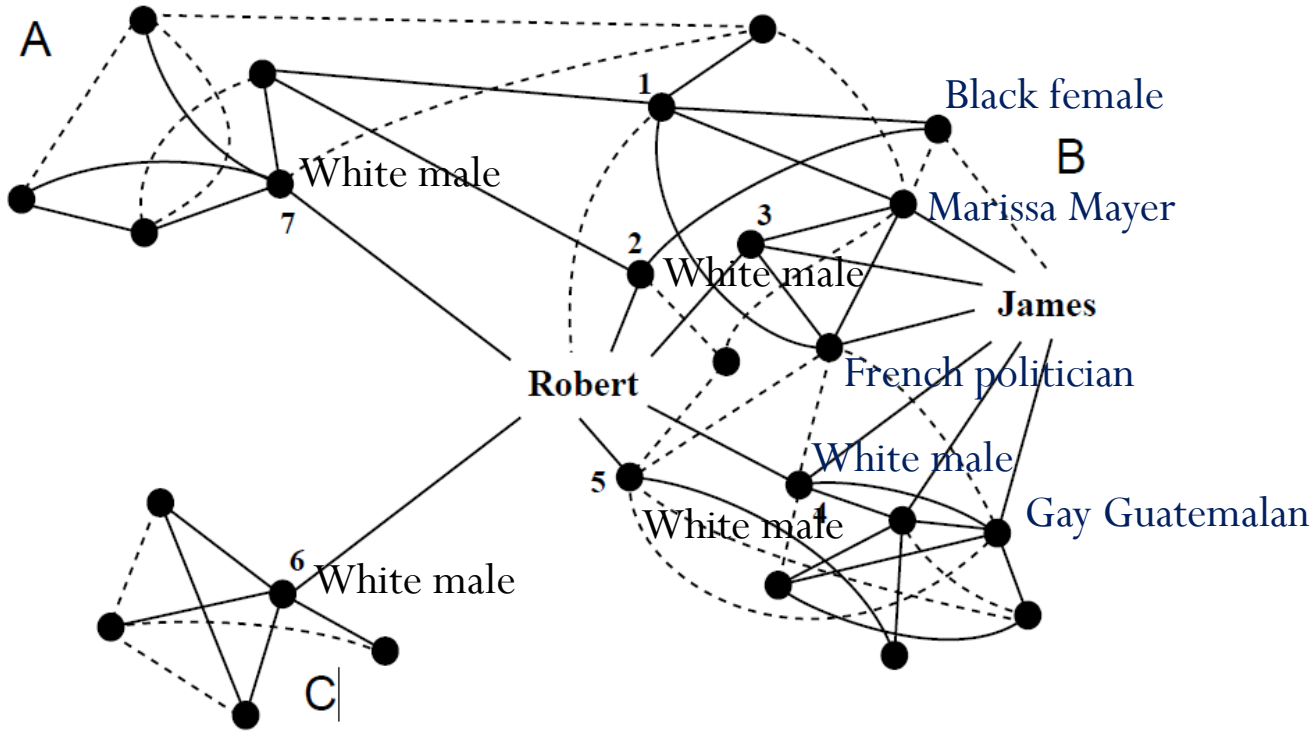
- Summarizing who is in a person's ego network
  - Examining attributes of ego's alters
    - How many of what kind?
- Categorical attributes
  - Gender, race
- Continuous attributes
  - Age, income



- Objective: summarize distribution of attribute values among ego's alters

Egonet composition

- Summarizing who is in a person's ego network
  - Examining attributes of ego's alters
- Categorical attributes
  - Gender, race
- Continuous attributes
  - Age, income
- Objective is to summarize distribution of attribute values among ego's alters



Usage

- Networ variables to add to database
- Is employee likely to perform better
  - if her personal network includes a lot of higher-ups?
  - Or is more diverse?
- What predicts % of women in network?

Id	zcomp	zeval	ic	mgr	size	c-suite contacts	%c-suite	Rank het	# fem	% fem	Avg age	Min age	Max age
1	0.26	0.46	0	2	8	1	0.13	0.72	3	0.38	34.00	25	55
2	1.51	0.36	0	3	5	2	0.40	0.64	1	0.20	43.20	39	46
3	-0.54	0.36	0	2	9	1	0.11	0.49	3	0.33	34.33	28	41
4	2.95	-0.94	0	4	23	6	0.26	0.70	1	0.04	46.52	39	56
5	-0.45	0.69	0	2	13	2	0.15	0.38	2	0.15	40.54	26	57
6	1.82	0.57	0	3	10	1	0.10	0.58	6	0.60	38.75	26	47
7	1.39		4	0	9	4	0.44	0.59	0	0.00	47.38	37	61
8	0.57		0	4	22	5	0.23	0.73	4	0.18	43.89	33	60
9	-1.01		0	3	2	2	1.00	0.00	0	0.00	49.00	44	54
10	0.78		3	0	9	0	0.00	0.49	1	0.11	42.33	34	52
11	-1.56	0.21	0	2	13	2	0.15	0.76	3	0.23	42.18	34	50
12	2.99	2.12	0	3	19	2	0.11	0.68	8	0.42	34.94	29	45
13	-0.49		0	2	12	1	0.08	0.67	3	0.25	39.75	30	51
14	-0.26	-0.94	2	0	7	1	0.14	0.69	4	0.57	45.00	43	48
15	0.20	-0.94	0	3	7	3	0.43	0.69	1	0.14	41.86	33	57
16	1.78	0.46	0	4	15	1	0.07	0.73	4	0.27	35.79	26	56

Composition / cat. alter attributes

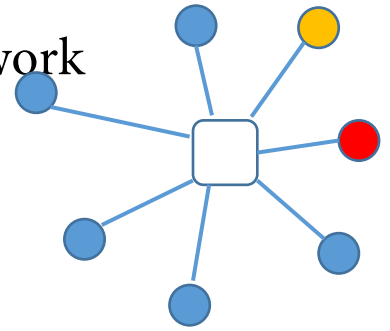
- Counts of the number of alters of different types
  - How many gay people is ego friends with?
  - How many white people does ego seek advice from?
  - Can characterize ego by modal type of attribute
    - Bill has a male network
    - Jane has a mixed network
- Frequencies or proportions?
  - Usually frequencies
    - It's the fact you have 3 computer savvy friends to draw on that matters, not what pct of your network they are
  - But proportions good for inferring preferences
    - Bill seems more comfortable with men …

Bill	freq	prop
Male	14	0.875
Female	2	0.125
	16	1

Jane		
Male	8	0.5
Female	8	0.5
	16	1

Categorical alter attributes

- Counts of the number of alters of different types
  - How many gay people is ego friends with?
  - How many white people does ego seek advice from?
  - Expressed as either frequencies or proportions
  - Can characterize ego by modal type of attribute
    - Bill has a male network
- For each ego, we count, say, # of men in their personal network. This variable, NumMen, can now be added to our database of information about ego
  - H: People who have MORE MEN in their network will be evaluated higher.



Bill	freq	prop
Male	14	0.875
Female	2	0.125
	16	1

Jane		
Male	8	0.5
Female	8	0.5
	16	1

egoid	female	netsize	nummen	eval
1	0	12	3	91
2	0	15	5	83
3	1	3	1	95
4	1	5	5	86
5	0	16	14	72
6	1	17	8	66
7	1	4	1	85
8	0	8	4	88

Continuous alter attributes

- How to summarize to what extent an ego is connected to rich people, or people with disposable time, or people willing to help?
  - Median or mean
  - Can use weighted mean where weight is the strength of tie to the alter
  - Total
  - Max and min

egoid	computer expertise	Highest expertise of any friend	Happiness
1	12	3	91
2	15	5	83
3	5	1	36
4	5	17	90
5	16	14	95
6	17	8	89
7	0	1	23
8	1	4	30

Measures for social resource theory

- Most common measure is **sum** of resources of alters

$$c_j = \sum_i w_{ij} r_i \qquad C = WR$$

  - r_i is amount of resource controlled by alter i, -- it is i's human capital
  - w_{ij} is the extent to which i values and needs j
  - c_j is the total amount of resource that j can indirectly access
  - $c_j + r_j$ = total capital j can access including their own human capital

- Alternatively, might use **maximum** resource of any alter, weighted by influence

$$c_j = \max_i (w_{ij} r_i)$$

For sum, can use matrix multiplication

- Nan Lin SRT $C = WR$ $c_j = \sum_i w_{ij} r_i$
- Count up total resources of different kinds available to each ego

Is owed by

	a	b	c	d	e	f	g
a	0	0	1	0	1	0	1
b	1	1	1	0	0	1	1
c	0	0	0	1	1	0	1
d	0	0	0	0	1	0	0
e	1	0	1	1	0	1	1
f	0	1	0	1	0	0	1
g	0	0	1	1	0	0	0

Resources available

	$	hrs
a	20	120
b	10	60
c	5	250
d	30	160
e	40	80
f	1	360
g	10	300

Social capital

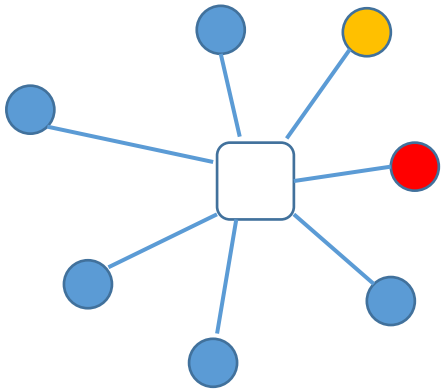
	$	hrs
a	55	630
b	46	1090
c	80	540
d	40	80
e	66	1190
f	50	520
g	35	410

Incoming or outgoing ties?

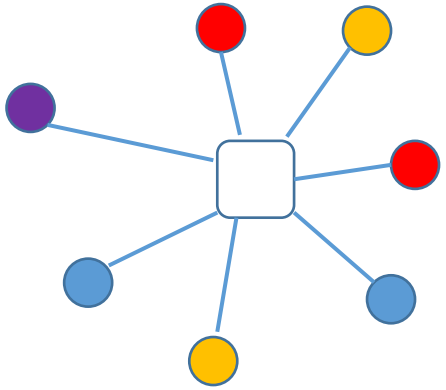
- For social capital research may often want to use incoming ties.
  - “who seeks you out for help at work?”
  - those who depend on you more likely to help you when you need them
  - “who do you respect and admire?”
- For influence research, may want to use outgoing ties
  - The people you like are the ones that have influence on you
- You can always transpose the matrix

Heterogeneity

- A summary of the diversity of ego's alters with respect to a given alter attribute



Homogenous



Heterogenous

Alter heterogeneity – categorical attribute

- Blau / Herfindhal / Hirschman index
  - P_k is proportion of ego's alters that fall in category **k**

$$H = 1 - \sum_k p_k^2$$

  - H = 0 if all alters in one category
  - H = 1 – 1/K if all categories have equal frequency
    - The measure cannot reach 1.0 unless there are infinite categories

Bill	freq	prop	prop^2	K =	2
Male	14	0.875	0.765625	H =	0.219
Female	2	0.125	0.015625	IQV =	0.438
	16	1	0.78125		

Jane					
Male	8	0.5	0.25	H =	0.5
Female	8	0.5	0.25	IQV =	1
	16	1	0.5		

Flaw in the measure: max score is 0.5; will only approach 1 if k -> infinity

Alter heterogeneity – categorical attribute

- Agresti's IQV
  - Divide H by 1 – 1/K so that measure runs between 0 and 1
    - Gives amount of diversity given the number of categories

$$IQV = \frac{H}{1 - {}^{1}/_{K}}$$

Normalization of Blau's index, not sensitive to # categories

- But H could be seen as best measure of diversity, because it is not satisfied until the number of categories →∞, which would imply massive diversity

Who should have higher heterogeneity in their ego-network?

1. **Person A, 3 friends, each different nationality**
2. **Person B, 30 friends, each different nationality**

Ha < Hb

IQVa = IQVb

Bill	freq	prop	prop^2	K =	5
Accounting	14	0.560	0.314	H =	0.618
Economics	5	0.200	0.040	IQV =	0.772
Finance	4	0.160	0.026		
Management	1	0.040	0.002		
Marketing	1	0.040	0.002		
total:	25	1.000	0.382		

Jane					
Accounting	5	0.200	0.040	H =	0.800
Economics	5	0.200	0.040	IQV =	1.000
Finance	5	0.200	0.040		
Management	5	0.200	0.040		
Marketing	5	0.200	0.040		
total:	25	1.000	0.200		

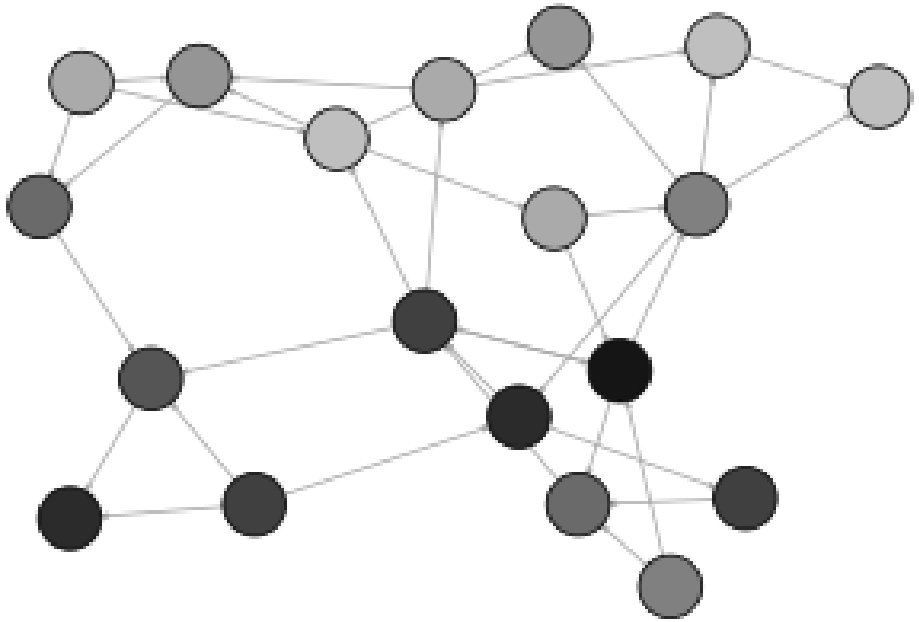
Continuous heterogeneity

- Standard deviation
- Coefficient of variation – sd/mean
  - Might be good idea. Egos vary in their alters' mean values

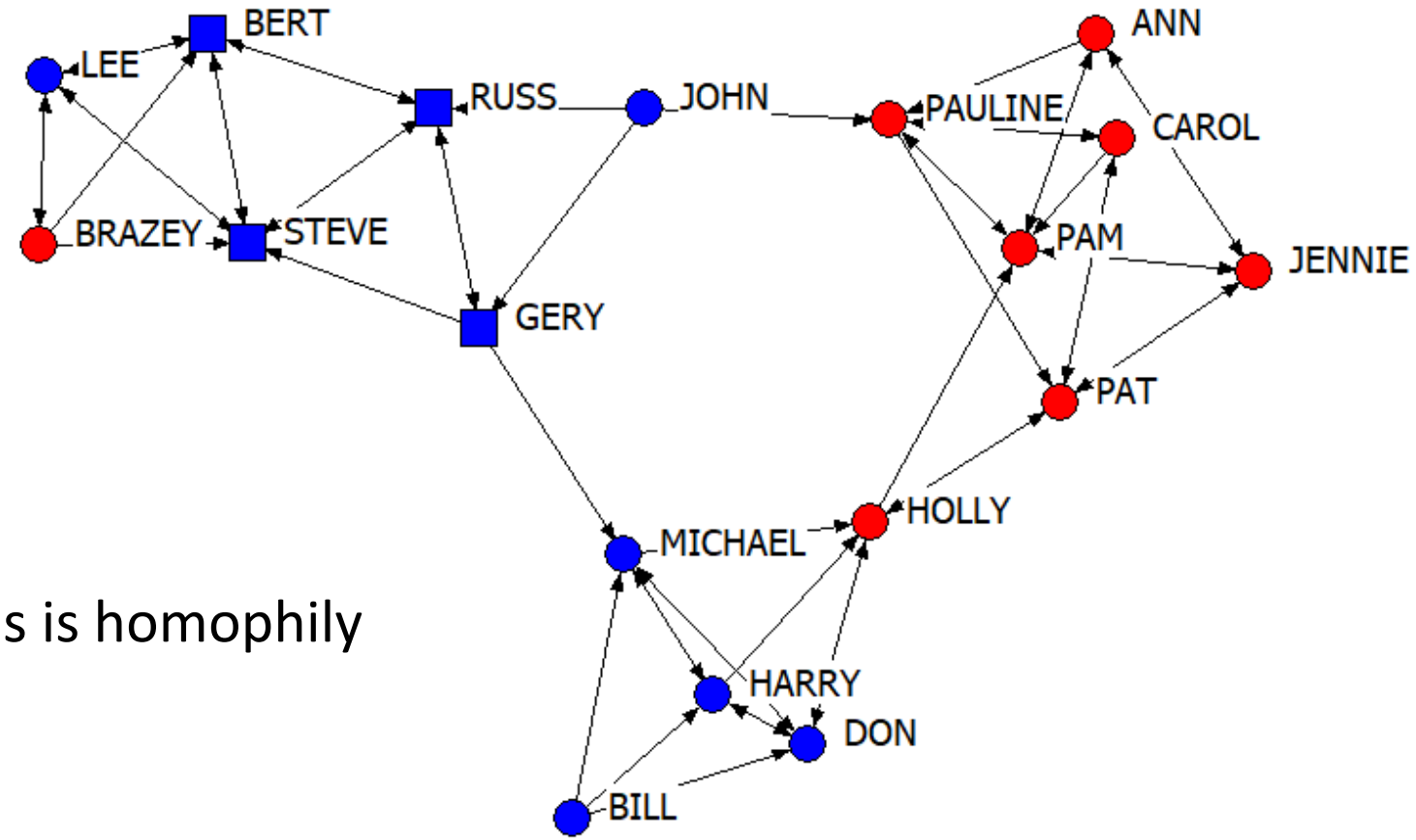
id	Avg	Sum	Min	Max	SD	EstSD	CV	Num
1	11.6	104.7	4.7	27.0	6.9	7.3	0.6	9
2	11.9	118.7	3.3	28.0	8.4	8.8	0.7	10
3	12.1	72.3	4.8	27.0	7.1	7.7	0.6	6
4	13.3	93.3	4.7	27.0	7.0	7.5	0.5	7
5	11.0	110.1	0.3	27.0	7.3	7.7	0.7	10
6	13.9	97.3	5.4	30.0	7.8	8.4	0.6	7
7	16.9	50.8	10.4	28.0	7.9	9.6	0.5	3
8	13.1	65.5	7.5	27.0	7.1	8.0	0.5	5
9	14.7	88.4	3.3	28.0	9.4	10.3	0.6	6
10	8.8	70.5	3.3	12.8	3.6	3.8	0.4	8
11	8.9	124.8	0.3	19.6	4.6	4.7	0.5	14
12	13.9	110.8	4.8	28.0	8.2	8.8	0.6	8
13	15.2	30.3	3.3	27.0	11.8	16.7	0.8	2
14	12.0	71.8	3.3	30.0	8.8	9.6	0.7	6
15	12.6	113.5	3.3	28.0	8.5	9.1	0.7	9
16	11.6	58.1	7.5	19.6	4.3	4.8	0.4	5
17	12.5	224.9	3.3	30.0	8.0	8.2	0.6	18
18	17.7	70.8	11.7	27.0	6.2	7.2	0.4	4
19	12.4	123.8	3.3	27.0	6.2	6.6	0.5	10
20	12.5	62.6	4.8	27.0	7.6	8.5	0.6	5
21	13.6	81.3	3.3	28.0	8.1	8.9	0.6	6

Ego-alter similarity

- Measuring the extent to which ego's alters are just like ego with respect to attributes of interest
- Depending on the direction of causality, can either reflect
  - **Homophily**. The tendency to form positive ties with those similar to one self
    - Demographics, attitudes, activities
  - **Influence**. The tendency to adopt the ways of those one is connected to



Sometimes direction of causality is clear



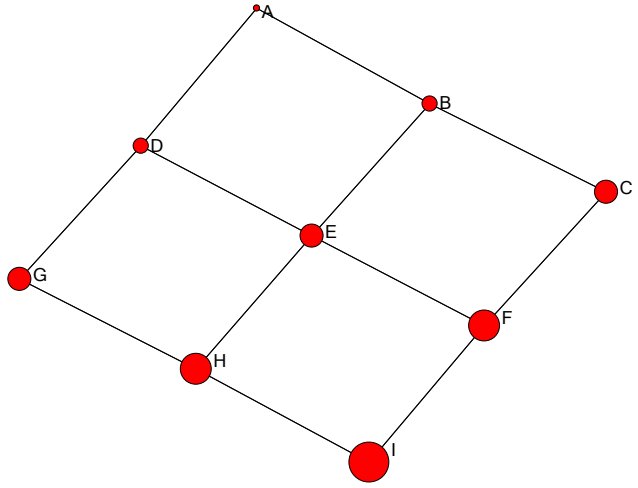
This is homophily

Homophily

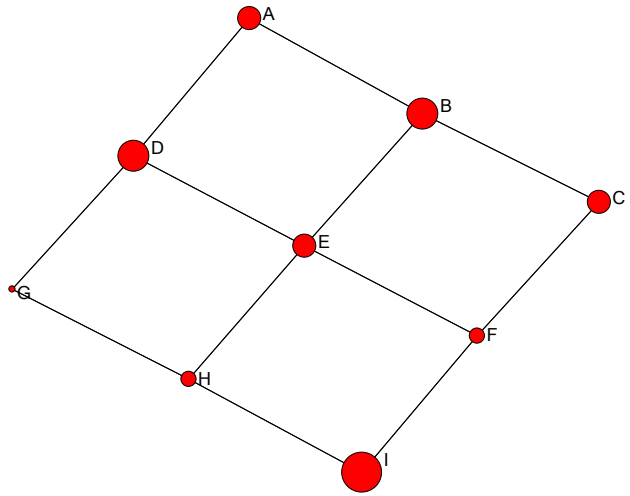
- Tendency for people to form positive ties with people similar to themselves on socially significant attributes
  - Race, Gender, Age, Education, Social class
  - Interests, activities, etc – social foci (Feld)
  - Behaviors, style, appearance
- Why?
  - Convenience of communication
  - Evolutionary psychology?
  - Status systems
    - If all seek ties to highest status possible, we will end up with ties to people of our own status

Testing for homophily/influence

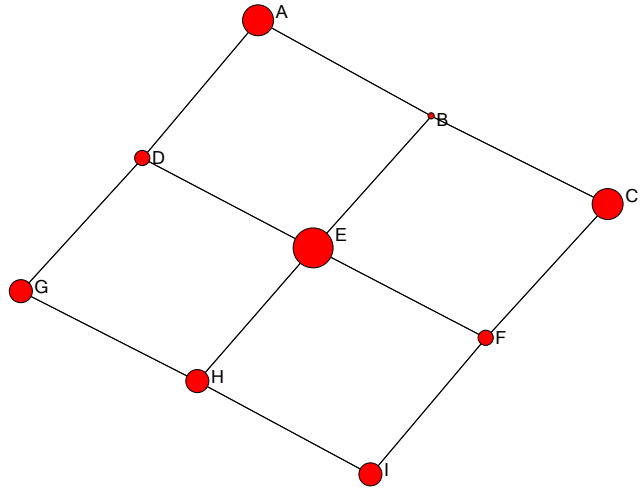
- **Problem**: they both look the same, they result in *autocorrelation*
- Autocorrelation is what we can test for
  - Direction of causality usually can't be determined statistically
- Essentially, we correlate presence/absence of tie with similarity on attribute



homophily



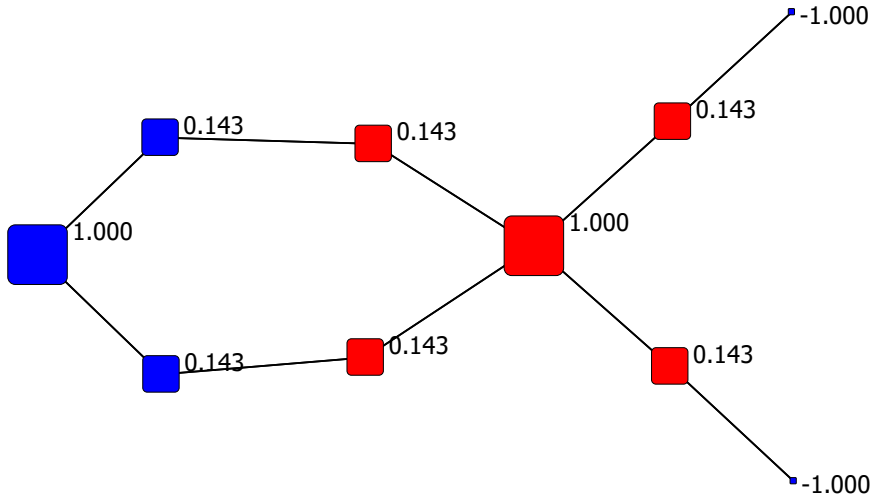
heterophily



random

Measuring homophily (ego-alter similarity)

- What we actually measure is the extent to which egos resemble their alters
  - Men's friends tends to be men, women's friends women
- Selection (homophily) or influence?
  - Is ego seeking out similars (selection/homophily), or is ego influencing alters to become like self (influence)?
  - In case of gender, probably selection
  - In case of smoking, probably both
- Multiple measures
  - In general can't distinguish selection from influence



Measures of node-level ego/alter similarity

- Pct of matches (M)
  - What proportion of ego's alters have same attribute value as ego does?
    - If ego is male, what proportion of alters are male?
      - If 0, then perfectly heterophilous. If 1, then perfectly homophilous
- E-I index (Krackhardt and Stern)
  - $\frac{E-I}{E+I}$, where E (external) is number of alters different from ego, and I (internal) is number of alters same as ego (measure of HETEROPHILY)
  - Is a linear rescaling of % matches: EI = 1 – 2M, where is M is % matches

Note that if most people are white, and if people chose others without regard for color, most whites will be homophilous, and most blacks will be heterophilous, regardless of intention

But homophily normally means preference

Supply

- Suppose ego is female, 8 of her 10 work friends are female. The M = 80% figure we observe is an outcome of both preference and availability
  - Suppose our female respondent is a nurse and her workplace is 80% female
  - If she is completely indifferent to gender, what percent of her alters would we expect to be female? 80%.
  - If her friends were only 60% female, it would probably mean a pref for men
- In a sociometric study, we know the proportion of men and women in the roster

Baseline Expectation

  - Given this information, can calculate index of homophily bounded between 0 and 1 that measures departure from the expected values given population
    - Newman's modularity (Q), Cohen's Kappa, Freeman's Segregation

Cross-tabbing dyads -> adding non-ties

	Ego	
id	has tie with	is same gender
a	1	1
b	1	0
c	0	1
d	1	0
ego		
f	0	1
g	1	0
h	0	1
i	1	1
j	0	0

		has tie with		
		1	0	
same gender	1	2	3	5
	0	3	1	4
		5	4	9

Corr = -0.350
Yules = -0.636
% Match = 0.400
E-I = 0.200

Importance of non-ties (whole network designs only)

- % matches and E-I take into account only ties that are present
  - They don't count non-ties (choices to not befriend someone)

		Has Tie 1	Has Tie 0	
Same group	1	10	50	60
	0	5	25	30
		15	75	90

There are 90 potential alters. Ego has a tie with 15 of them. Of these 15, 10 are the same race/gender/etc as ego

  - The above measures only use the values in the "1" column, so it looks like this person prefers own kind 2 to 1
  - But if you look at who they are not tied with (0 column), you see they also prefer to NOT have a tie with their own kind – also at odds of 2 to 1
    - So they have no **preference** for own kind. It is just that there are more of them

Odds ratio and Yule's Q

		Has Tie 1	Has Tie 0	
Same group	1	a	b	a+b
	0	c	d	c+d
		a+c	b+d	T

- Odds ratio is just $\frac{a/b}{c/d} = \frac{ad}{bc}$

ad = concordant pairs
bc = discordant pairs

  - When this is 1, then we have independence – no preference for in-group ties
  - Odds ratio is unbounded
- Yule's Q bounds the odds ratio between -1 and 1

A measure that includes non-ties & is bounded

  - $\frac{ad-bc}{ad+bc}$, where 0 indicates no association, 1 is perfect homophily and -1 is perfect heterophily
- Key advantage of these measures is they are **not fooled by group sizes**

		Has Tie 1	Has Tie 0	
Same group	1	10	10	20
	0	10	10	20
		20	20	40

Yule's Q = 0

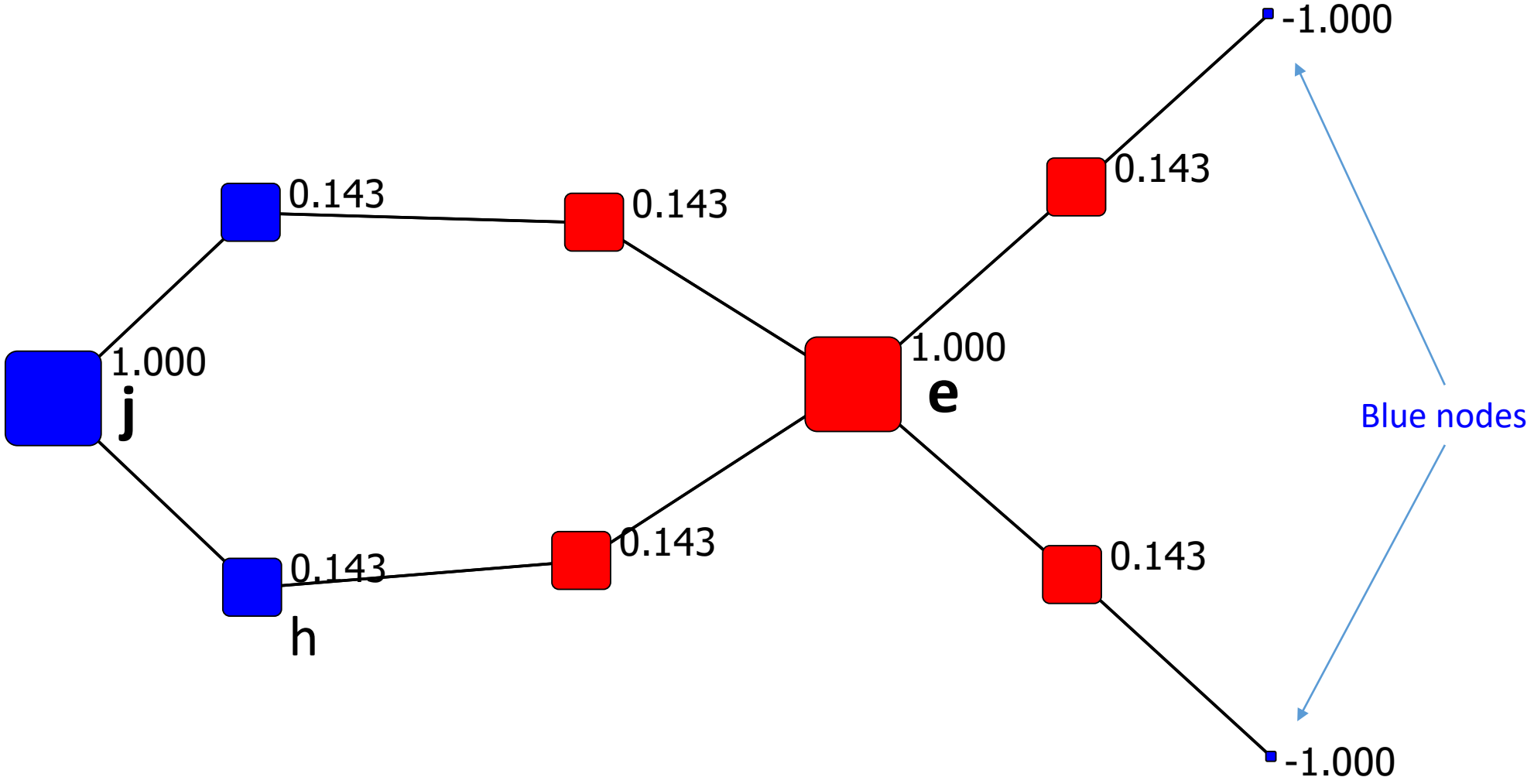
- Multiply 1st row by 10
- Multiply 1st col by 5
- Yule's Q unchanged

		Has Tie 1	Has Tie 0	
Same group	1	500	100	600
	0	50	10	60
		550	110	660

Yule's Q = 0

Network|Egonet|eg-alter similarity
~ campnet ~campattr

Example of Yule's Q



Yule's Q and the correlation coefficient (phi)

- Yule's Q is $\frac{ad-bc}{ad+bc}$,
- Pearson correlation (phi) is

		Has Tie: 1	Has Tie: 0	
Same group	1	a	b	a+b
	0	c	d	c+d
		a+c	b+d	T

$$\frac{ad-bc}{\sqrt{(a+c)(b+d)(a+b)(c+d)}}$$

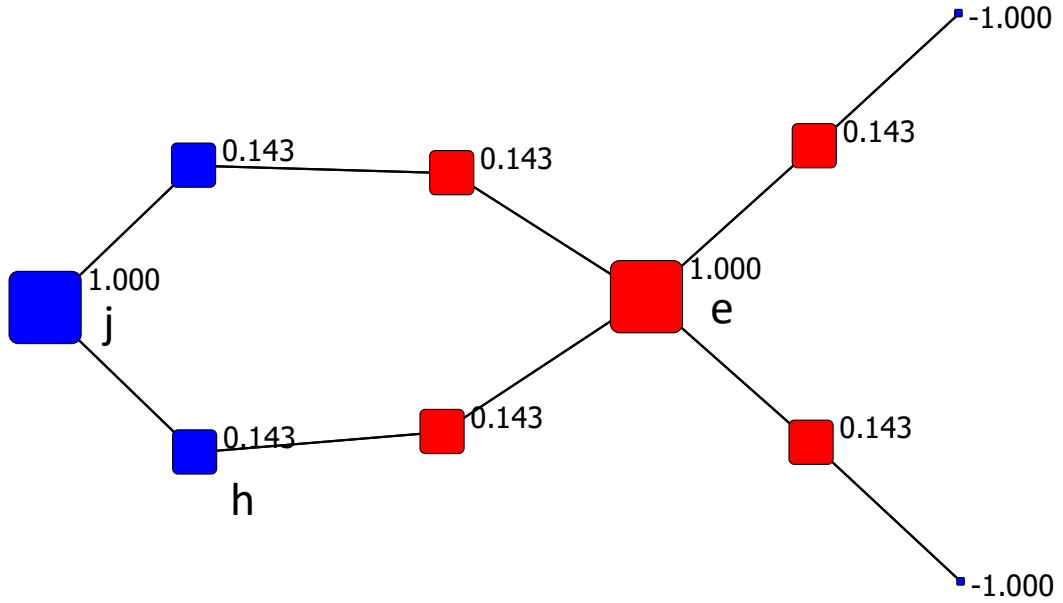
- Just a different denominator from Yule's Q
- Yule's Q has advantage that it equals 1 if nodes only have ties to same group, whereas correlation adds condition that nodes must have ties to all members of their group
  - In real settings, correlation can never achieve it's maximum value of 1

Yule's Q example

Node *h*	Tie	No Tie
Same	1	3
Different	1	4

- Yule's Q gives 1.0 to both nodes *e* and *j*, but corr only gives 1.0 for *e*
- On average, the reds are more homophilous

	% Same	EI Index	Matches	Yules Q	Cohen Kappa	Corr	fInGroup	fOutGroup
a	0.000	1.000	0.444	-1.000	-0.216	-0.316	0	1
b	0.000	1.000	0.444	-1.000	-0.216	-0.316	0	1
c	0.500	0.000	0.556	0.143	0.053	0.060	1	1
d	0.500	0.000	0.556	0.143	0.053	0.060	1	1
e	1.000	-1.000	1.000	**1.000**	1.000	**1.000**	4	0
f	0.500	0.000	0.556	0.143	0.053	0.060	1	1
g	0.500	0.000	0.556	0.143	0.053	0.060	1	1
h	0.500	0.000	0.556	0.143	0.053	0.060	1	1
i	0.500	0.000	0.556	0.143	0.053	0.060	1	1
j	1.000	-1.000	0.778	**1.000**	0.526	**0.598**	2	0



Why does *h* have a positive score? Shouldn't it be a perfect 0.0? No, because it shuns 4/5 reds but only 3/4 blues

Yules Q

- A better measure of homophily that takes into account who ego did not choose

		Same	
		1	0
Tie	1	a	b
	0	c	d

$$EI = \frac{b-a}{b+a}$$

$$Q = \frac{ad-bc}{ad+bc}$$

	B	
	blu	or
a	1	0
b	0	1
c	0.75	0.25
d	0.5	0.5
e	0	1
f	0.2	0.8
g	1	0

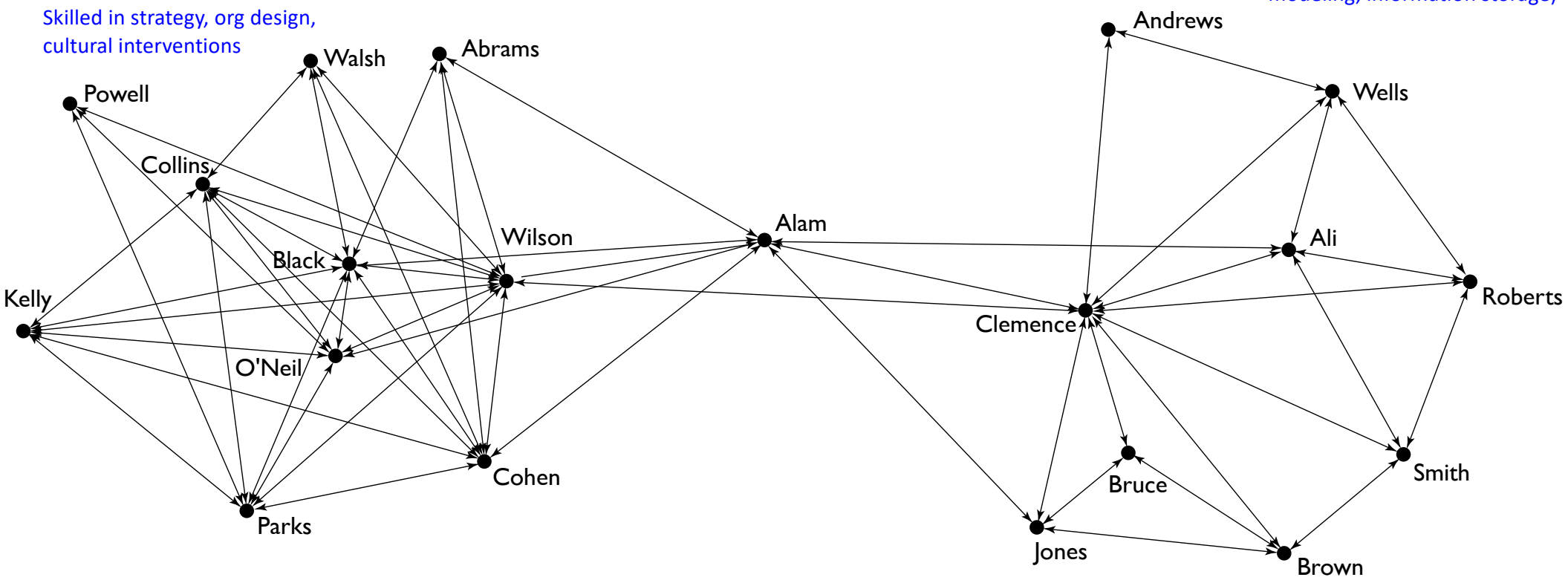
Summarizing ego-alter similarity measures

- Measures that only look at a person's ties …
  - Don't correct for availability of different groups
  - Measure realized levels of contact with ingroup vs outgroup members
  - Are the only measure available in personal network designs, where the alters aren't interviewed
- Measures that look at both ties and non-ties …
  - Correct for availability
  - Measure underlying preferences for others, so that even having just two fellow Eskimo friends is indicative of homophily
  - Can only be calculated on data collected in whole network research designs
- Remember that the measures capture ego/alter similarity
  - Whether it is due to homophily or influence must be determined elsewhere

Information & Success

Global consulting organization had group dedicated to provide thought leadership and specialized support to to the organization's knowledge management consultants. Group was composed of people with industry experience in **(1) organizational design (soft-skills)** and **(2) technical fields (data warehousing)**.
USP: **holistic knowledge management solution.** However, they were not delivering. Why?

SNA intervention – information sharing network.

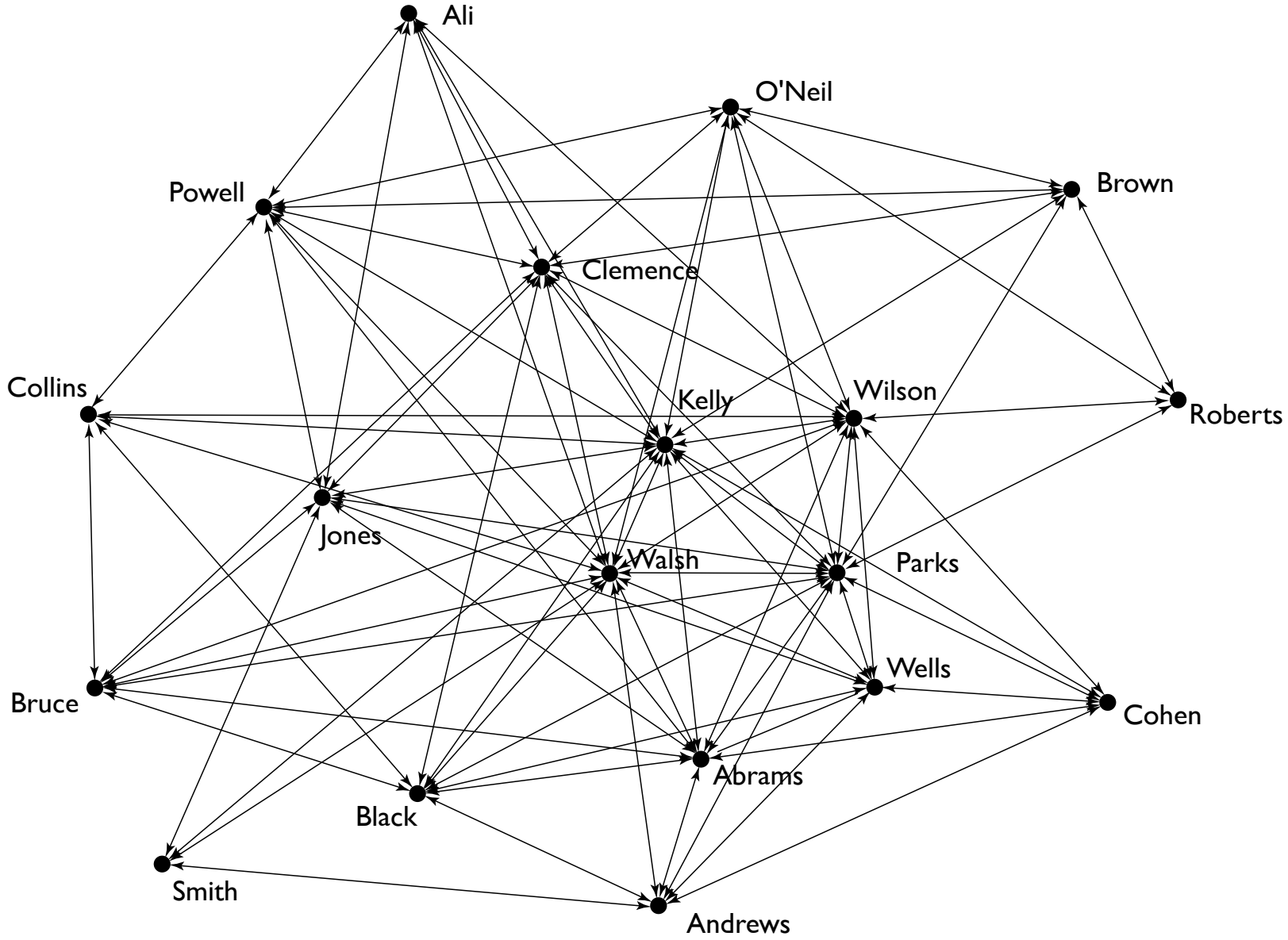


Cross, Borgatti & Parker (2002), "Making Invisible Work Visible: using social network analysis to support strategic collaboration".

Changes Made

- Cross-staffed new internal projects
  - white papers, database development
- Established cross-selling sales goals
  - managers accountable for selling projects with both kinds of expertise (forced people to integrate their approaches to addressing client problems)
- New communication vehicles
  - project tracking db; weekly email update
- Personnel changes

9 Months Later



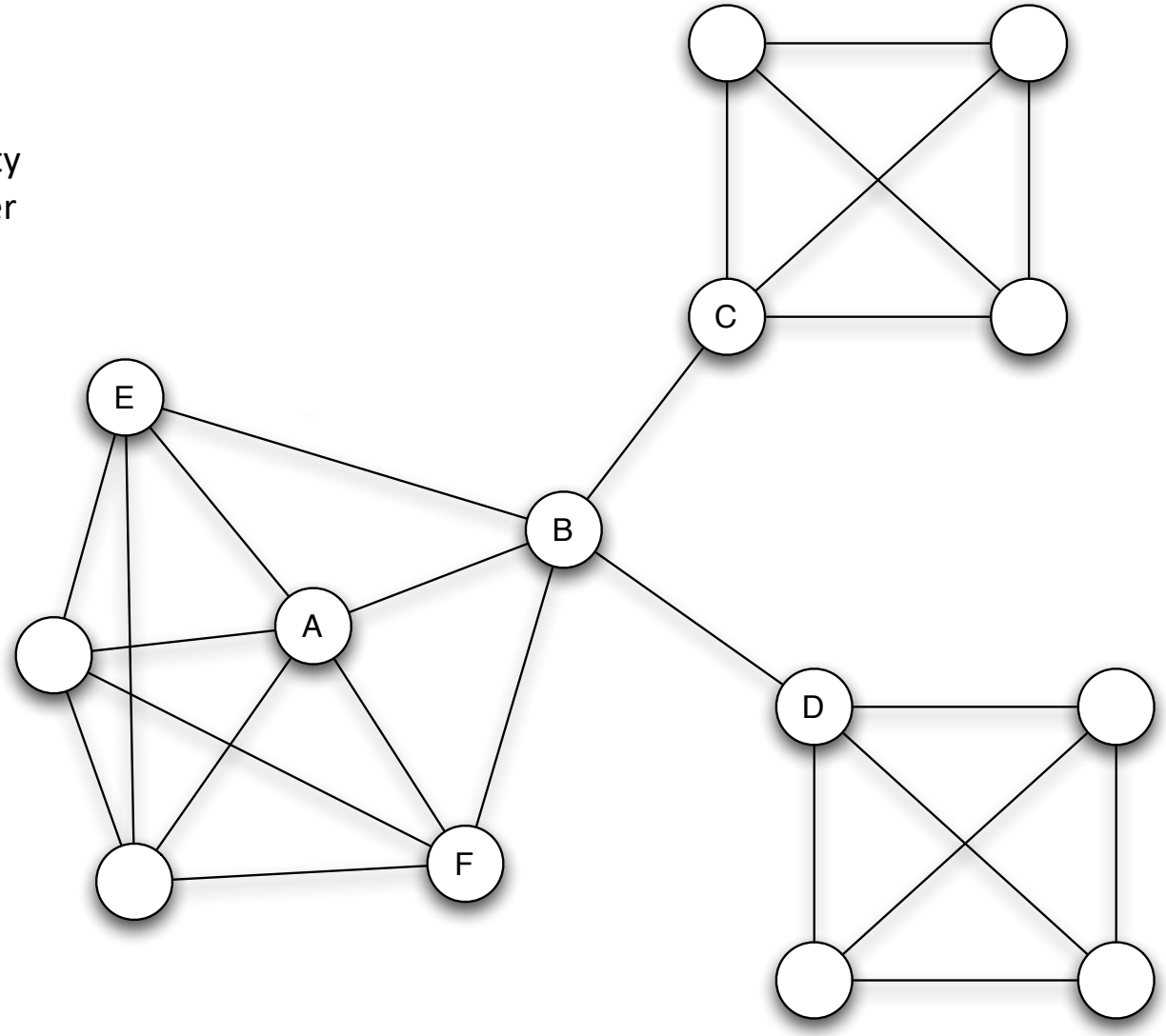
a contrasting experience: embeddedness and bridging

Bridging:

- riskier;
- Brokerage
- amplifies creativity
- Gatekeeper power

Embeddedness (# of common neighbours of edge):

- greater trust
- old, repackaged information?



strength of weak ties

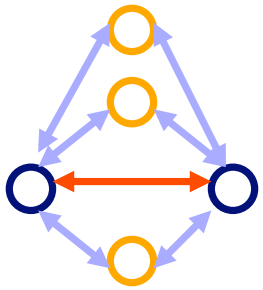
Granovetter thesis that, under many circumstances, strong ties are less useful than weak ties:

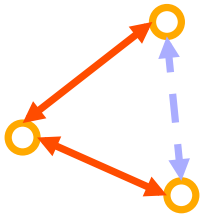
- interviewed people in Amherst, MA across professions to determine how they found out about their jobs;
- recorded whether they used social contacts and strength of the relationship;
- surprising proportion (~20%) of jobs were found through “weak ties”

Why?

- individuals involved in weak ties less likely to overlap in their neighborhoods;
- weak ties form bridges across groups that have fewer connections to each other (plays role in disseminating information).
- weak ties hold communities together;

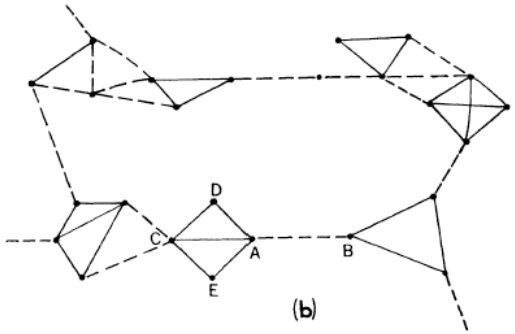
- A strong tie
  - frequent contact
  - affinity
  - many mutual contacts



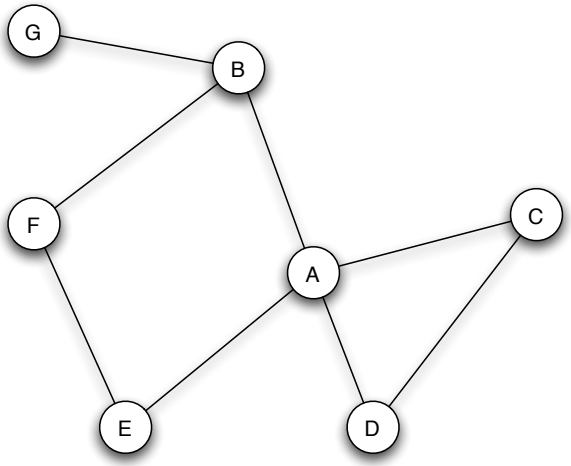


“forbidden triad”: strong ties are likely to “close”

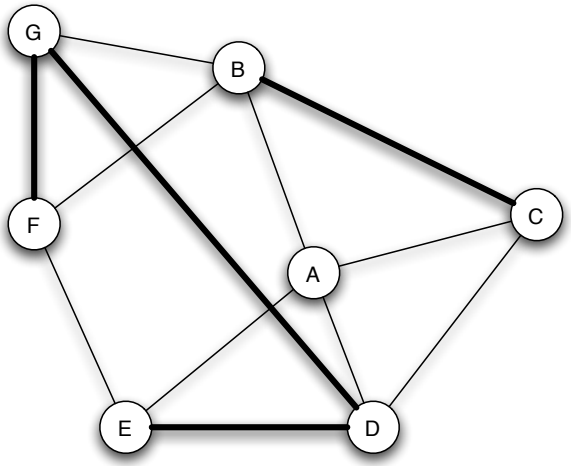
- Less likely to be a bridge (or a local bridge)



triadic closure, local bridges and weak ties



(a) *Before new edges form.*



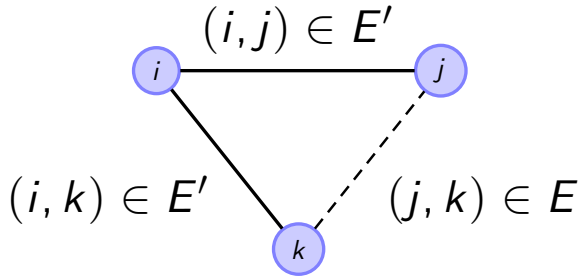
(b) *After new edges form.*

Why are we likely to observe a tie forming between B and C ?

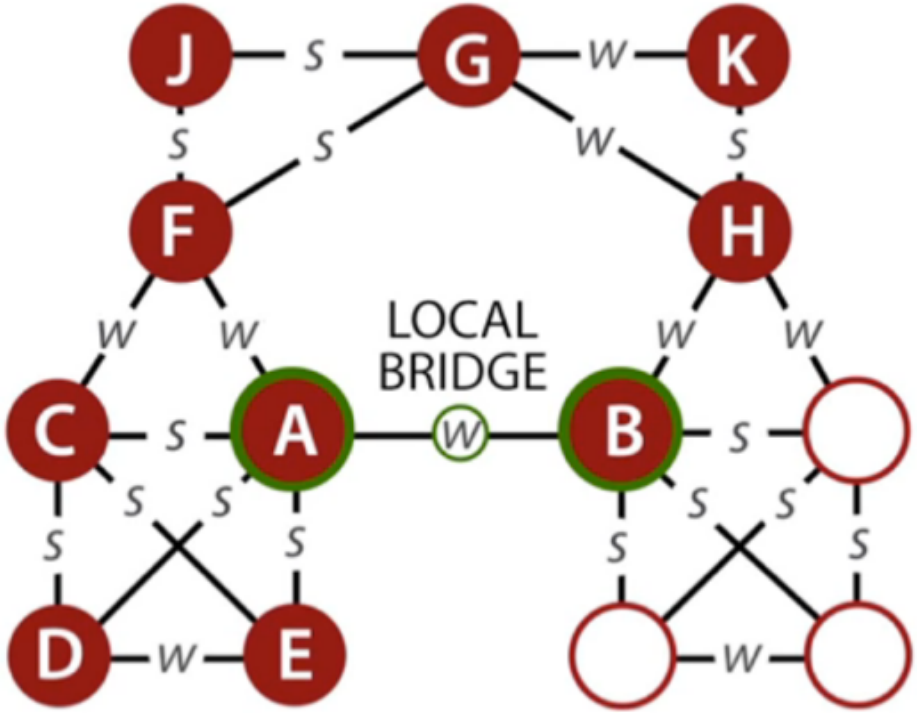
- Opportunity;
- Similarity;
- Incentive

Strong Triadic Closure Property

if $(i, j) \in E'$ and $(i, k) \in E'$, then $(j, k) \in E$.



Any local bridge will necessarily be a weak tie…
[**proof by contradiction**]



edges are either **embedded** or **bridging** (Social Capital)

"the ability of actors to secure benefits by virtue of membership in social networks or other social strucures"

Social capital is viewed as **property of a group** (favorable structures contribute to higher social capital) or as **property of an individual** (depends on position of the individual in the network). Different approaches highlight different aspects:

- **Coleman** values **embedded** edges (enable enforcement of norms, have reputational effects, enhance trusting mechanisms)
- **Burt** sees it as a tension between ***closure*** (as in Coleman's embeddedness) and ***brokerage*** (ability to broker interactions between different groups).
- **Putnam** harmonizes both views when he discusses ***bonding*** *capital* and ***bridging*** *capital*.

"Social capital is at once the **resources contacts hold** and **the structure of contacts in a network**. The first term describes *whom* you reach. The second describes *how* you reach."

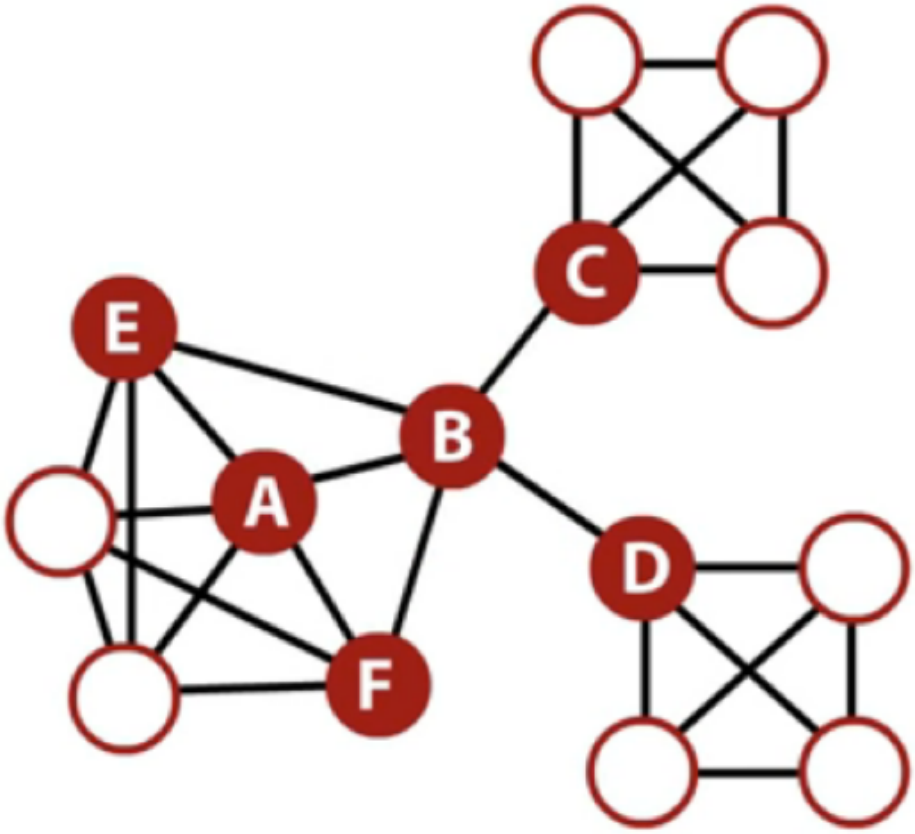
1. **"Who you reach"** – network provides an actor with access to people with specific resoruces and functions as a conduit; establishes a correlation between your resources and theirs. Relates to concept of *power* and *prestige*.

2. **"How you reach"** – social structure is capital itself that is meaured in terms of network **range** and **size**. The value of the "rate of return" can be boosted given the structure of the network and the location of the actor's contacts within that structure. The benefits include: **information** and **control**.

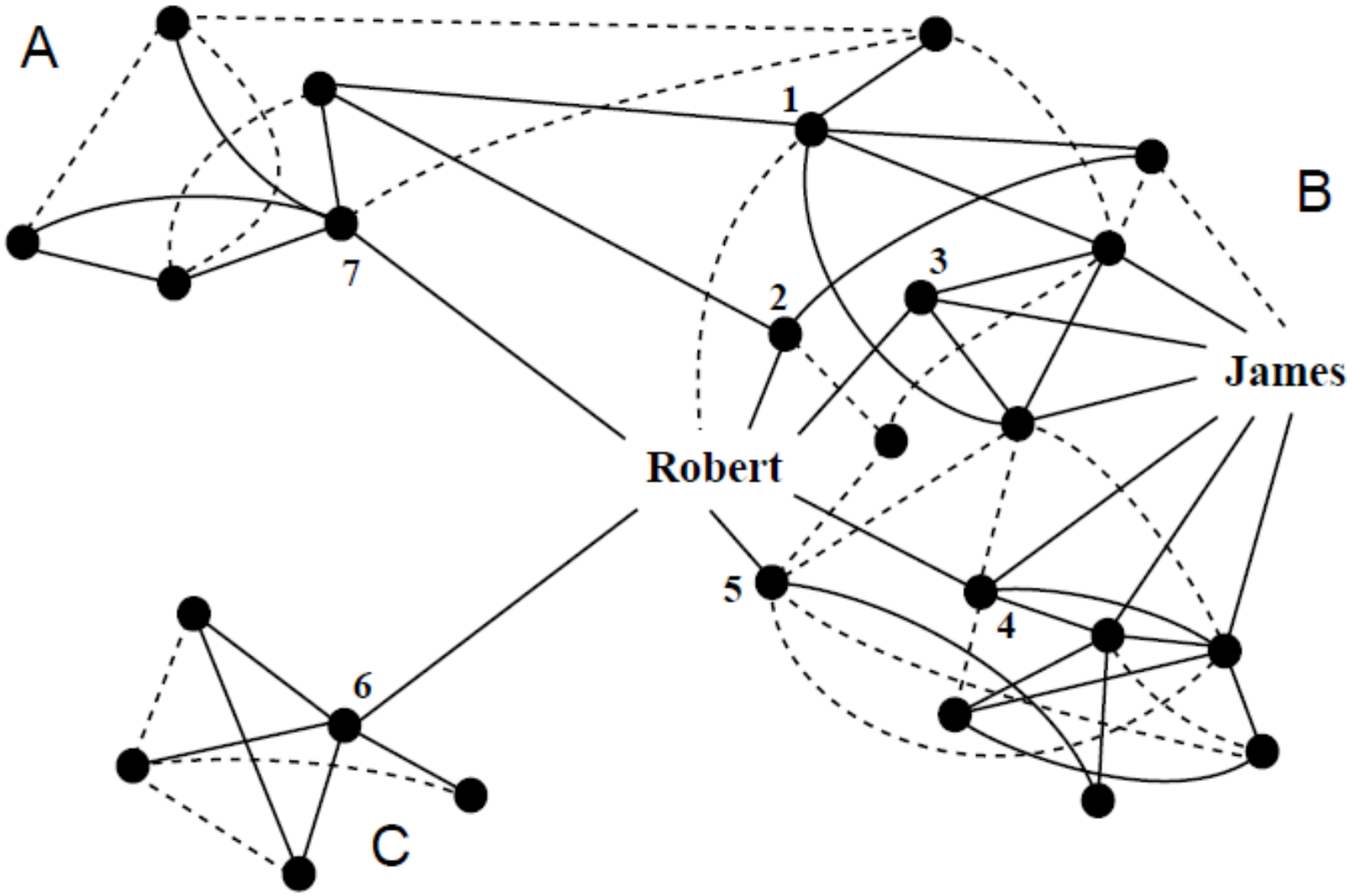
Structural holes

Structural Holes (Ron Burt)

The distribution of bridging edges among the nodes is unequal in a network…



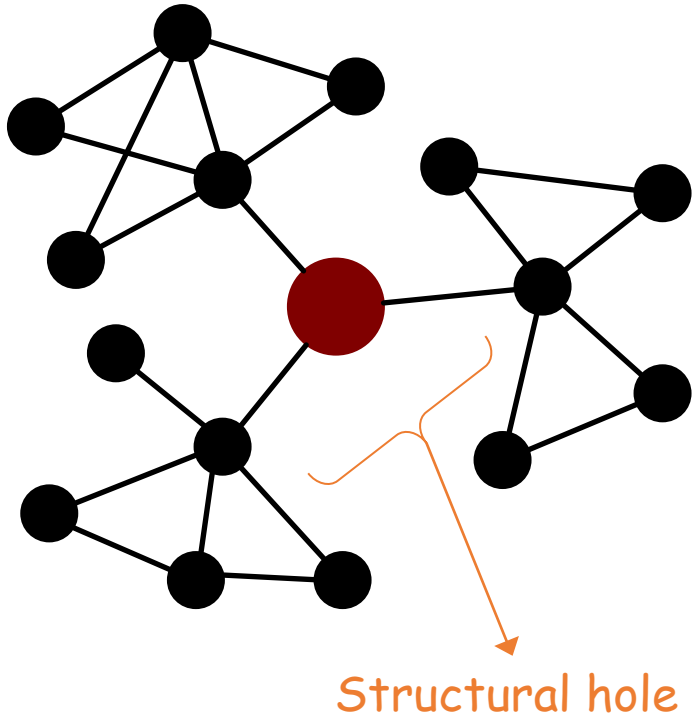
A and B have different sources of relative advantages. B spans structural holes in the network.



Who is better off Robert or James?

Structural holes

- Burt ’92 theory of individual social capital
- Structural advantage
  - Not based on the attributes of ego’s alters, but on the structure of the ego network
  - How you are connected, not who you are connected to
- Specifically, the lack of ties among alters
- Benefits (’92 version)
  - Autonomy
  - Control
  - Information



Burt study

- Managers asked to come up with an idea to improve the supply chain
- Then asked:
  - whom did you discuss the idea with?
  - whom do you discuss supply-chain issues with in general
  - do those contacts discuss ideas with one another?

  - 673 managers (455 (68%) completed the survey)
  - ~ 4000 relationships (edges)

Structural Holes (Ron Burt)

- Managers asked to come up with an idea to improve the supply chain
- Then asked:
  - whom did you discuss the idea with?
  - whom do you discuss supply-chain issues with in general
  - do those contacts discuss ideas with one another?

  - 673 managers (455 (68%) completed the survey)
  - ~ 4000 relationships (edges)

Hypotheses:

1. Opinions within groups are homogenous;
2. People who extend themselves across structural holes are exposed to new information
3. New ideas emerge from having diverse pool of options

after intervention...

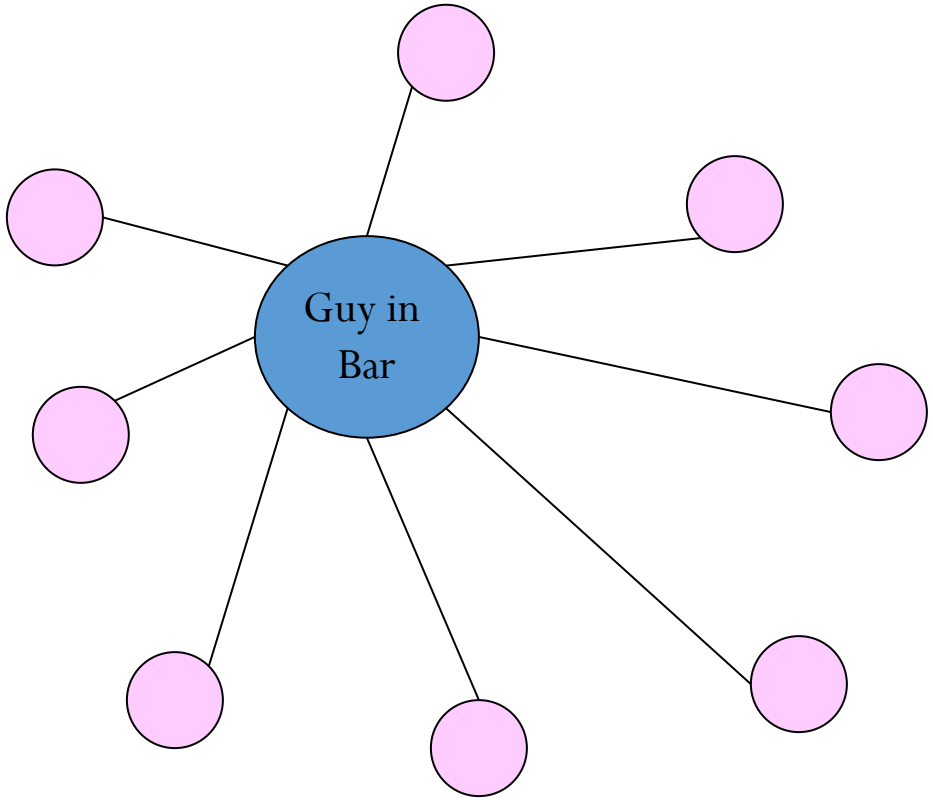
results

- people whose networks bridge structural holes have
  - higher compensation
  - positive performance evaluations
  - more promotions
  - more good ideas

- these brokers are
  - more likely to express ideas
  - less likely to have their ideas dismissed by judges
  - more likely to have their ideas evaluated as valuable

Autonomy

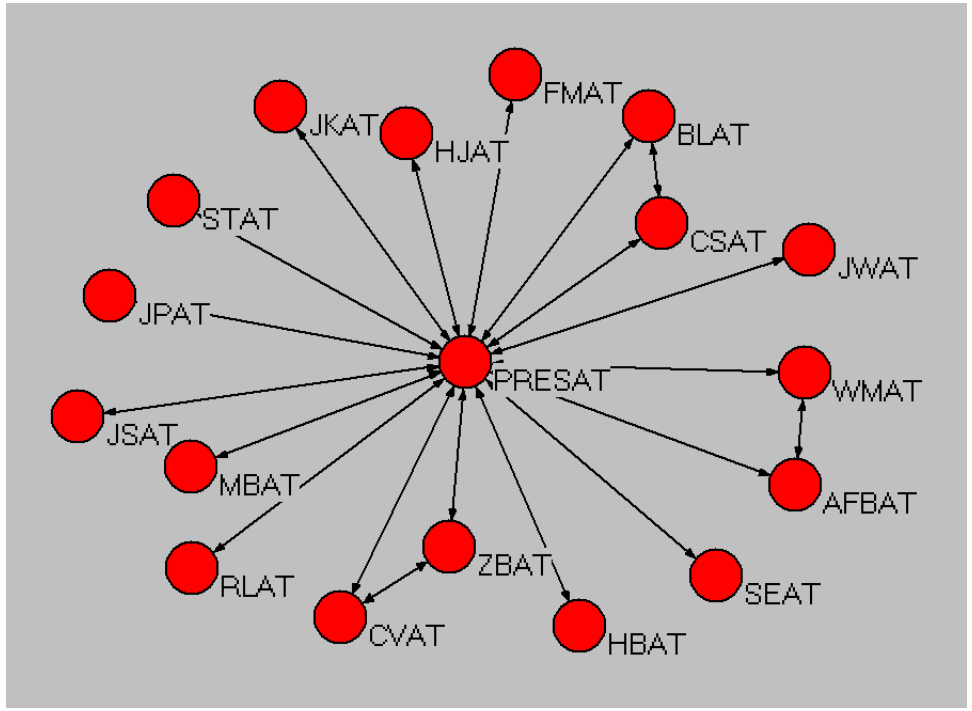
- Independence
- Freedom of action
- Fluid identity



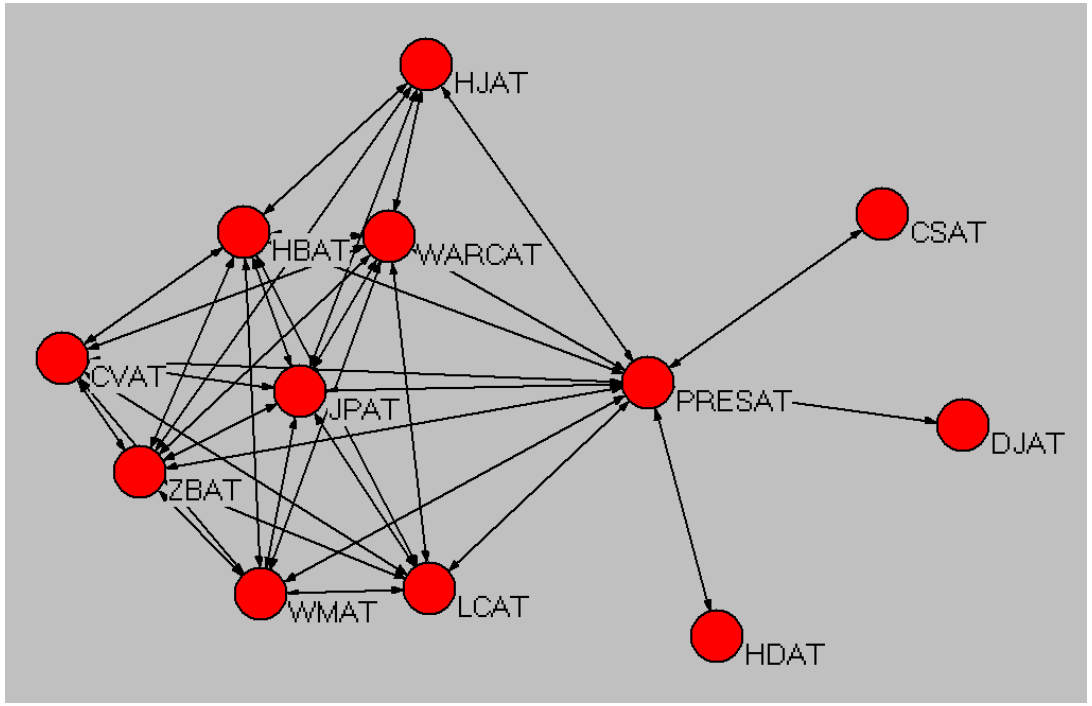
Control Benefits

- Collusion/unionizing
- Divide and conquer

White House Diary Data, Carter Presidency



Year 1

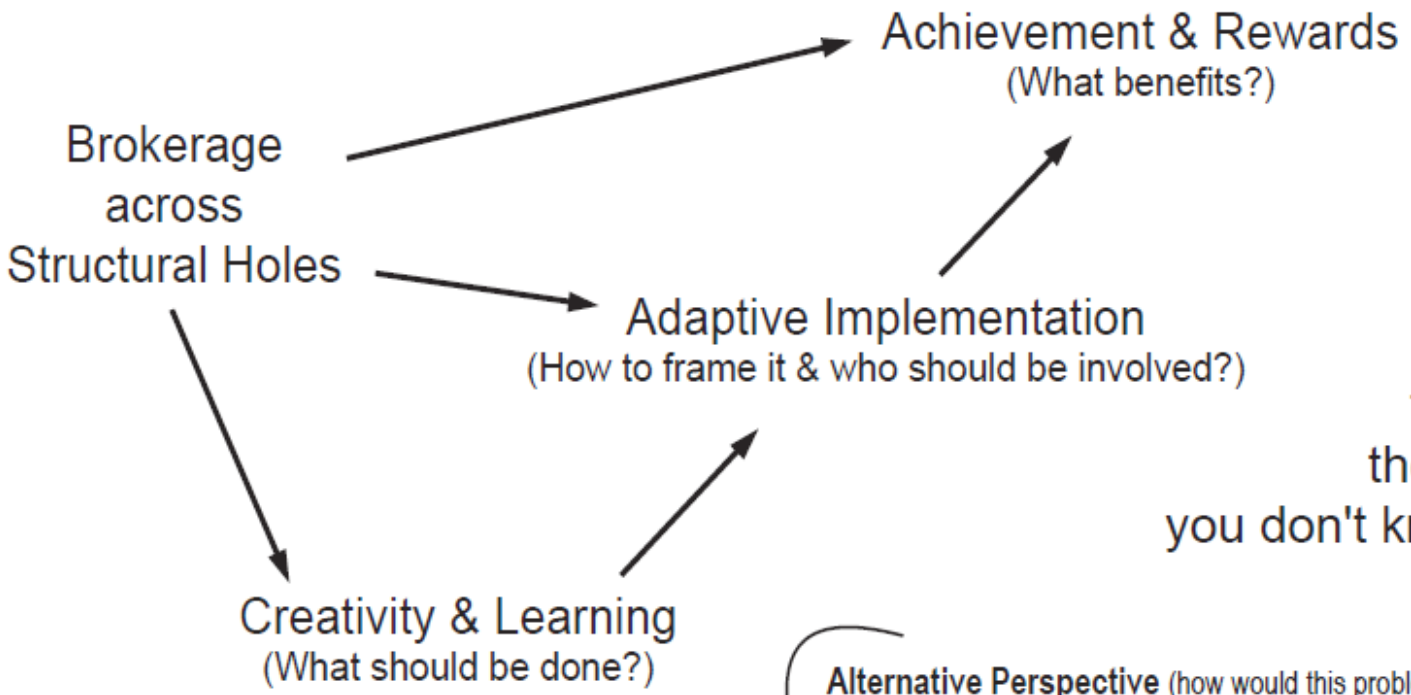


Year 4

Data courtesy of Michael Link

Information benefits

The vision advantage



What in your work improves the odds that you will discover the value of something you don't know you don't know?

Alternative Perspective (how would this problem look from the perspective of a different group, or groups — thinking "out of the box" is often less valuable than seeing the problem as it would look if you were inside a specific "other box")

Best Practice (something they think or do could be valuable in my operations)

Analogy (something about the way they think or behave has implications for how I can enhance the value of my operations; i.e., look for the value of juxtapositioning two clusters, not reasons why the two are different so as to be irrelevant to one another — you often find what you look for)

Synergy (resources in our separate operations can be combined to create a valueable new idea/practice/product)

Strategic Leadership
Network Brokerage (page 13)

CHICAGO BOOTH

from Burt, "The social capital of structural holes" (2002, *The New Economic Sociology*).

More recently proposed mechanisms

Strategic Leadership
Network Brokerage (page 24)

MECHANISM: But Personal Engagement is Essential to the Advantage

There is no advantage or disadvantage to affiliation with network brokers. Advantage comes from personal access to structural holes. Advantage does not result from exclusive access to the information of diverse contacts so much as it results from personal skills developed from translating information between diverse contacts. Brokers develop skills of analogy and metaphor for seeing and communicating across diverse ways of thinking and behaving. Network advantage affects performance less for who you know than for who you are. In other words, social capital is a forcing function for human capital, transforming network brokers into people stronger than they would otherwise be.

CHICAGO BOOTH

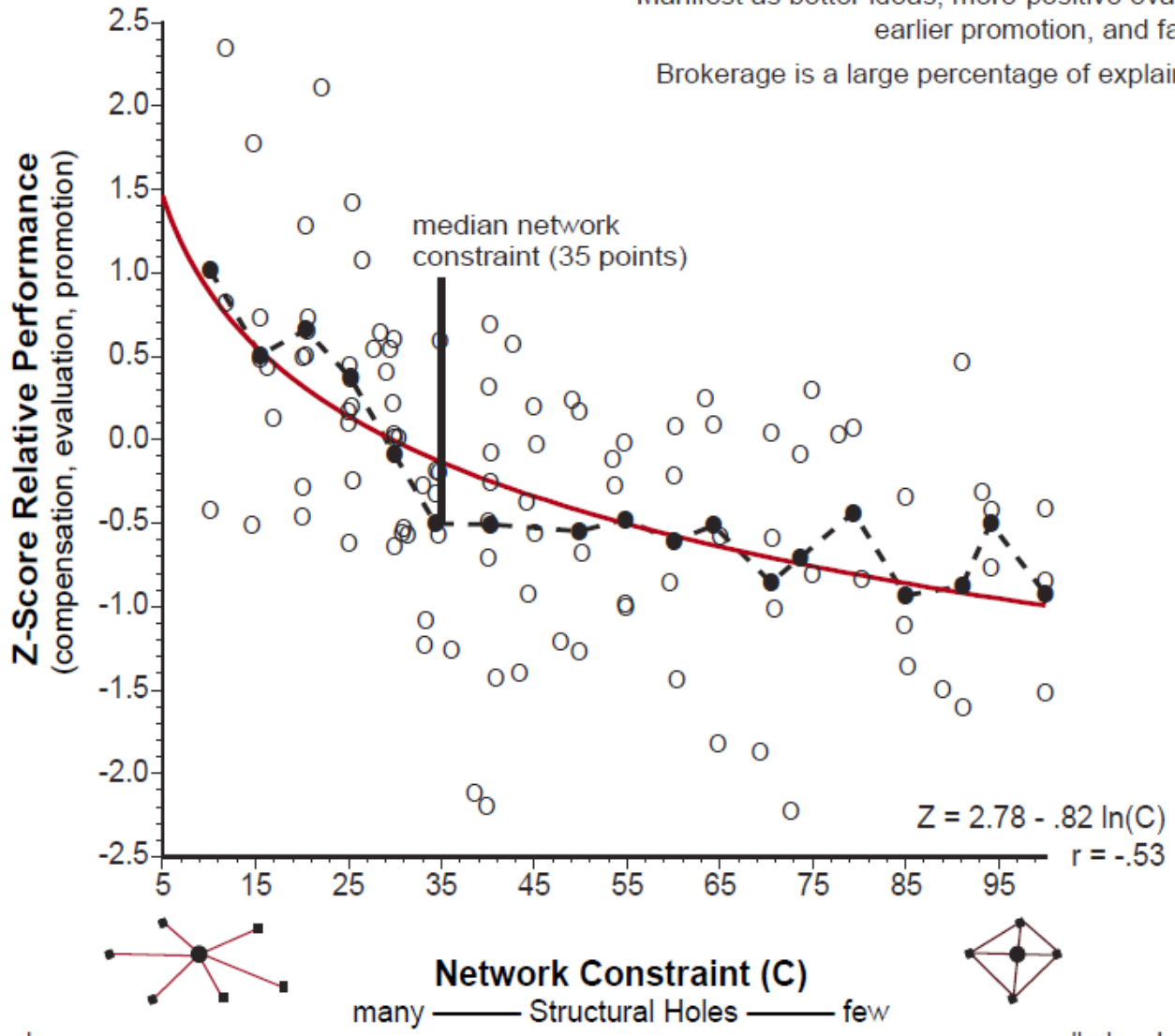
Courtesy Ron Burt

SH associated with Performance

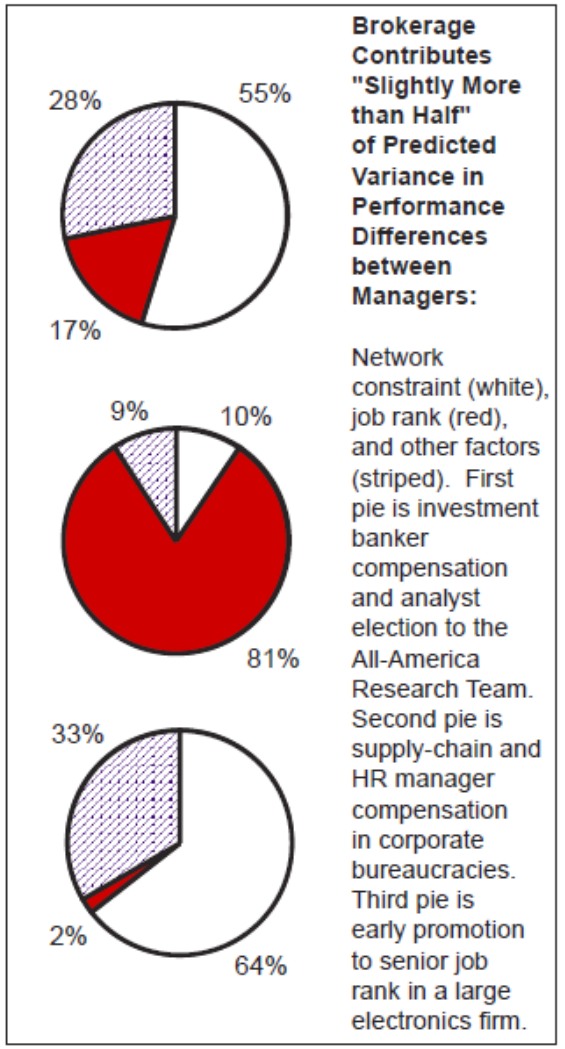
Social Capital of Brokerage

Manifest as better ideas, more-positive evaluations, higher compensation, earlier promotion, and faster teams.

Brokerage is a large percentage of explained performance differences.



Circles are average z-score performance (Z) for a five-point interval of network constraint (C) within each of eight study populations. Dashed line goes through mean values of Z for intervals of C. Bold line is performance predicted by the natural log of C.



Brokerage Contributes "Slightly More than Half" of Predicted Variance in Performance Differences between Managers:

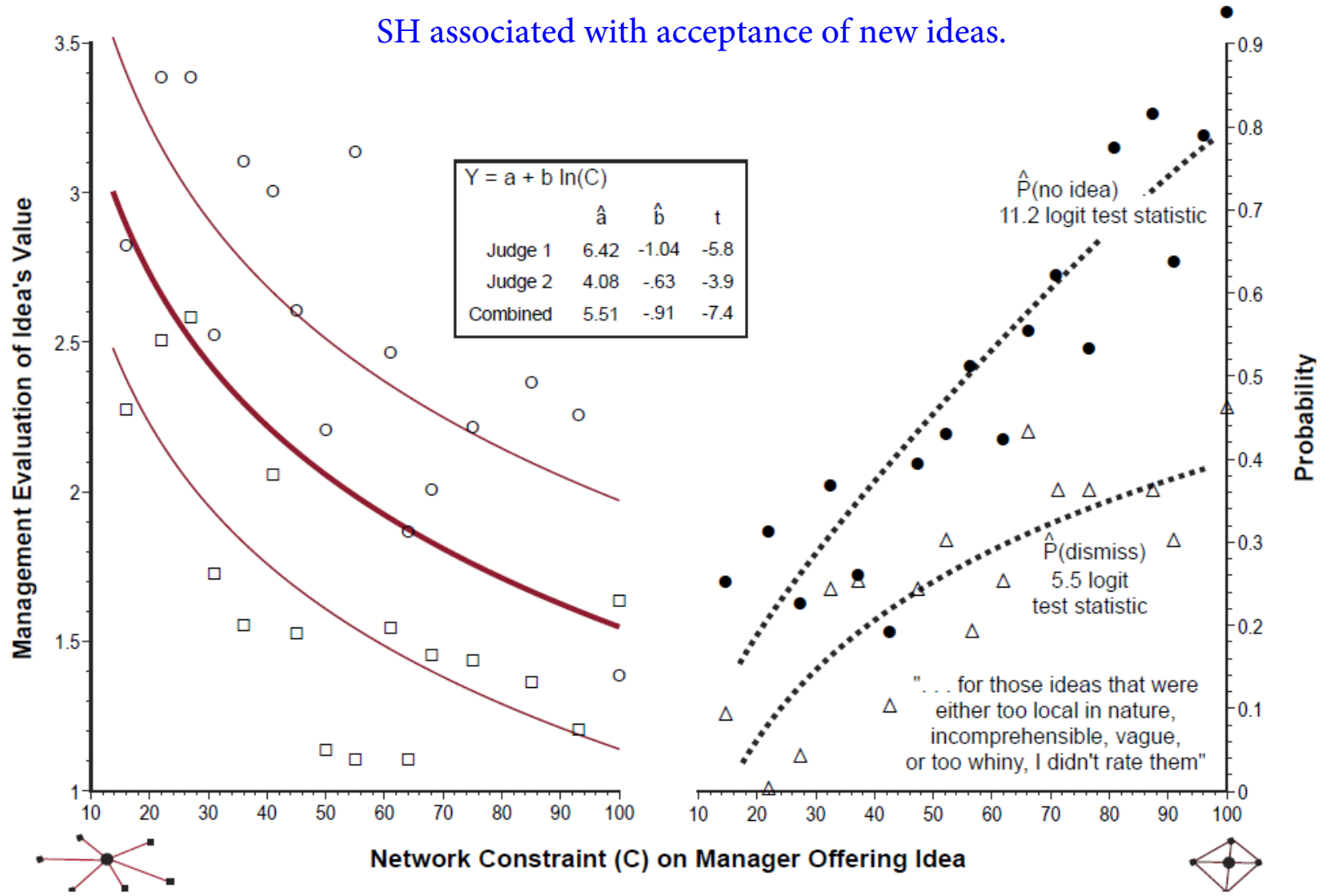
Network constraint (white), job rank (red), and other factors (striped). First pie is investment banker compensation and analyst election to the All-America Research Team. Second pie is supply-chain and HR manager compensation in corporate bureaucracies. Third pie is early promotion to senior job rank in a large electronics firm.

Graph is from Figure 1.8 in *Brokerage and Closure*. Data are pooled across eight management populations. Pie charts are from Figure 2.4 in *Neighbor Networks*.

CHICAGO BOOTH

Structural holes & good ideas

SH associated with acceptance of new ideas.

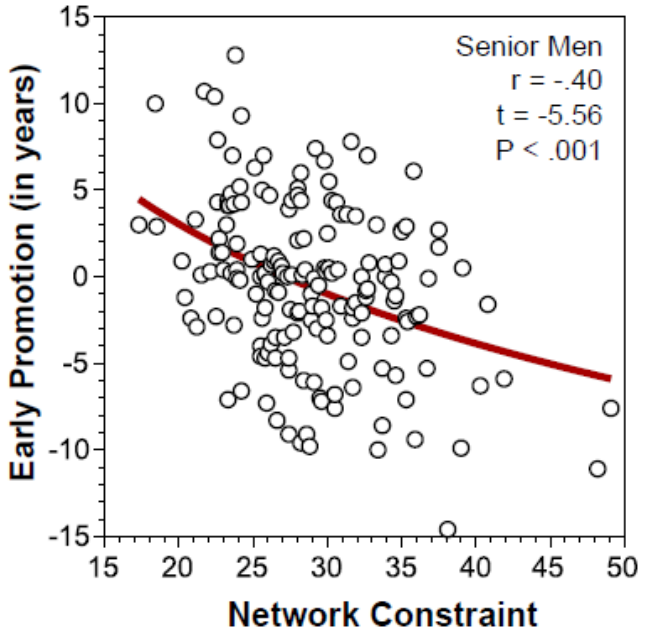


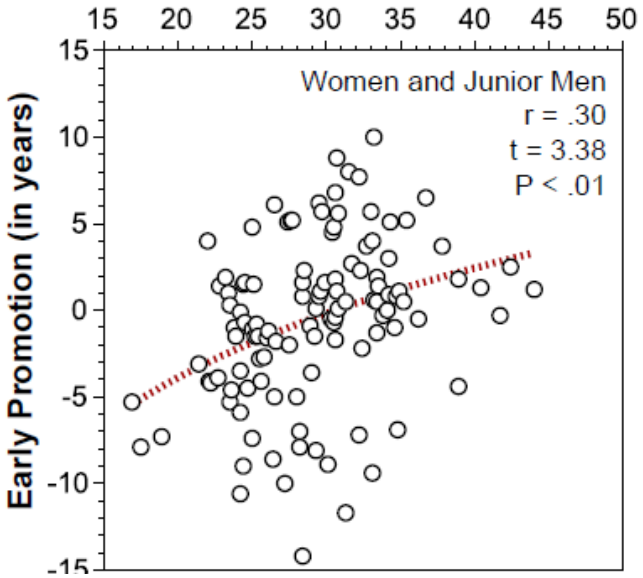
Courtesy of Ron Burt

Gender differences in benefits of social capital

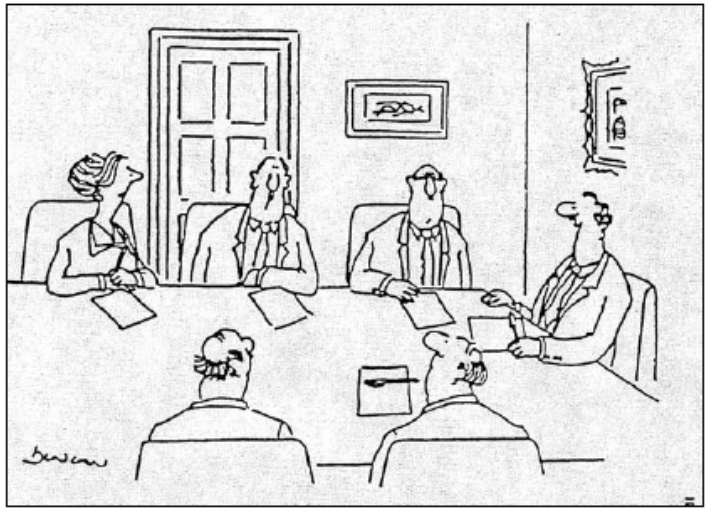
- Women and junior men don't benefit as much from structural holes
- Why?

Women & Junior need constraint to climb up promotion ladder





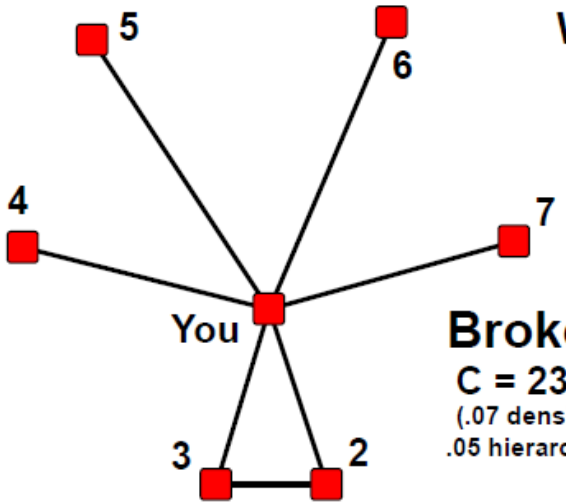
Courtesy of Ron Burt



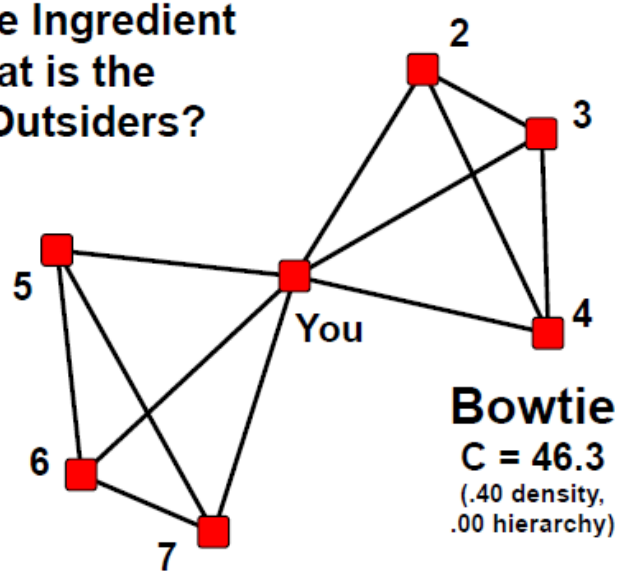
"That's an excellent suggestion, Miss Triggs. Perhaps one of the men here would like to suggest it." (Punch, 8 January, 1988)

Common Network Forms

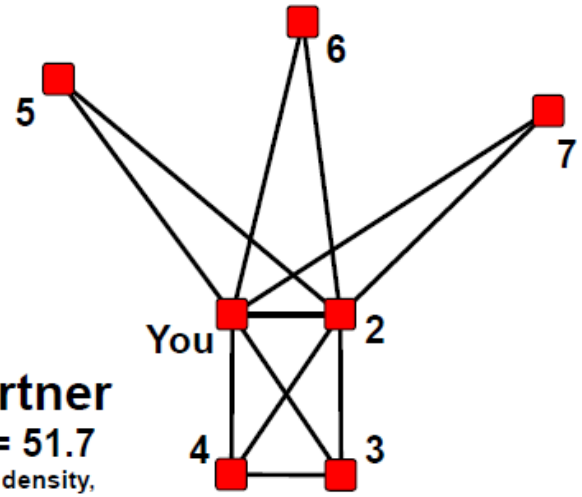
What Is the Active Ingredient in Closure that is the Advantage for Outsiders?



Broker
C = 23.6
(.07 density,
.05 hierarchy)

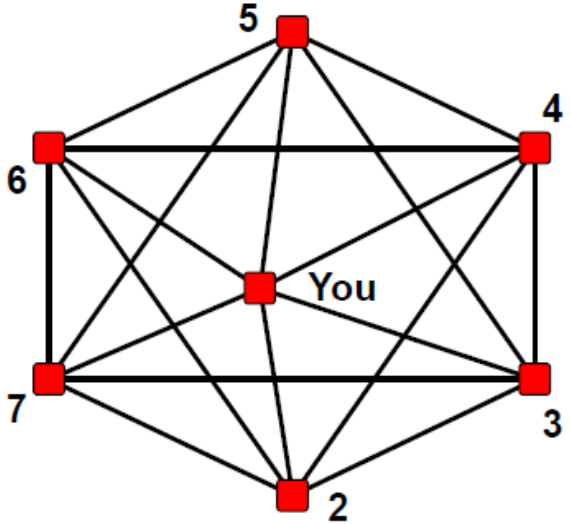


Bowtie
C = 46.3
(.40 density,
.00 hierarchy)



Partner
C = 51.7
(.40 density,
.21 hierarchy)

Imp for people who don't have their own legitimacy (e.g. women)



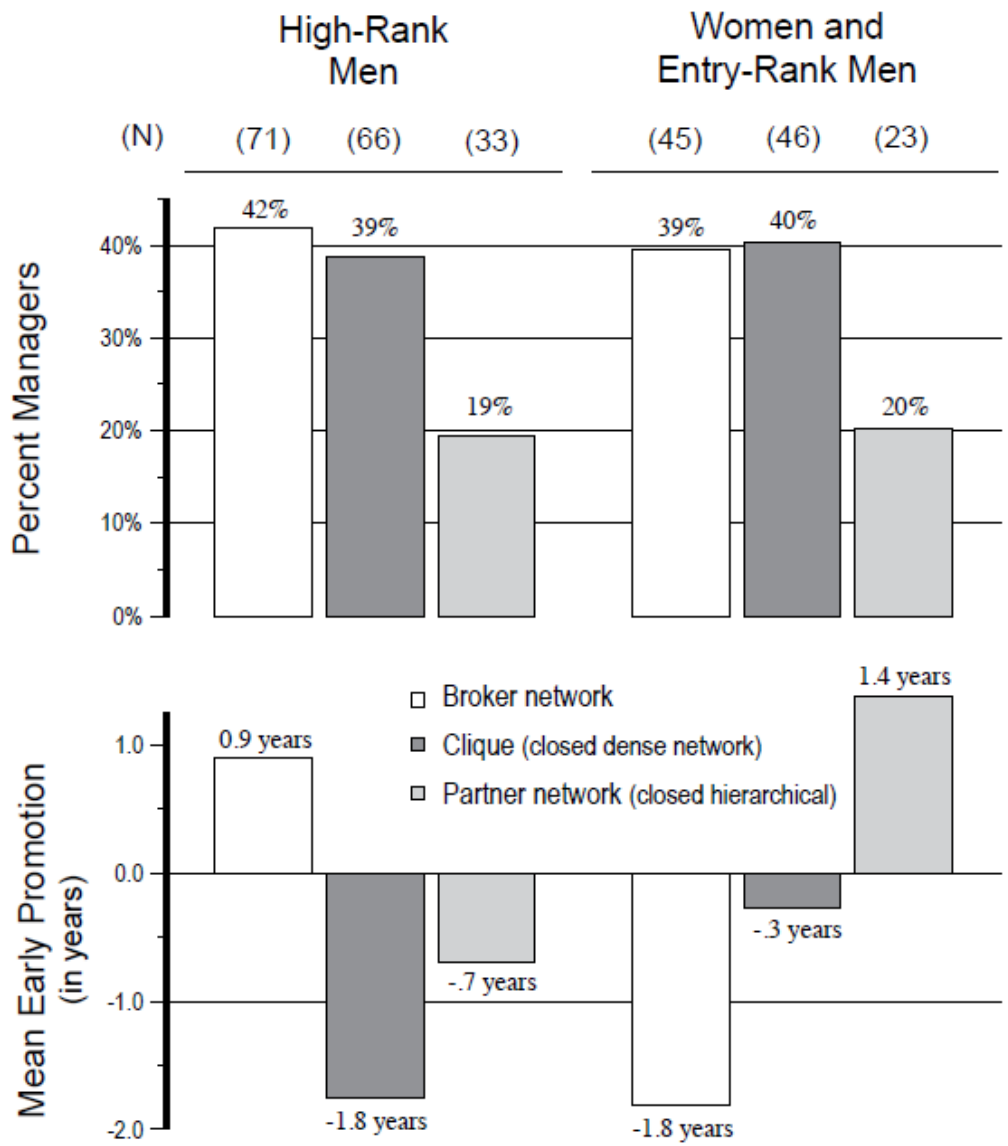
Clique
C = 54.0
(.80 density,
.00 hierarchy)

Coleman?

from Burt, "Sometimes they don't want to hear it from a person like you," (2012, *L'Impresa*)

Courtesy of Ron Burt

Partnering Is the Active Ingredient that Links Network Constraint with Success for People Excluded from Brokerage

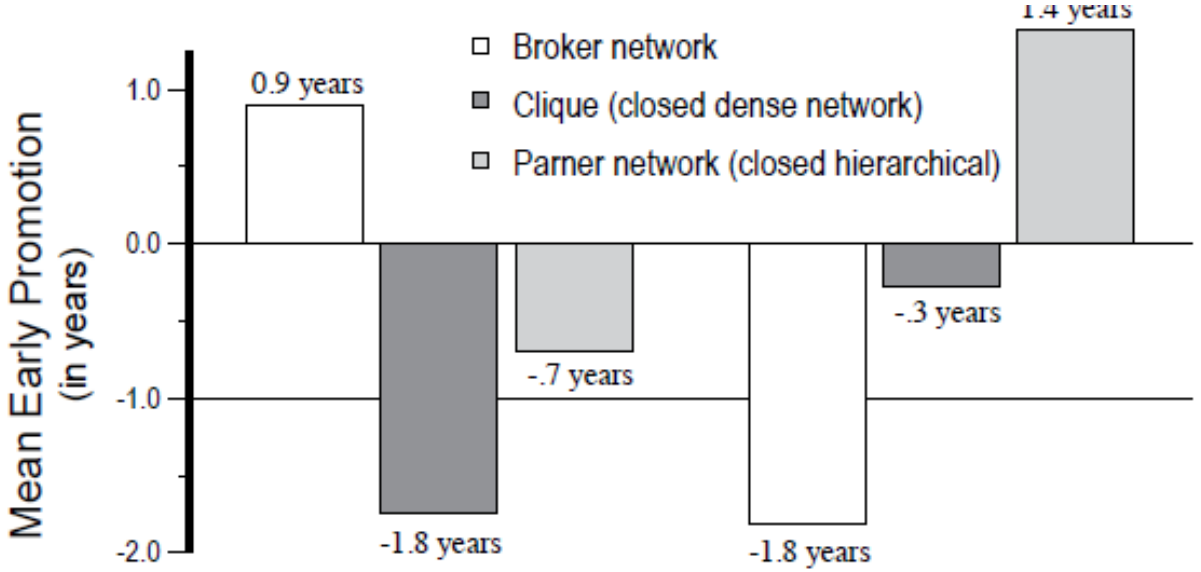


Kinds of Networks Are Similarly Likely across Kinds of Managers
($\chi^2 = 0.15$, 2 d.f., P = .93)

(In other words, pick a network for what it can do, not for the kind of people who picked it in the past.)

Kinds of Networks Have Different Consequences for Kinds of Managers
(F = 3.77, 5-278 d.f., P < .01)

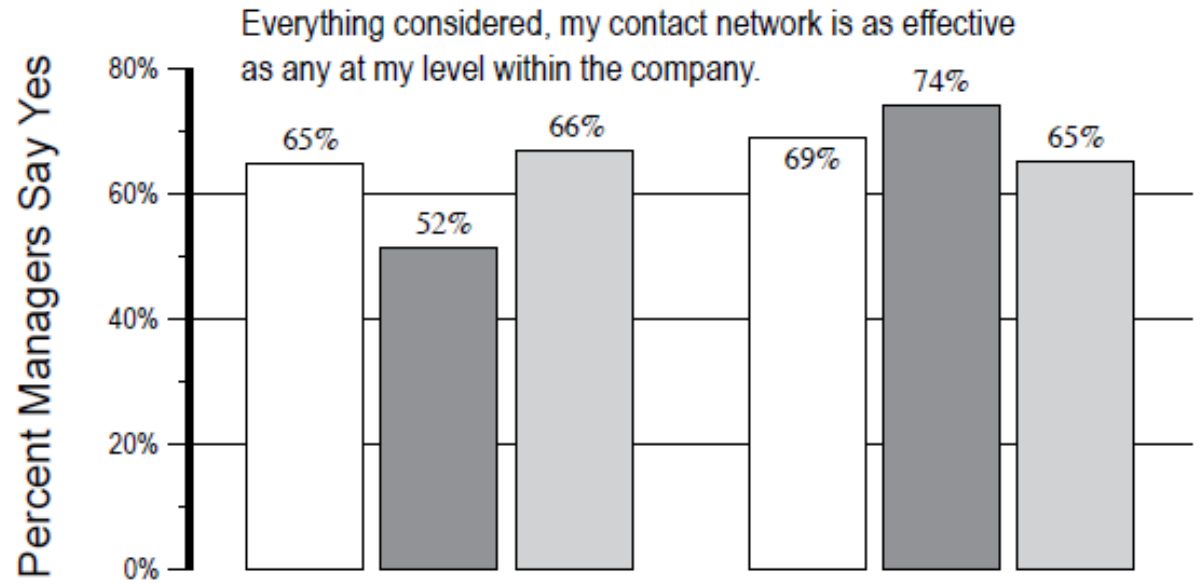
Men & Women seem to create the same type of networks around them....



....albeit with different consequences.

Kinds of Networks Have Different Consequences for Kinds of Managers

($F = 3.77$, 5-278 d.f., $P < .01$)



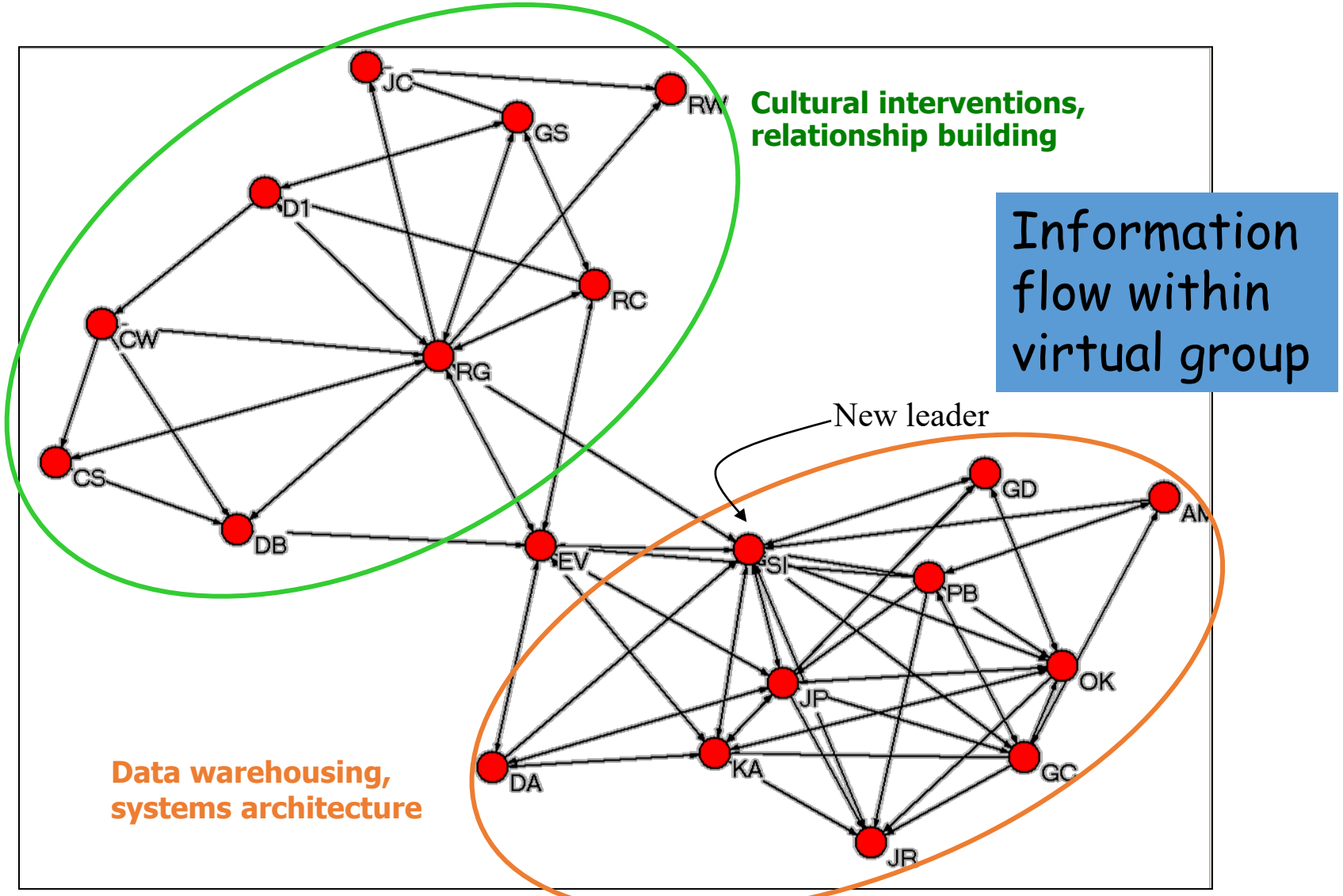
Regardless, Managers Believe That They Have an Effective Network

($\chi^2 = 6.97$, 5 d.f., $P = .22$)

(and no association between early promotion and manager's belief that his or her network is effective; 1.63 t-test, $P = .20$)

Courtesy of Ron Burt

Information & Success

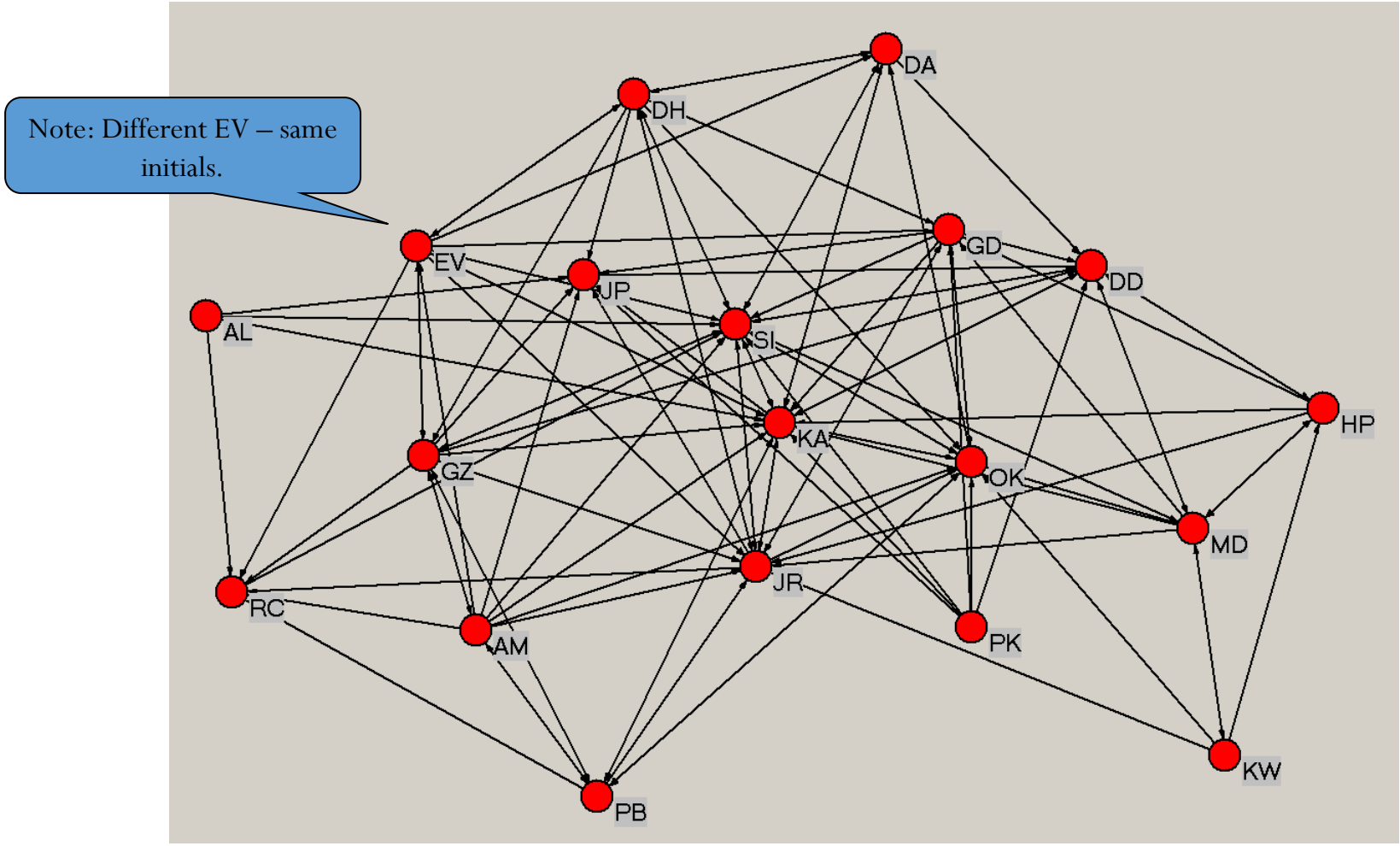


Cross, Parker, & Borgatti, 2002. Making Invisible Work Visible. *California Management Review*. 44(2): 25-46

Changes Made

- Cross-staffed new internal projects
  - white papers, database development
- Established cross-selling sales goals
  - managers accountable for selling projects with both kinds of expertise
- New communication vehicles
  - project tracking db; weekly email update
- Personnel changes

9 Months Later

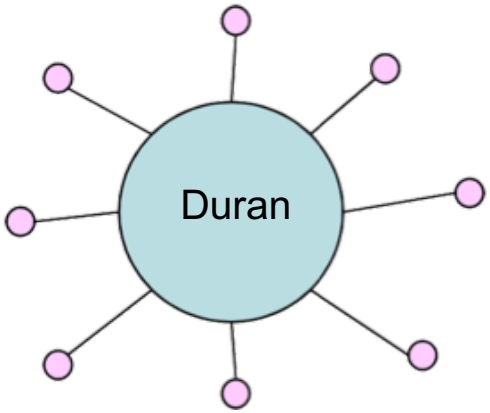


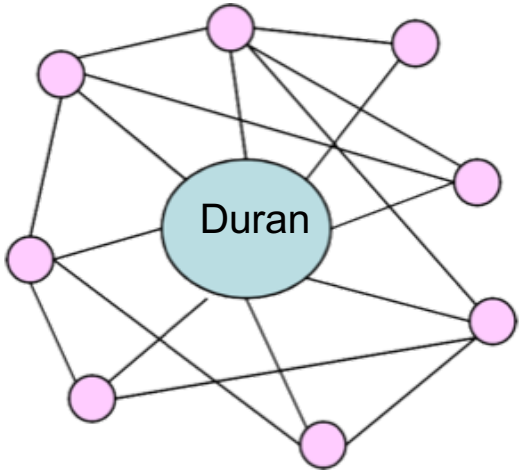
Cross, Parker, & Borgatti, 2002. Making Invisible Work Visible. *California Management Review*. 44(2): 25-46

Measures of Structural Holes

- **Effective size**;
- **Efficiency**;
- **Constraint**;
- **Hierarchy**;

Redundancy: dyadic redundancy calculates, for each actor in ego's neighborhood, how many of the other actors are also tied to the other. What % of Ego's network is redundant? Correlates with embeddedness.





Effective Size

m_{jq} = j's interaction with q divided by j's strongest relation with anyone

p_{iq} = proportion of i's energy invested in relation with q

$$ES_i = \sum_j \left[1 - \sum_q p_{iq} m_{jq} \right], \quad q \neq i, j$$

$$ES_i = \sum_j 1 - \sum_j \sum_q p_{iq} m_{jq}, \quad q \neq i, j$$

- Effective size is network size (N) minus redundancy in network

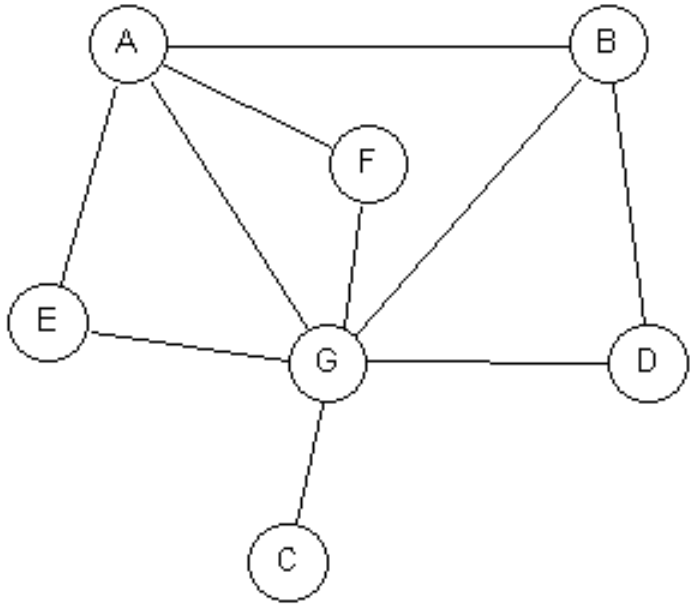


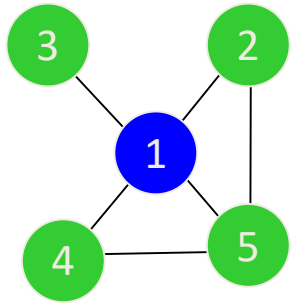
Figure 1. Adapted from Burt (1995:56)

Effective Size formula

Effective Size:

$$\sum_j \left[1 - \sum_q p_{iq} m_{jq} \right]$$

P_{iq} is the proportion of actor i's relations that are spent with q.



Adjacency

	1	2	3	4	5
1	0	1	1	1	1
2	1	0	0	0	1
3	1	0	0	0	0
4	1	0	0	0	1
5	1	1	0	1	0

P

	1	2	3	4	5
1	.00	.25	.25	.25	.25
2	.50	.00	.00	.00	.50
3	1.0	.00	.00	.00	.00
4	.50	.00	.00	.00	.50
5	.33	.33	.00	.33	.00

Effective Size formula

Effective Size:

$$\sum_j \left[1 - \sum_q p_{iq} m_{jq} \right]$$

m_{jq} is the marginal strength of contact j's relation with contact q. Which is j's interaction with q divided by j's strongest interaction with anyone. For a binary network, the strongest link is always 1 and thus m_{jq} reduces to 0 or 1 (whether j is connected to q or not - that is, the adjacency matrix).

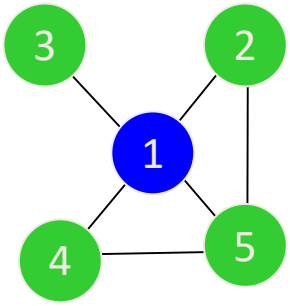
The sum of the product $p_{iq}m_{jq}$ measures the portion of i's relation with j that is redundant to i's relation with other primary contacts.

Effective Size formula

Effective Size:

$$\sum_j \left[1 - \sum_q p_{iq} m_{jq} \right]$$

Working with 1 as ego, we get the following redundancy levels:



P

	1	2	3	4	5
1	.00	.25	.25	.25	.25
2	.50	.00	.00	.00	.50
3	1.0	.00	.00	.00	.00
4	.50	.00	.00	.00	.50
5	.33	.33	.00	.33	.00

PM_{1jq}

	1	2	3	4	5
1	---	---	---	---	---
2	---	.00	.00	.00	.25
3	---	.00	.00	.00	.00
4	---	.00	.00	.00	.25
5	---	.25	.00	.25	.00

Sum=1, so

Effective size = 4-1 = 3.

Effective Size in 1/0 Data

- M_{jq} = j's interaction with q divided by j's strongest tie with anyone
  - So this is always 1 if j has tie to q and 0 otherwise
- P_{iq} = proportion of i's energy invested in relationship with q
  - So this is a constant 1/N where N is ego's network size
- Effective size reduce to network size minus the average degree of the alters **redundancy**

$$ES_i = \sum_j \left[1 - \sum_q p_{iq} m_{jq} \right], \quad q \neq i, j$$

$$ES_i = \sum_j \left[1 - \frac{1}{n} \sum_q m_{jq} \right], \quad q \neq i, j$$

$$ES_i = \sum_j 1 - \sum_j \frac{1}{n} \sum_q m_{jq}, \quad q \neq i, j$$

$$ES_i = n - \underbrace{\frac{1}{n} \sum_j \sum_q m_{jq}}, \quad q \neq i, j$$

Average degree of alters (net of ego) within the ego-network

Simplifying effective size for case of 1/0 data

- M_{jq} = j's interaction with q divided by j's strongest tie with anyone
  - So this is always 1 if j has tie to q and 0 otherwise
- P_{iq} = proportion of i's energy invested in relationship with q
  - So this is a constant 1/N where N is ego's network size
- Effective size reduces to network size minus the average degree among the alters

$$ES_i = \sum_j \left[1 - \sum_q p_{iq} m_{jq} \right], \quad q \neq i, j$$

$$ES_i = \sum_j \left[1 - \frac{1}{n} \sum_q m_{jq} \right], \quad q \neq i, j$$

$$ES_i = \sum_j 1 - \sum_j \frac{1}{n} \sum_q m_{jq}, \quad q \neq i, j$$

$$ES_i = n - \frac{1}{n} \sum_j \sum_q m_{jq}, \quad q \neq i, j$$

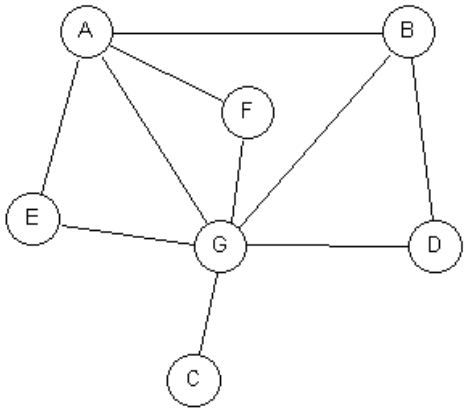


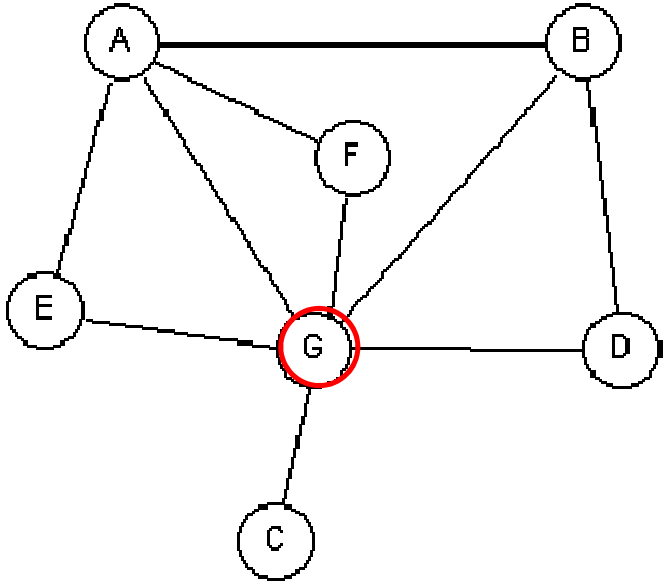
Figure 1. Adapted from Burt (1995:56)

Effective Size of G = Number of G's Alters – Sum of Redundancy of G's alters

= 6 – 1.33 = 4.67

Borgatti, S. P. 1997. Structural holes: Unpacking Burt's redundancy measures. Connections, 20(1): 35-38. [pdf]

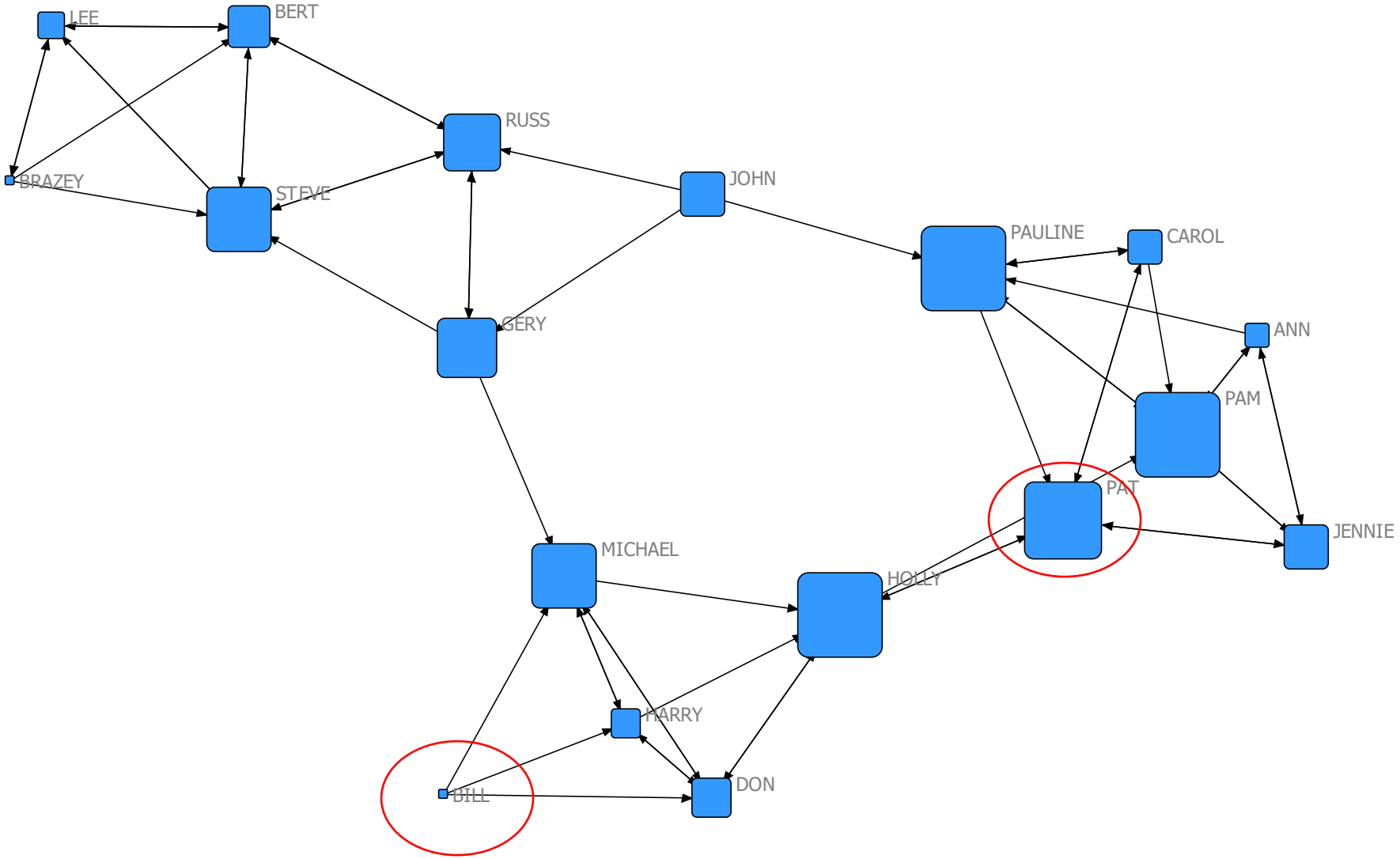
Effective Size



Node "G" is EGO	A	B	C	D	E	F	Total
Redundancy with EGO's other Alters:	3/6	2/6	0/6	1/6	1/6	1/6	1.33

Effective Size of G = Number of G's Alters – Sum of Redundancy of G's alters
= 6 – 1.33 = **4.67**

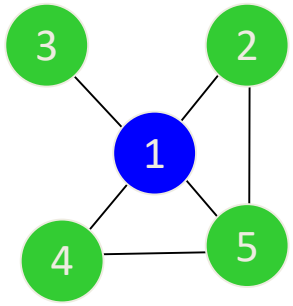
Sized by Effective Size



Efficiency

Efficiency is the observed size divided by the effective size:

degree/effective size



Node	Size	Effective Size:	Efficiency
1	4	3	.75
2	2	1	.5
3	1	1	1.0
4	2	1	.5
5	3	1.67	.55

Constraint

Mjq= j's interaction with q divided by j's strongest relationship with anyone

So this is always 1 if j has tie to q and 0 otherwise

P_{iq} = proportion of i's energy invested in relationship with q

So this is a constant 1/N where N is network size

$p_{iq} = 1/k_i$, where k_i is ego degree

$$c_{ij} = \left(p_{ij} + \sum_q p_{iq} p_{qj} \right)^2, \quad q \neq i, j$$

extent to which EGO is invested in alter q who is invested in alter j

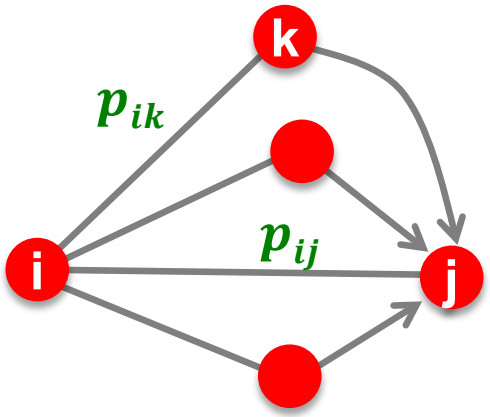
$$c_i = \sum_j c_{ij}$$

inverse measure of SH: small Sum = big SH

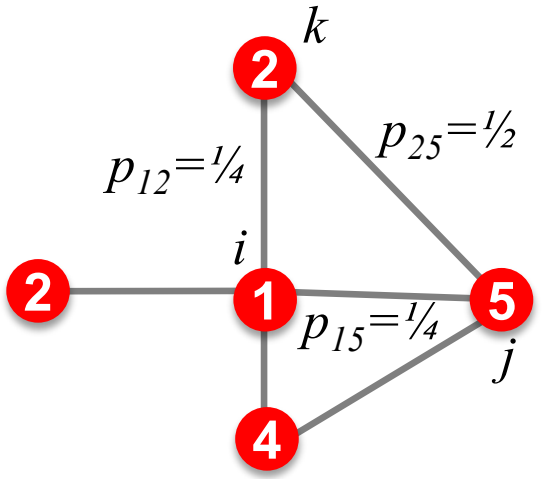
- Reverse-coded measure of structural holes: large values = fewer holes
- Alter j constrains i to the extent that
  - i has invested in j
  - i has invested in people (q) who have invested heavily in j. That is, i's investment in q leads back to j.
- Even if i withdraws from j, others in i's network still invested in j
- Overall constraint is the extent to which i is invested in those that other contacts of i are invested in
- **Father-in-law effect**: i has many friends, but they're all friend's with i's father-in-law

Constraint – formula

- **To what extent are person's contacts redundant**



$$p_{uv} = 1/d_u$$



p_{uv}

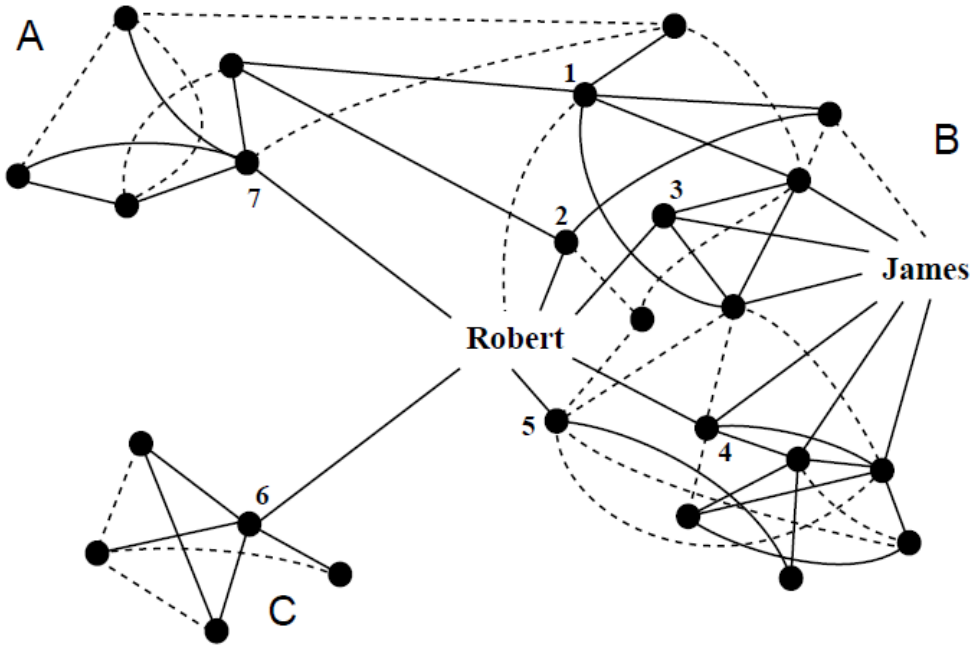
	1	2	3	4	5
1	.00	.25	.25	.25	.25
2	.50	.00	.00	.00	.50
3	1.0	.00	.00	.00	.00
4	.50	.00	.00	.00	.50
5	.33	.33	.00	.33	.00

- **Low**: disconnected contacts
- **High**: contacts that are close or strongly tied

$$c_i = \sum_j c_{ij} = \sum_j \left[p_{ij} + \sum_k \left(p_{ik} p_{kj} \right) \right]^2$$

p_{uv} prop. of u's "energy" invested in relationship with v

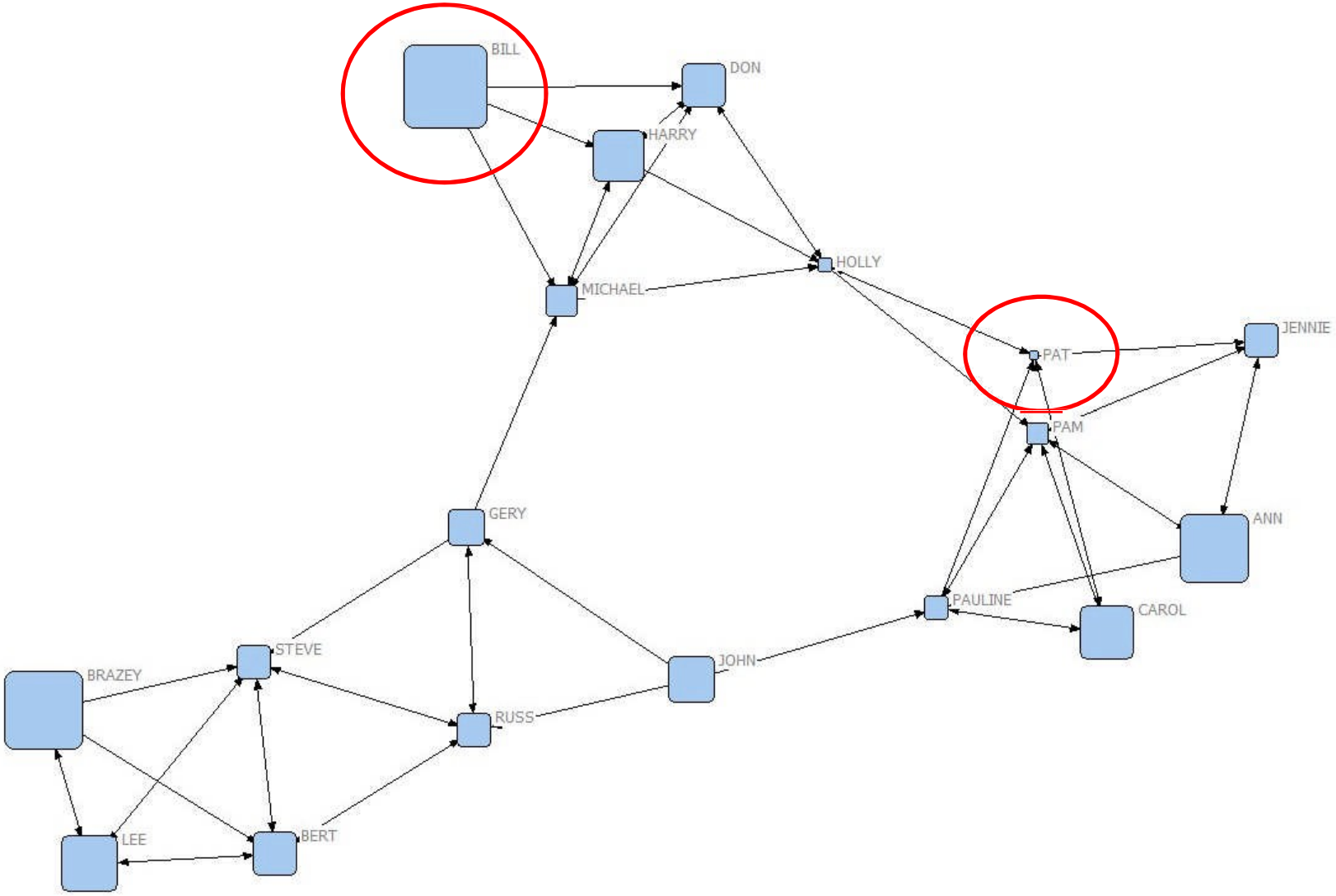
Constraint – formula



- **Network constraint:**
  - James: $c_J = 0.309$
  - Robert: $c_R = 0.148$

- **Constraint:** To what extent are person's contacts redundant
  - **Low:** disconnected contacts
  - **High:** contacts that are close or strongly tied

Sized by Constraint

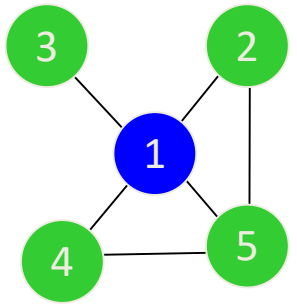


Hierarchy

Conceptually, hierarchy (for Burt) is really the extent to which **constraint** is concentrated in a single actor. It is calculated as:

$$H = \frac{\sum_j \left(\frac{C_{ij}}{C/N} \right) \ln \left(\frac{C_{ij}}{C/N} \right)}{N \ln(N)}$$

	2	3	4	5	C
C:	**0.11**	**0.06**	**0.11**	**.25**	**0.53**
$\left(\frac{C_{ij}}{C/N} \right)$	**.83**	**.46**	**.83**	**1.9**	



H=.514

SUMMARY

Name:	Description:	Relation to Social Capital:
Effective Size (Burt, 1992)	The number of alters, weighted by strength of tie, that an ego is directly connected to, minus a "redundancy" factor.	Positive. The more different regions of the network an actor has ties with, the greater the potential information and control benefits.
Constraint (Burt,1992)	The extent to which all of ego's relational investments directly or indirectly involve a single alter	Negative. The more constrained the actor, the fewer opportunities for action.
Compositional Quality (e.g., Lin)	The number of alters with high levels of needed characteristics (e.g., total wealth or power or expertise or generosity of alters)	Positive. The more connected to useful others, the more social capital.
Heterogeneity (e.g., Burt, 1983)	The variety of alters with respect to relevant dimensions (e.g., sex, age, race, occupation, talents).	Positive (except when it conflicts with compositional quality)
Brokerage Roles (Gould & Fernandez, 1989)	There are different roles that ego can play depending on network structure and composition	Depends on the situation

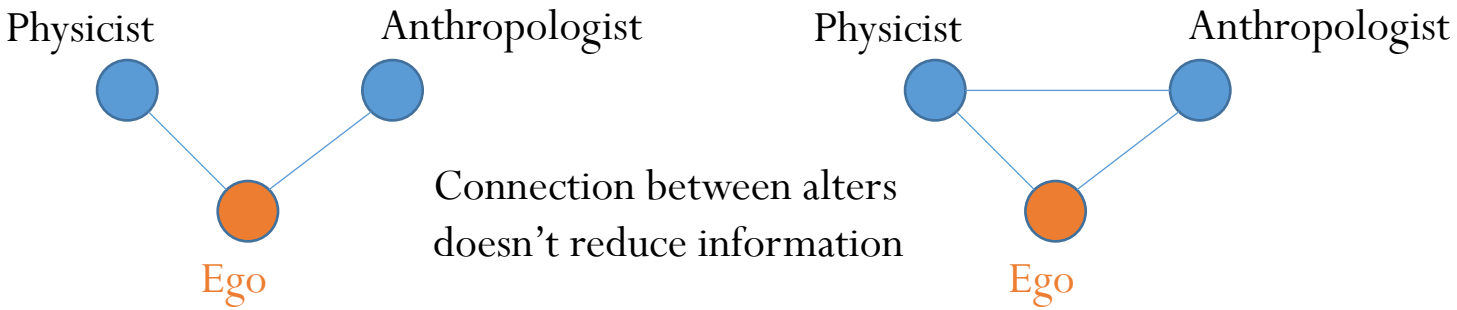
Controlling for size

- Should one control for degree when using measures of structural holes?
- Burt designed structural holes to measure social capital, and therefore included elements of degree and density – both contribute to social capital
  - access to novel information
- Controlling for degree reduces structural hole measures to something like density
  - Can’t call the result structural holes
- Better idea is to directly include degree and density as separate variables and let regression sort out relative importance

information-based

Deriving value from structural holes

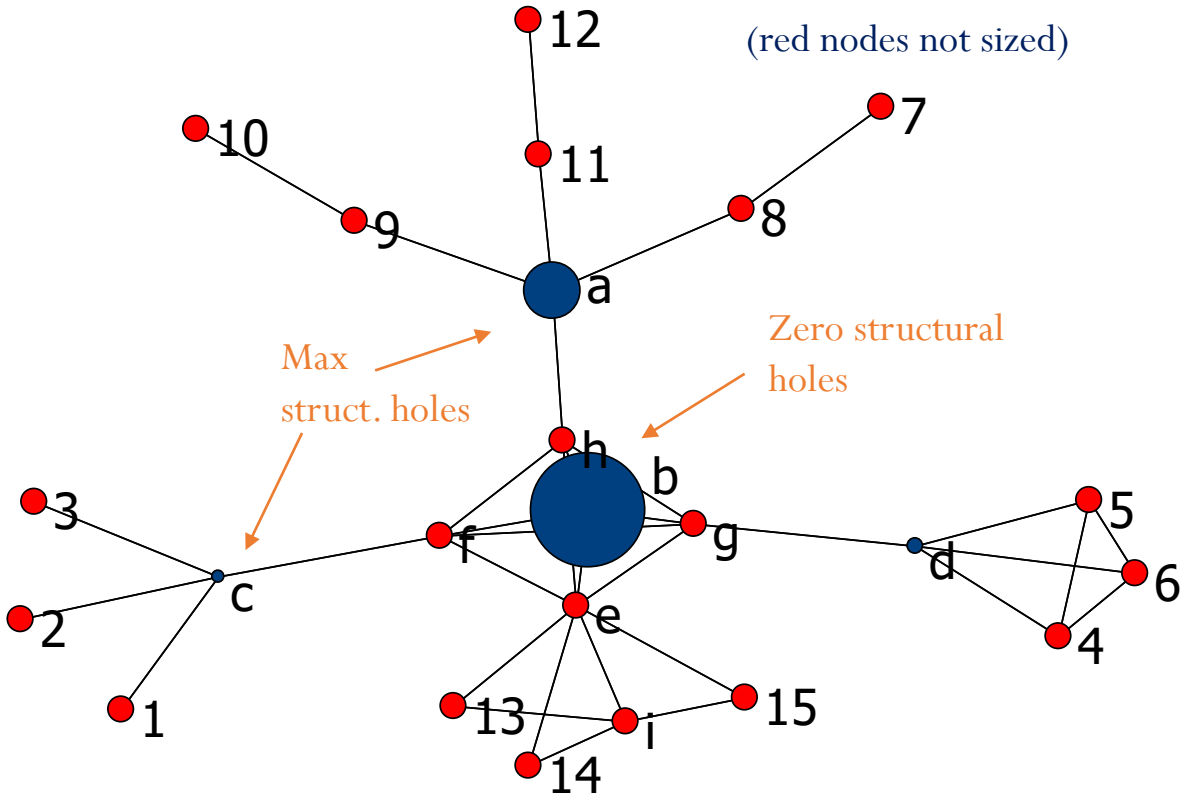
- Synthesizing solutions to problems by combining heterogeneous knowledge sources
  - If it is fundamental disciplinary knowledge, why do we care about structural holes?



  - If it is news and sense-making kind of opinion that is circulating, then ties among alters reduce the range of novel points of view available to ego
    - But perhaps the SH of my alters can make up for that

Simulation of gossip process (centrality v SH)

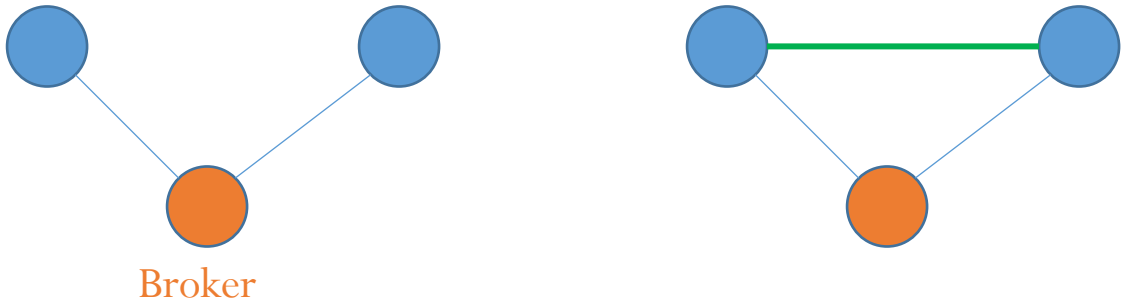
- Let news originate at random nodes and spread probabilistically to adjacency nodes over time
- At some time t measure number of unique bits of information that node has received
  - Will nodes with more holes have more info?
    - Only for very small t (few iterations)
    - For large t, it is position in global network that matters



Nodes a, b, c, d sized in proportion to amt of info

SH & actor agency

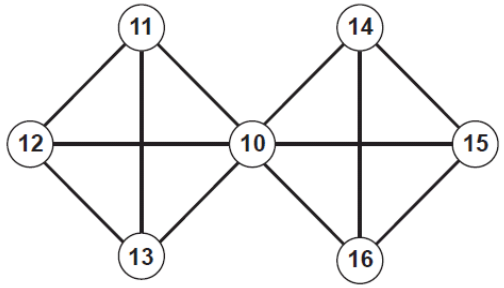
- Sometimes the value is extracted by **closing the hole**
  - Marriage broker; real-estate brokers; lawyers
  - Typically, the tie to be created between alters is a different kind of tie. Often an event, such as a sale



- It requires one to know they can close a structural hole and capitalize their position

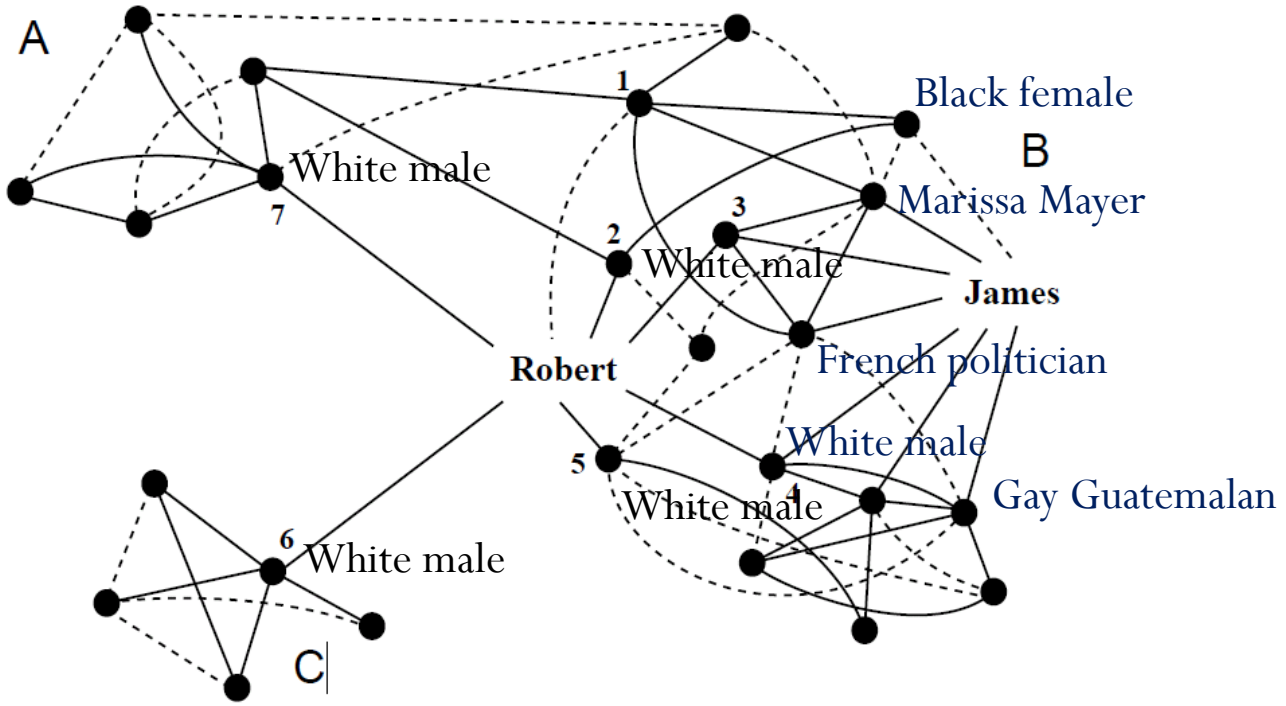
Do actors need to be aware of structural holes to benefit from them?

- For receiving non-redundant information, no
  - Whether I know it or not, I am positioned to receive non-redundant info
  - But, in terms of capitalizing on those information benefits it might help to recognize that your group 1 friends have solutions that group 2 doesn't
- For **control benefits**, yes. Knowledge of network seems more important, but can still benefit passively

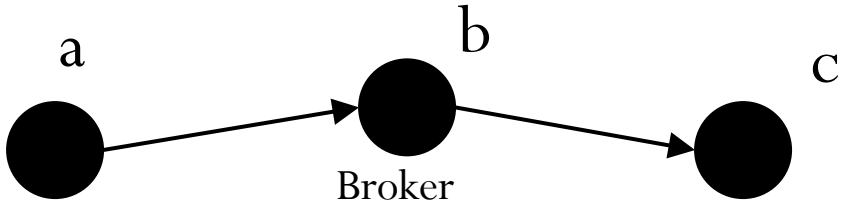


Combining Lin and Burt

- Perhaps should invent measure that takes into account both the diversity of alters and their lack of connection to each other

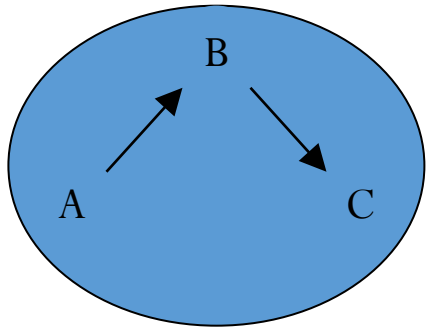


Gould & Fernandez brokerage roles

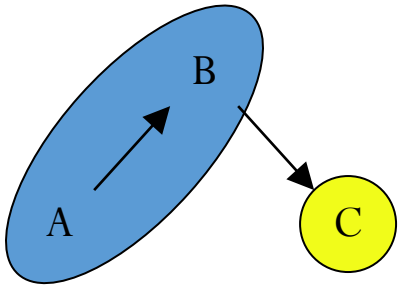


- Gould & Fernandez (1989)
- Broker is middle node of directed triad (note: a is NOT connected to c)
- What if nodes belong to different groups?
  - Categorical node attribute such as dept, ethnic group

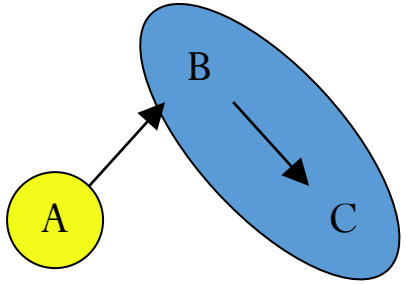
G&F Brokerage Roles



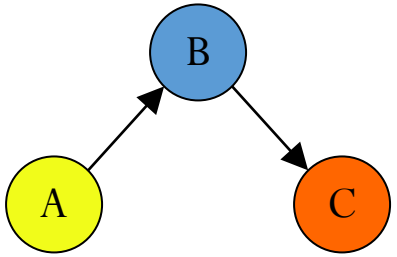
Coordinator



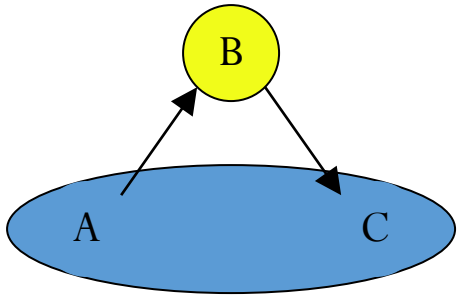
Representative



Gatekeeper



Liaison

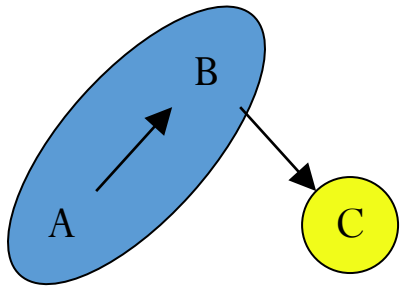


Consultant

- We can count how often a node enacts each kind of brokerage role

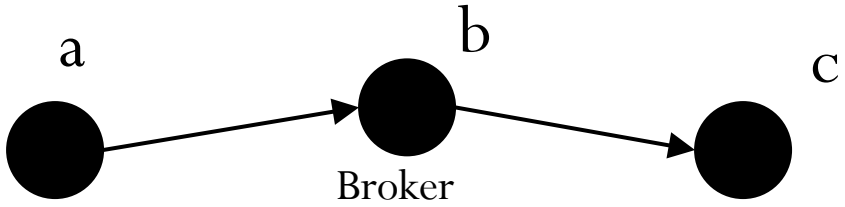
Caveats on G & F brokerage roles

- Just because B is in the structural position to be a representative doesn't mean she ever does the associated behaviors
  - Is the tie "gives information to" or something irrelevant like "likes"?
  - Even if tie is gives info to, no evidence that what A gives B is the same thing that B gives C



Representative

Gould & Fernandez brokerage roles

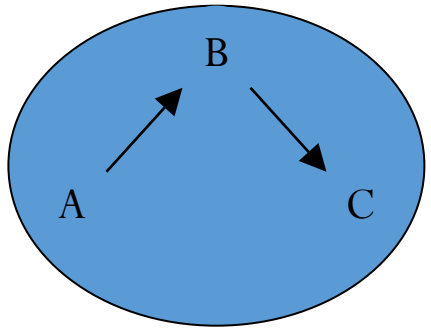


- Gould & Fernandez (1989)
- Broker is middle node of directed shortest path of length 2
  - a not directly connected to c

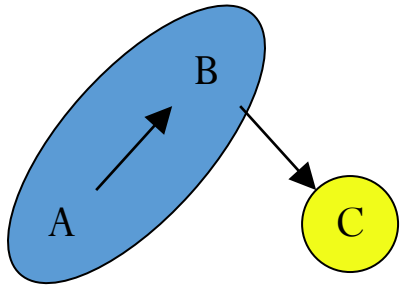
Algorithm:

- List all open paths of length 2 (A→B→C)
- If Cat(A) = Cat(B) = Cat(C), add to coordinator count
- If Cat(A) = Cat(C) != Cat(B), add to consultant count
- Etc.

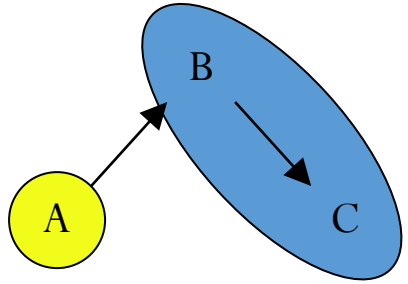
G&F Brokerage Roles



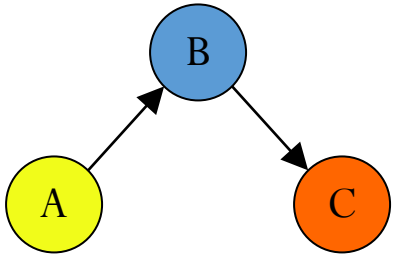
Coordinator



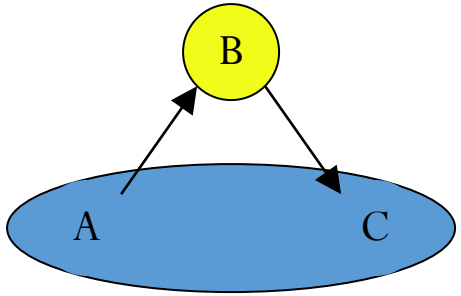
Representative



Gatekeeper



Liaison



Consultant

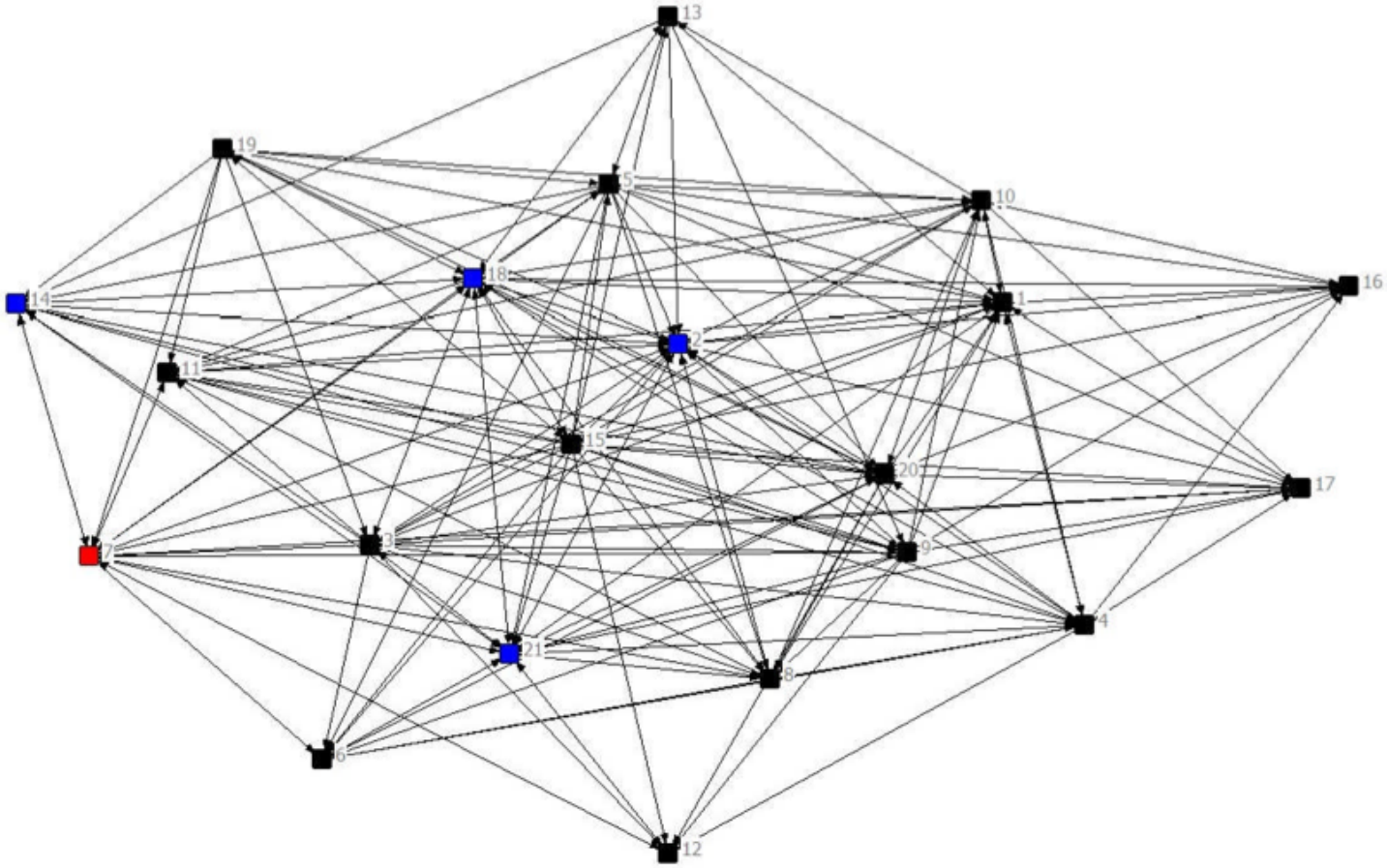
- We can count how often a node enacts each kind of brokerage role

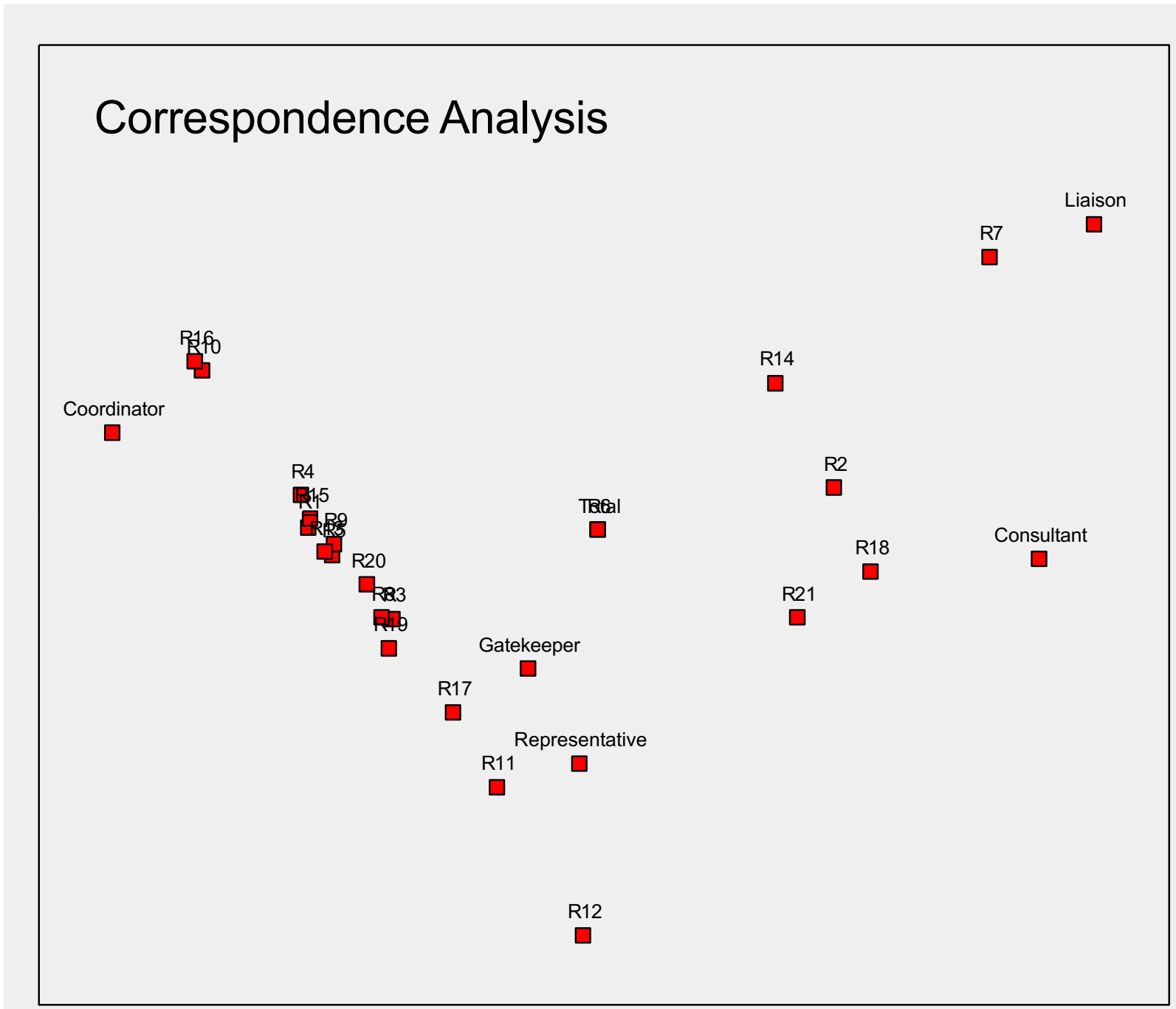
Counting of Role Structures

ID	Coordinator	Gatekeeper	Representative	Consultant	Liaison
7 (CEO)	0	0	0	17	21
21(Mgr)	2	11	16	35	8
18(Mgr)	0	9	22	72	18
14(Mgr)	0	2	0	0	2
2	0	5	2	7	6
6	0	0	0	0	0
5	14	2	6	0	0
3	9	7	4	0	0
9	8	3	2	0	0
10	44	1	0	0	0
1	17	0	7	0	0
12	0	0	2	0	0
13	2	0	1	0	0
4	21	7	2	0	0
15	18	3	5	0	0
16	2	0	0	0	0
17	3	3	4	0	0
8	8	3	5	0	0
19	2	0	2	0	0
20	12	7	4	0	0
11	1	1	3	0	0

Advice Network:

Nodes Colored by Level (CEO / Manager/ Line Staff)




Correspondence Analysis
Liaison
R7
R16
R10
R14
Coordinator
R2
R4
R15
R9
R20
Consultant
R18
R21
R19
Gatekeeper
R17
Representative
R11
R12