What is a “Good” Social Network for a System?: Knowledge Flow and Organizational Change

*Kenneth Frank
Michigan State University
*Ann Krause
University of Toledo
Bill Penuel
SRI International
*equal co-authors, authors listed alphabetically

AEFP San Antonio, March 2014
Resubmitted to Journal of Policy Management
Work supported by NSF: Human Social Dynamics program

Discussant comments

- Bi-directionality (not for me)
- Formal vs informal: paper with Jim
- Interaction between concentration of ties within clusters and effect of entropy
- Weak ties?
- Flow through formal, all faculty pd?
- Explain frequency change outcome
- Causal language?

acknowledgements

This work has been supported by National Science Foundation grants #0231981 and #0624307. All opinions expressed herein are the sole responsibility of the authors. We wish to thank core members of the data collection and analysis team for their efforts in making these analyses possible: Christine Korbak, Judi Fusco, Christopher Hoadley, Joel Galbraith, Amy Hafter, Aasha Joshi, Amy Lewis, Margaret Riel, Willow Sussex, and Devin Vodicka. The authors thank Federico Aime, Ebony Bridwell-Mitchell, Spiro Maroulis, Ray Reagans, Brian Uzzi; Arthur Stinchombe, Brian Uzzi, and Ezra Zuckerman for their thoughtful comments on earlier drafts. Draft, do not cite without permission of authors

Abstract

This study concerns how intra-organizational networks affect the implementation of policies and practices in organizations. In particular, we attend to the role of the informal subgroup or clique in cultivating and distributing locally adapted and integrated knowledge, or know-how. We develop two hypotheses based on the importance of intra-organizational coordination for an organization's capacity for change. The first emphasizes the importance of distributing know-how evenly *to* potential recipient subgroups. The second emphasizes the importance of restricting know-how to flow from high know-how subgroups. We test our hypotheses with longitudinal network data in 21 schools, finding stronger support for the second hypothesis than the first. Our findings can help managers cultivate know-how flows to contribute to organizational change.

Study Context: Resource Flows and Organizational Change

- Resource flows critical for organizations
  - overall performance
  - product innovation
  - transfer of best practices
  - workplace learning
  - Coordination: relates to stratification
- Limitation:
  - Many focus on aggregate culture, not specific flows through networks

Specific Focus on Implementation of New Initiatives in Schools

- Challenging problem reformers and theorists
- Consistent focus of organizational theory
  - control theory
  - contingency theory
  - new institutionalism

Baseline Assumption: Teachers Need Local Knowledge to be Effective for Complex Production

- Must adapt external, general knowledge to context of the school
- Local knowledge allows teachers to comply with local norms (Kennedy, 2002, calls craft knowledge)
- Local knowledge (if made explicit) can be shared with others to improve school
  - Frank et al., 2011

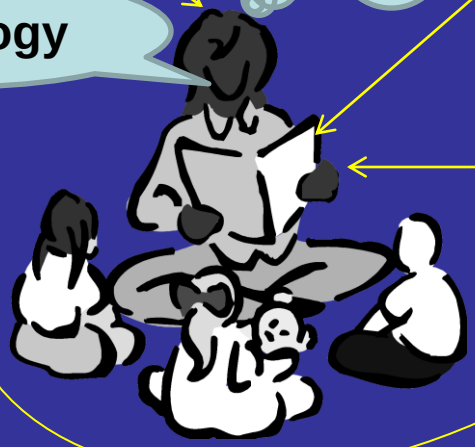
The Workforce: School Teachers

Teaching as Complex Production

Previous Training

Curriculum

pedagogy



Coordination

Assessments

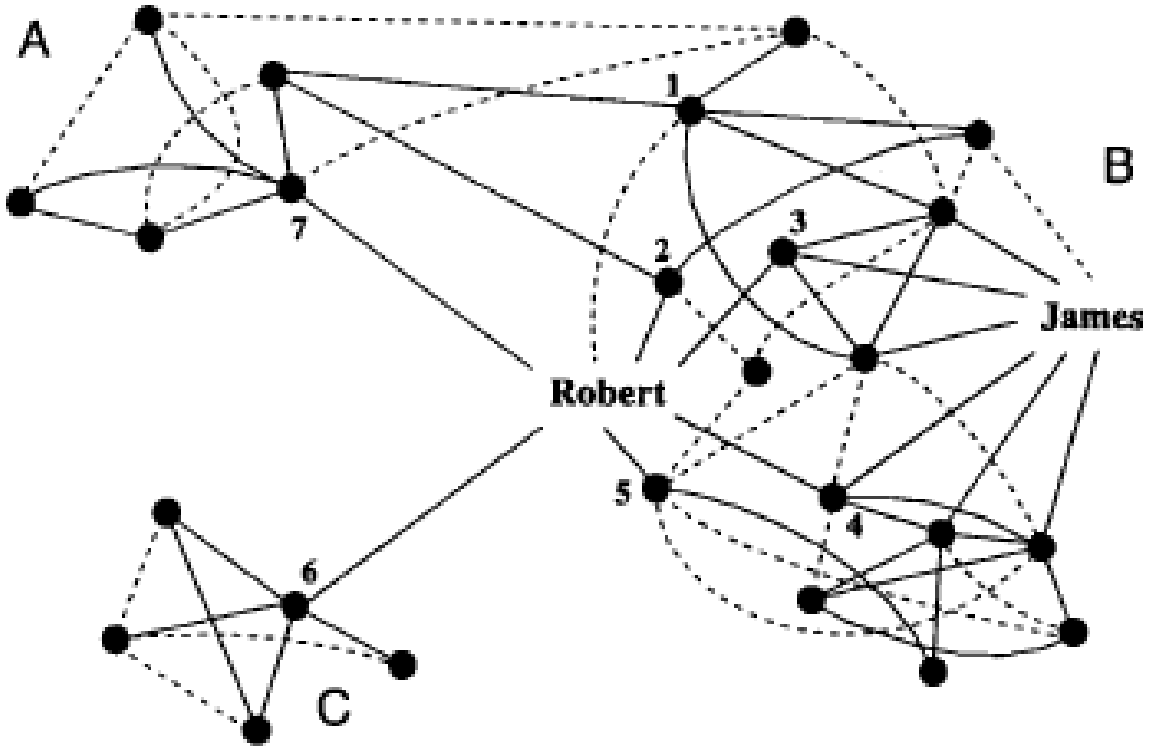
Student Composition

Thompson;Woodward; Bidwell

How can a Teacher Access Relevant Knowledge?

- Jackson & Brueger
- Within subgroups
  - Others understand context and language
- Between subgroups (Burt)
  - Realize the strategic advantage of being a bridger in a network
- Individual bridger more likely to advance within organization
    - Information
    - Framing
    - Opportunities for creativity

The Power of the Bridger: Access to Unique Information and Opportunities

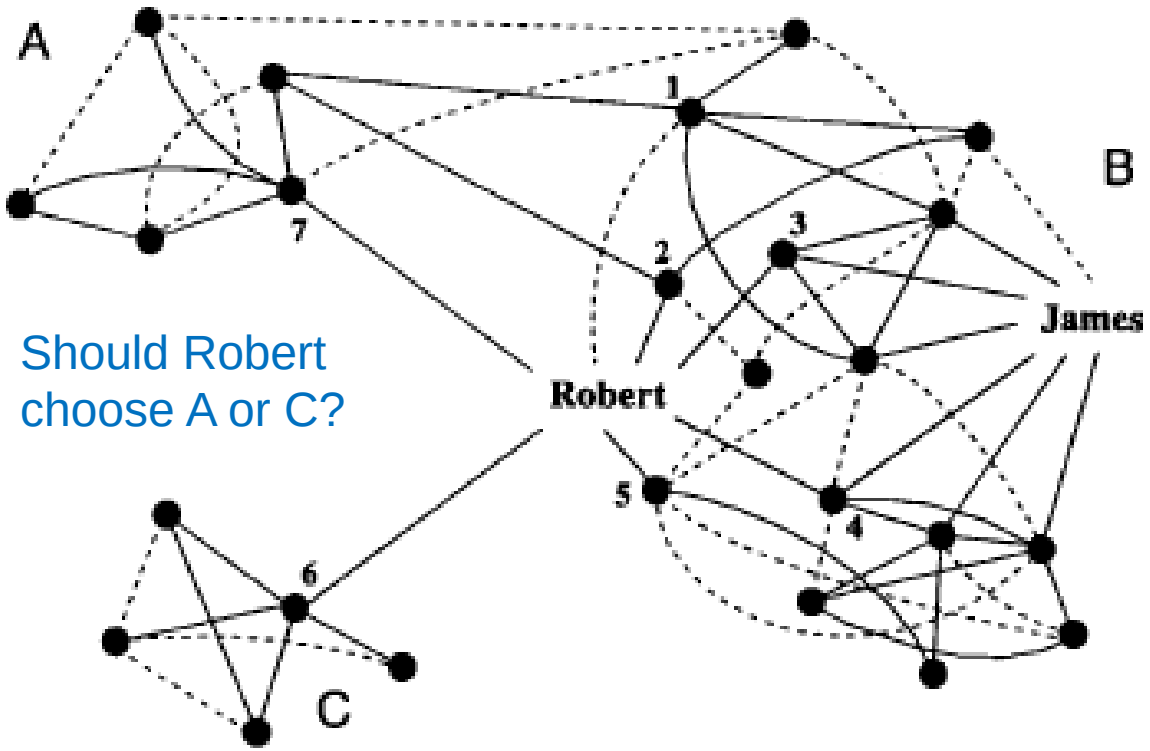


(2000) "The network structure of social capital," Pp. 345-423 in *Research in Organizational Behavior*, Volume 22, edited by Robert I. Sutton and Barry M. Staw. Greenwich, CT: JAI Press. **Page 349**

But the Bridge to Where? Difficult to Scale to the System

- If actors should advance by building bridges (brokering), then from the system perspective
  - What patterns of knowledge flow at the school level are most related to overall change in behavior?
    - Which subgroups should broker?
    - From which subgroups to which subgroups?
  - Policy: What patterns of knowledge flow at the school level are most related to overall change in behavior?
- Formal Research Questions
  - Should social capital flow uniformly to all subgroups?
  - What is the optimal distribution of social capital from subgroups?

The Power of the Bridger: Access to Unique Information and Opportunities



Should Robert choose A or C?

(2000) "The network structure of social capital," Pp. 345-423 in *Research in Organizational Behavior*, Volume 22, edited by Robert I. Sutton and Barry M. Staw. Greenwich, CT: JAI Press. **Page 349**

Moving from Individual to Organization level with Yong Zhao

- Individual access to knowledge changes implementation
  - Frank, Zhao and Borman 2004; Zhao and Frank 2003; Penuel Frank and Krause
- Influence for Individual change, but how about at the organizational level?
- Everybody needs lots of access to knowledge?
  - Just demands more resources
    - More experts
    - More interaction
- Get smarter about knowledge distribution: What patterns of knowledge flow at the school level are most related to overall change in behavior?
  - Why are some schools able to leverage their sources of knowledge to change behaviors/practices, and others not?

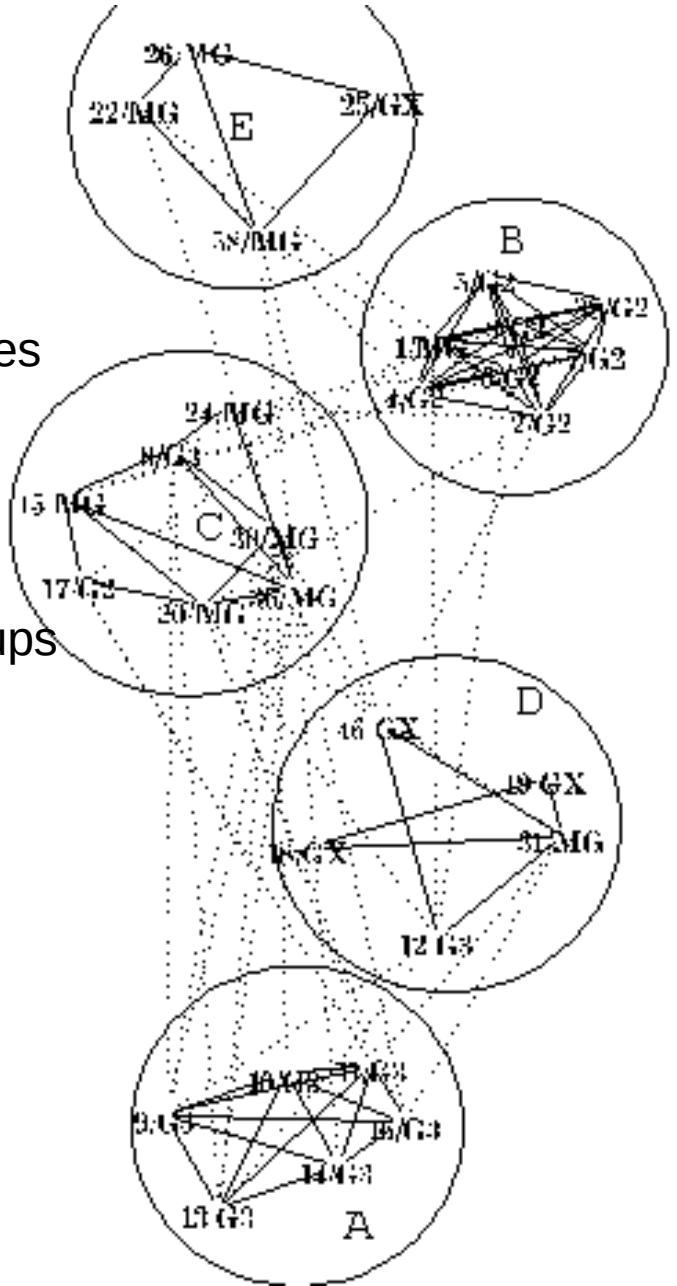
A Crystallized Sociogram in Westville

Number is teacher ID
Text indicates grade:
G2=grade 2

Circles indicate subgroups,
Lines indicate close colleagues
"who are your closest colleagues in this school?"

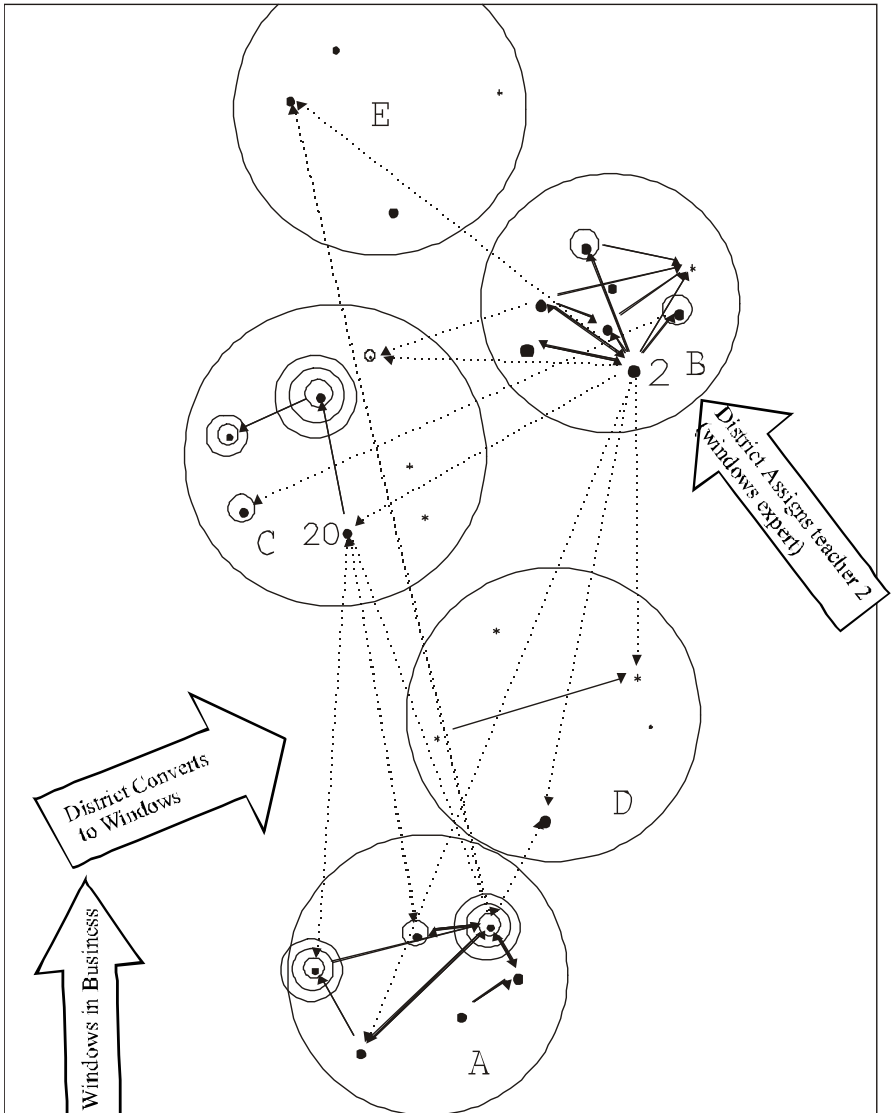
Solid lines within subgroups,
dotted lines between subgroups

Subgroups identified by
Frank (1995) algorithm



Distance between A and B reflects history of school: SES integration

A Ripple Plot of Westville



ID's replaced by •

Size of • represents extent of implementation of technology at time 1

○ indicates change in implementation between time 1 and time 2

Lines indicate help with technology

Same social geography as crystalized sociogram

Interpretation: Diffusion and Embeddedness

- Resource flows structured by the underlying pattern of close collegial ties.
  - Teacher 2's expertise flowed within subgroup B (bonding social capital)
  - To subgroup C (bridging)
  - To Subgroup A (bridging)
- Need both figures to tell the story
  - Crystalized sociogram: Underlying social structure
    - Do not need a separate social structure for each diffusion process
  - Rippled plot: Resource flows and changes in behaviors

From Single Case to Cross-case comparison

- Crystalized sociogram and ripple plot good for single case study
- But what systematically affects implementation of innovation across schools?

Implementation of Reforms Across 21 Schools

Data

425 school staff from 21 schools from a single state in the Pacific West
2 time points, roughly one year apart (Spring 2004, Spring 2005)

Measures

Implementation of localized reform (α = .92) . How much have you implemented the reform in …

curriculum materials used
instructional strategies and activities used with students
assessment strategies used
standards and topics covered
performance levels expected of students
complexity of work assigned to students
classroom management techniques employed
student grouping methods used
professional development sought out
roles and relationships in the school

Social Network:

Who are your closest professional colleagues in this school?

Who has helped you in the past twelve months implement the primary schoolwide initiative (frequency of interaction: once or twice a year = 1, monthly = 9, weekly = 40, daily = 160)

Table 1

Characteristics of Schools in the Sample

School	Enrollment	Percent Free/Reduced Price Lunch	Native American	Asian	African American	Hispanic	White	Multiple Races
A	668	0.15%	0.3%	8%	2%	4%	84%	2%
B	730	85.89%	0.0%	8%	10%	82%	0%	0%
C	526	24.71%	0.0%	10%	3%	42%	40%	6%
D	354	20.90%	0.9%	9%	3%	28%	52%	7%
E	501	1.80%	0.3%	8%	1%	13%	58%	21%
F	301	Missing	0.3%	8%	0%	7%	71%	14%
G	527	18.79%	1.0%	13%	3%	68%	15%	0%
H	1036	36.49%	0.9%	13%	10%	29%	43%	5%
I	274	45.99%	0.0%	9%	23%	24%	32%	11%
J	315	70.16%	0.0%	6%	66%	26%	1%	2%
K	321	8.41%	0.6%	17%	4%	19%	55%	6%
L	480	22.29%	0.4%	8%	1%	17%	65%	9%
M	455	65.89%	0.8%	19%	3%	64%	10%	3%
N	633	66.82%	0.1%	3%	1%	55%	37%	4%
O	611	52.22%	0.0%	3%	0%	90%	6%	1%
P	434	5.30%	0.7%	31%	2%	8%	58%	0%
Q	784	8.29%	0.2%	11%	0%	10%	70%	8%
R	771	100%	0.0%	19%	9%	67%	3%	2%
S	526	49.81%	0.4%	2%	1%	69%	24%	4%
T	1088	9.10%	0.4%	18%	17%	57%	7%	1%
U	473	74.21%	0.0%	16%	2%	76%	4%	2%
V	627	69.17%	0.0%	3%	7%	86%	2%	2%
W	476	4.62%	0.8%	13%	1%	13%	52%	20%
X				Missing				
SAMPLE AVERAGE	*561*	*38.23%*	*0.4%*	*11.1%*	*7.4%*	*41.4%*	*34.2%*	*5.6%*
STATE AVERAGE	**676**	**50.6%**	**0.7%**	**8.6%**	**6.5%**	**52.0%**	**26.1%**	**2.1%**

Characteristics of Faculty Members in the Sample

	N	%
Gender		
Male	73	17.2%
Female	352	82.8%
Race/Ethnicity		
White	338	79.5%
African American	11	2.6%
Hispanic/Latino	43	10.1%
Asian	18	4.2%
Other/Unknown	15	3.5%
Certification Status		
Provisional	57	13.4%
Emergency	5	1.2%
Clear	341	80.2%
National Board	9	2.1%
Missing	13	3.1%

Teaching Assignment

	N	%
PreK	3	0.7%
K	31	7.1%
1	61	14.0%
2	48	11.0%
3	57	13.1%
4	28	6.4%
5	54	12.4%
6	37	8.5%
7	10	2.3%
8	40	9.2%
9	2	0.5%
10	12	2.8%
11	6	1.4%
12	32	7.4%
Other	14	3.2%

Deep Dives

- Crystalized sociogram and ripple plot in each school
- Incorporate information about characteristics of teachers by subgroup
- Interpret with field researchers
  - 2 -3 hours per school
- Understand school as a case

See: Penuel, W., Riel, M., Krause, A., & Frank, K. A. 2009. Analyzing Teachers' Professional Interactions in a School as Social Capital: A Social Network Approach. Teachers College Record Volume 111 Number 1, 2009, p. - http://www.tcrecord.org/Content.asp?ContentID=15174.

Hypothesis Generation 1

- Distance from knowledge base a key to organizational change
- Scaling up from individual level
- H_1: The more *even* the flow of know-how *to* potential recipient subgroups the greater will be the systemic implementation of practices dependent on the know-how.

Hypothesis Generation 2

- Small number of sources of knowledge
  - Those with higher expertise/know-how
  - Develop more explicit knowledge through focused interaction within the subgroups
    - will test this assumption
  - Exert more pressure on others to conform
  - Provide unambiguous messages
- Members of the organization will access similar know-how → coordination
  - All else being equal, it is better for one subgroup to provide know-how to three other subgroups than for three different subgroups to provide the know-how separately to each subgroup.
- H_2: The more know-how is *restricted* to flow *from* subgroups with high levels of know-how, the greater will be the systemic implementation of practices dependent on that know-how.

Comparison of Hypotheses

- Hypothesis 1 assumes quality of resource is constant, therefore focus on distribution to recipients

- Hypothesis 2 assumes variation in quality of resource, therefore focus on distribution from providers

Quantifying Distribution of Resources

- Average access to knowledge?
  - No, focus on distribution
- Standard deviation of access to knowledge
  - Didn’t work to predict school level change in implementation

Ann Krause Suggests: Shannon Indices

- From physics, communication and ecology
- How much entropy is there in resource flow
  - A lot of entropy means flows could occur across any path
  - A little entropy means flows are targeted
- Shannon:

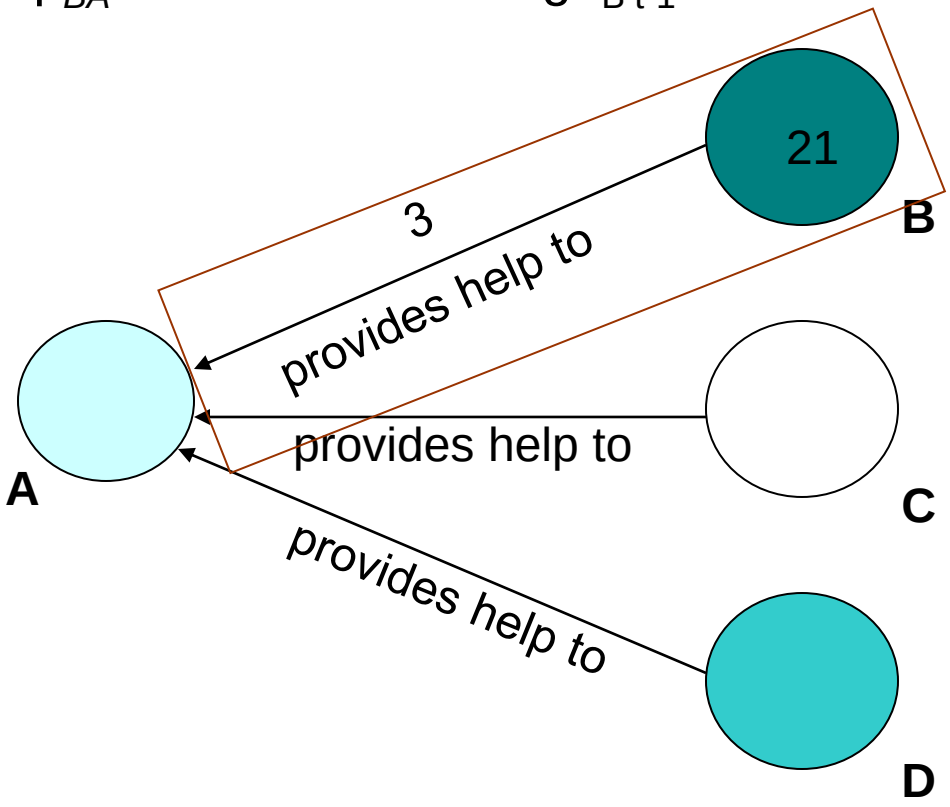
  "The fundamental problem of communication is that of reproducing at one point, either exactly or approximately, a message selected at another point."

Constructing Measures of Entropy: Start with Social Capital

$w_{ii'}$, =Flow of knowledge from *i'* to *i*

=$help_{ii}$ x level of knowledge$_{i' t-1}$

e.g., $help_{BA}$ x level of knowledge$_{B t-1}$



If A talks to B 3 times/week and B has 7 units of knowledge, then A has potential access to 21 units of knowledge through interactions with B:
3 x 7 =21

Convert to Probability that Resource Flows over a Given Path

$$p(i,i') = \frac{w_{i,i'}}{\sum_{i,i'} w_{i,i'}}$$

Probability of flow over a path = flow over that path/all flows over all paths

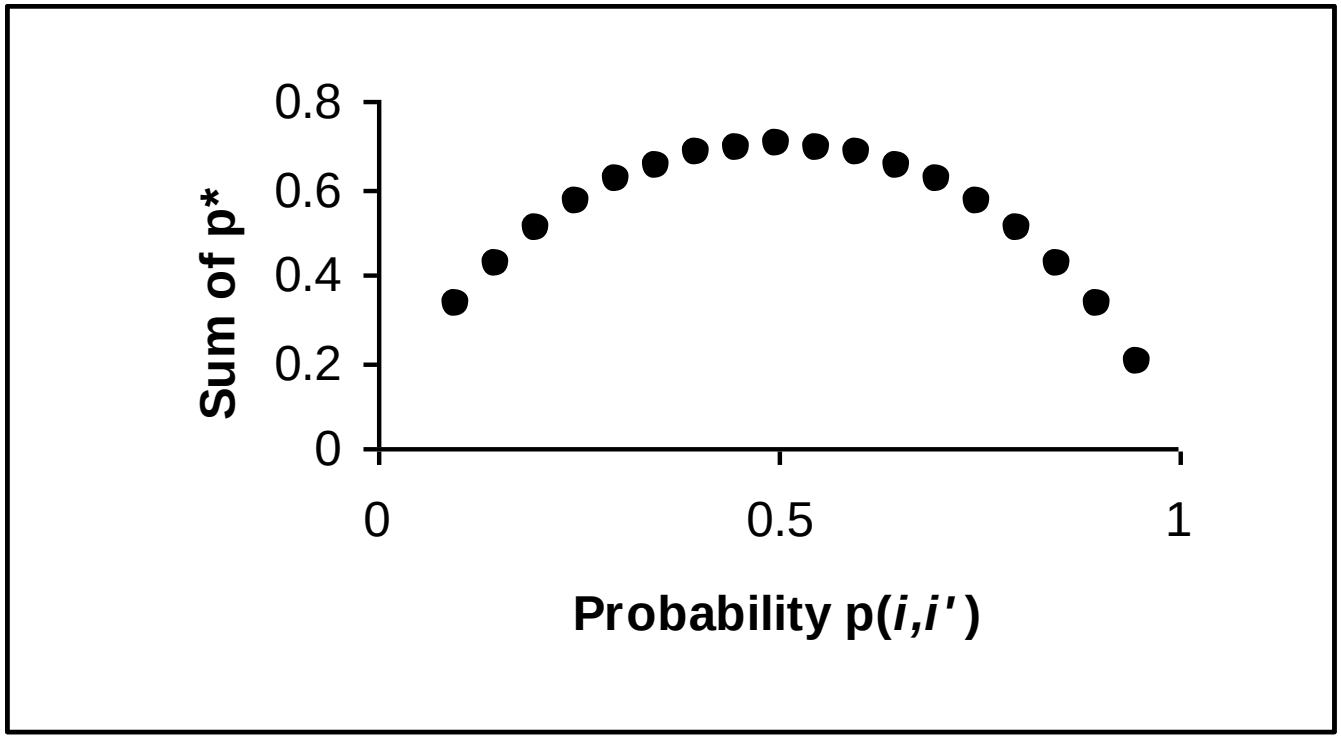
Shannon's measures

$$p^{*}(i,i') = -p(i,i')Ln[\,p(i,i')]$$

Transformation of probability

For single flow, doesn't mean much, but for system…

The *Sum* of Two Transformed Resource Flow Probabilities



Measure has high value when equal probability of flow over either path
Symmetric, smooth.

Organizational Change as a function of Sink and Source Entropy

- Sink entropy: are resources evenly received throughout system
  - Coordination as units access equal knowledge, can act in unison
- Source entropy: are resources evenly sent throughout the system
  - Coordination via control of quality and type of knowledge

Separating Sources from Sinks

$$E_{sources} = -\sum_{i,i'} p(i,i') \log \sum_{i'} p(i,i')$$

Entropy from sources: large if flow from many sources of knowledge

$$E_{sinks} = -\sum_{i,i'} p(i,i') \log \sum_{i} p(i,i')$$

Entropy from sinks: large if flow to many sinks or receivers

Subgroup as the Unit of Analysis

- Relevant variation is *between* subgroups: Bridging social capital
- Within subgroups
  - Strong pressure to conform
  - Many opportunities for information to flow
  - Homogeneity within subgroups
- Use subgroup as unit of analysis:
  - How should resources flow between subgroups to most effect change at school level?

Restatement of Hypotheses in terms of Entropy Across Subgroups

H_1 [restated in terms of entropy]: The *greater* the entropy of the potential flow of know-how *to* subgroups the greater will be the systemic implementation of behaviors dependent on that know-how.

H_2 [restated in terms of entropy]: The *less* the entropy of potential flow of know-how *from* subgroups the greater will be the systemic implementation of behaviors dependent on that know-how.

School Level Regression

School level Regression

$$\text{Mean change in implementation}_j = \beta_0 + \beta_1 \text{potential receiver entropy} + \beta_2 \text{potential provider entropy} + e_j$$

Where the e_j are assumed $N(0,\sigma^2)$.

A positive value of β_1 indicates that the greater the entropy of the flow of resources to potential recipients, the greater the change in implementation (hypothesis 1).

A negative value of β_2 indicates that the less the entropy of the flow of resources from potential providers, the greater the change in implementation (hypothesis 2).

Outcome defined only for teachers in school at both time points
Entropy based on networks at time 2, teacher know-how at time 1

Multilevel Version of Model

For teacher i in school j:

Level 1 (teacher level):

$$\text{change in implementation}_{ij} = \beta_{0j} + e_{ij};$$

Level 2 (school level):

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\text{potential receiver entropy} + \gamma_{02}\text{potential provider entropy} + u_{0j}.$$

Where the u_{0j} are assumed N(0,τ). A positive value of γ_{01} indicates that the greater the entropy of the flow of resources to potential recipients, the greater the change in implementation (hypothesis 1). A negative value of γ_{02} indicates that the less the entropy of the flow of resources from potential providers, the greater the change in implementation (hypothesis 2).

Inference strategy

- Change in teacher's behaviors
  - Equivalent to fixed effects for teachers: controls out anything that is constant to the teacher over time
- Key predictors are emergent properties of school
  - Not a function of a single actor
  - Actors may not be aware of key forces (potential knowledge flow) and their effects
  - Distribution of resource flow unlikely to be attributed to behavior or characteristics of a single actor

  → Control for

  structural characteristics of school: Size, duration of reform, perceived explicitness of initiative, *Perceived pressure to implement the school-wide initiative, School size,*
  - *structure of network: density of network, extent of clustering*
- Quantify how strong an omitted variable would have to be to change an inference
  - Informs discussion about inference

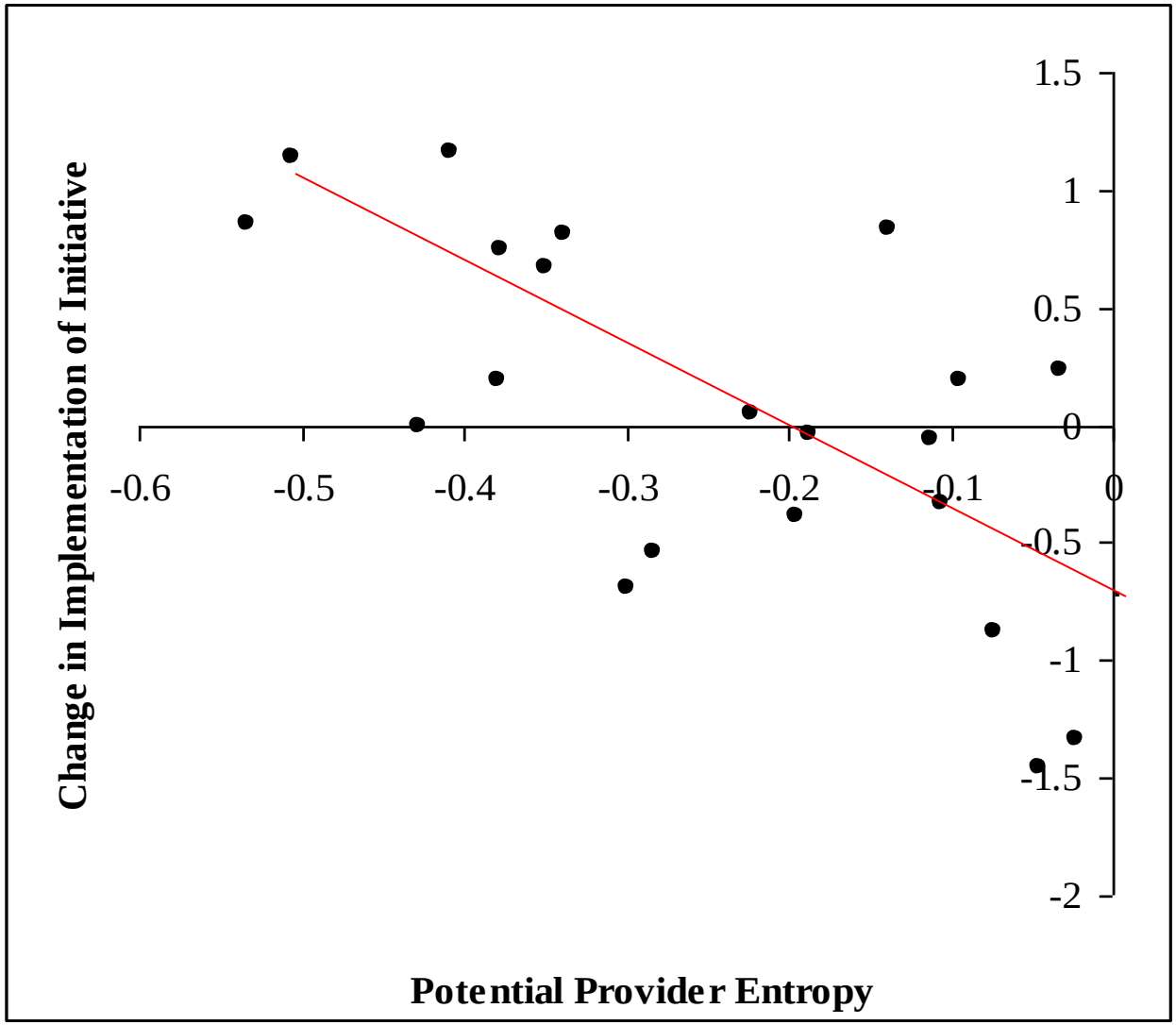
Table 3

Correlations among Change in Level of Implementation of the School-wide Initiative and Measures of Entropy

	Change in Implementation of the School-wide Initiative	Potential receiver entropy	Potential provider entropy
Potential receiver entropy	-.059		
Potential provider entropy	-.654**	.236	
Mean	.056	-.329	-.244
Std	.759	.197	.160

* $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$

Linear Trend between Potential Provider Entropy and Change in School-wide Initiative Implementation



Regression of School Level Change in Implementation of the School-wide Initiative on Measures of Entropy

Independent Variable	Model 1	Model 2 (Potential Confounding)
Potential receiver entropy (H1)	0.39	.28
	(0.70)	(.71)
Potential provider entropy (H2)	-3.22**	-3.18**
	(.86)	(.87)
Duration of the school-wide initiative		-.05
		(.06)
R^2	.44	.46
(adjusted R^2)	.38	.37

* $p \leq .05$; ** $p \leq .01$. n=21 schools.

Hypotheses Evaluation

H_1 [restated in terms of entropy]: The *greater* the entropy of the flow of know-how *to* subgroups the greater will be the systemic implementation of behaviors dependent on that know-how. NO

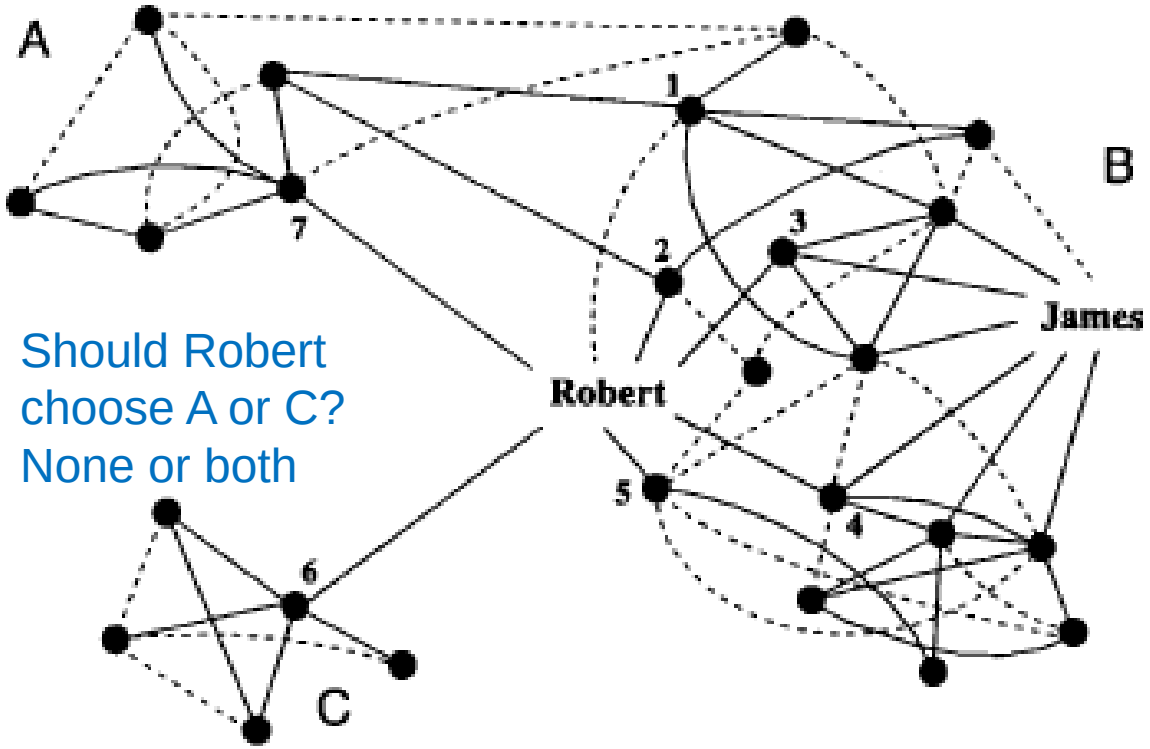
H_2 [restated in terms of entropy]: The *less* the entropy of potential flow of know-how *from* subgroups the greater will be the systemic implementation of behaviors dependent on that know-how. YES

- Alternative factor would have to be correlated about .6 with source entropy and .6 with change in implementation to invalidate the inference that source entropy causes change in implementation
  - (impact of omitted variable necessary to invalidate the inference =.6 x .6 =.36).
- duration of the school-wide initiative correlated -.16 with potential provider entropy and -.18 with change in implementation. Product=-.16 x -.18=.03.
- To invalidate inference, the impact of omitted variable must be 10 times stronger than impact of duration of the initiative.

Interpretation

- The larger the number of subgroup sources (greater entropy), the *less the change. Strong effect*
  - few subgroup sources creates large effect
- Why?
  - Better coordination
  - More explicit knowledge/know-how
  - Committed experts can exert pressure
  - Coherent message
    - Can’t shop around for sources that condone what you do

The Power of the Bridger: Access to Unique Information and Opportunities



Should Robert choose A or C?
None or both

(2000) "The network structure of social capital," Pp. 345-423 in *Research in Organizational Behavior*, Volume 22, edited by Robert I. Sutton and Barry M. Staw. Greenwich, CT: JAI Press. **Page 349**

Theoretical Implications

- Filling structural holes
  - Good for individual
  - What happens if everybody follows Ron Burt's advice?
  - Maybe not good for the system
- Locate knowledge flow in social network of organization
  - Not just social structure or level of knowledge
- Between subgroup flows are key
  - Bridging social capital
  - There is a qualitative difference in knowledge possessed by subgroups.

Policy Implications

- Cultivate interactions with knowledge expert subgroups
  - Concern: could elevate voice and status
  - But subject specific (everyone gets their 15 minutes of fame)
- How to cultivate interactions?
  - Provide venues
  - Provide release time and other resources

EFFECTIVE TEACHERS AND LEADERS

OUR APPROACH

- **Flexibility with results.** Flexible formula grant funding conditioned on SEA and LEA improvements in teacher and leader effectiveness and equity.
- **Fair, rigorous evaluation systems.** Focus on teacher effectiveness and improved teacher evaluation through requirements that LEAs implement a state-approved evaluation system that uses multiple rating categories, takes into account student achievement results, and provides meaningful feedback and support to teachers for improvement.
- **Strengthen the profession.** Treat teachers like the professionals they are by providing time for collaboration, implementing performance-based pay and advancement, and providing on-the-job learning opportunities with peers and experts linked to evaluations and to student needs.
- **Equity.** More equitable distribution of qualified and effective teachers and leaders through better data, an equity plan, and a requirement that Title II funds be directed toward improving equity where LEAs are not meeting performance targets.
- **Data for transparency and decision-making.** Use of meaningful data and accountability for results through program performance measures, state and district human-capital report cards, and tracking the effectiveness of professional development and teacher-preparation programs.

ON-LINE APPENDIX B

Multilevel Estimation of Main Models

The estimates in Table 4 can be obtained alternatively by a multilevel model for teacher i in school j:

Level 1 (teacher level):

$$\text{change in implementation}_{ij} = \beta_{0j} + e_{ij};$$

Level 2 (school level):

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\text{potential receiver entropy} + \gamma_{02}\text{potential provider entropy} + u_{0j}.$$

Where the u_{0j} are assumed $N(0,\tau)$. A positive value of γ_{01} indicates that the greater the entropy of the flow of resources to potential recipients, the greater the change in implementation (hypothesis 1). A negative value of γ_{02} indicates that the less the entropy of the flow of resources from potential providers, the greater the change in implementation (hypothesis 2).

Estimates from the multilevel model are given below. The parameter estimates for potential provider entropy were well within 10% of those reported in the main text, and none of our inferences change from the model we report in the main text. Therefore we present the single level regressions in the main text for their interpretability in terms of the model coefficients and the graphical and sensitivity analyses.

Alternative to Table 4

Multilevel Regression of School Level Change in Implementation of the School-wide Initiative on Measures of Entropy

Independent Variable	Model 1	Model 2 (Potential Confounding)
Intercept	-.69	-.53
	(.43)	(.49)
Potential receiver entropy	0.31	.21
	(0.93)	(.94)
Potential provider entropy	-3.11**	-3.12**
	(1.09)	(1.09)
Duration of the school-wide initiative		-.05
		(.08)
Overall R^2	15%	.15%

* $p \leq .05$; ** $p \leq .01$. n=21 schools.

6% of the variance is at the school level in an unconditional model.

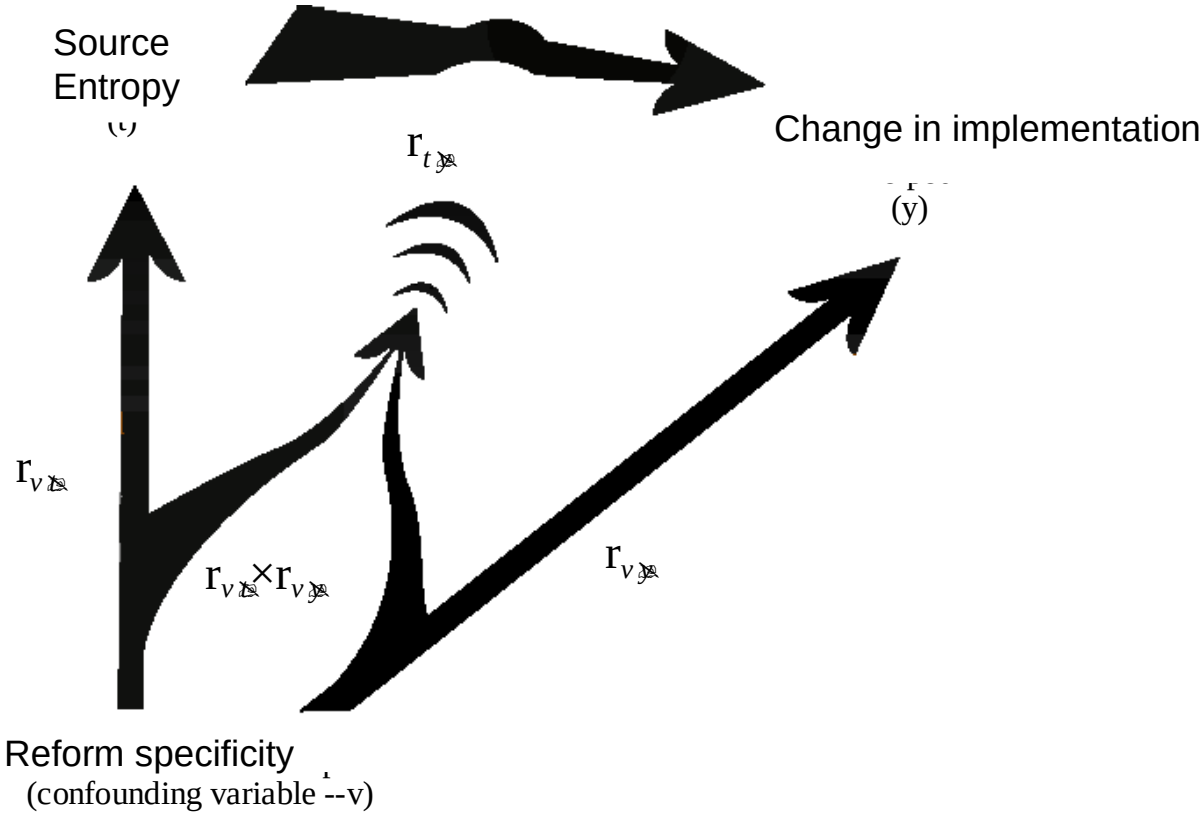
END HERE

Concern over Missing Confound (Internal Validity)

- Causal Inference concern: How much of the difference would have to be attributed to other factors to invalidate the causal inference? Maybe schools that have more tightly specified reforms have fewer subgroups supporting the reform and increase more in implementation.

- We may never know ,but we can quantify the concern
  - What would the impact of a confound (e.g, inclination to help) have to be to alter our Inference? (Frank, 2000)

Impact of a Confounding Variable on a Regression Coefficient



What must be impact to Invalidate the Inference?

y = outcome (change in implementation)

x = predictor of interest (source entropy)

$r_{x \cdot y}$ = the correlation between *x* (source entropy) and *y* (change in implementation)=.65

$r_{x \cdot cv}$ = correlation between *x* and an unmeasured confounding variable

$r_{y \cdot cv}$ = correlation between *y* and an unmeasured confounding variable

$k = r_{x \cdot cv} \times r_{y \cdot cv}$ = Impact of Unmeasured Confound

Question:

What must be k to invalidate the inference that source entropy affects how much a school changes in implementation?

Procedure for Robustness Indices

- 1) Define a threshold for inference
- 2) Express focal statistic in components reflecting unknown
- 3) Set focal statistic equal to threshold and solve for threshold
  - make assumptions like maximizing.

→ conditions necessary to invalidate inference

1) Defining a Threshold for Inference

- Define $r^{\#}$ as the value of r that is just statistically significant:

$$r^{\#} = \frac{t_{critical}}{\sqrt{(n - q - 1) + t^2_{critical}}}$$

n is the sample size
q is the number of parameters estimated

$r^{\#}$ can also be defined in terms of effect sizes or correlation coefficients.

2) Express focal statistic in components reflecting unknown

$$r_{x \cdot y|cv} = \frac{r_{x \cdot y} - r_{x \cdot cv} \times r_{y \cdot cv}}{\sqrt{1 - r_{y \cdot cv}^2}\sqrt{1 - r_{x \cdot cv}^2}} = \frac{r_{x \cdot y} - k}{1 - k}$$

Assuming $r_{x \cdot cv} = r_{y \cdot cv}$ (which maximizes the impact of the confounding variable) (Frank 2000):

3) Set focal statistic equal to threshold and solve for threshold

Set $r_{x \cdot y|cv} = r^{\#}$ and solve $(r_{xy} - k)/(1-k)$ for k to find the threshold for the impact of a confounding variable (TICV).

$$TICV = \frac{r_{xy} - r^{\#}}{1 - |r^{\#}|}$$

impact of an unmeasured confound > TICV → inference invalid
impact of an unmeasured confound ≤ TICV → inference valid.

Application to Estimated Treatment Effect for Source Entropy on Change in Implementation

$$r^{\#} = \frac{2.09}{\sqrt{19 - 2.09^2}} = .44$$

$r_{source\ entropy \cdot change\ implementation} = .65$

$$TICV = \frac{.65 - .44}{1 - |.44|} = .38$$

https://www.msu.edu/user/k/e/kenfrank/web/research.htm#causal

Quantifying the Robustness of the Inference

- Inference is invalid if the impact (impact=$r_{x\,cv}$ x r_{ycv}) of an unmeasured confound is greater than .36.
- Assuming $r_{x\,cv} = r_{y\,cv}$, *each would have to be greater than.60 (.60=.38$^{.5}$) to invalidate the inference.*
- *Compare with impacts of observed covariates*
  - *Impact of longevity of reform is .029*
  - *Impact of school size is .003*
- *impact of unmeasured confound would have to be 12 times greater than impact of a strong theoretical contender (longevity of reform) to invalidate inference*
- *We hope you will concur with our inference*
- Can also say: 35% of the estimate must be due to bias to invalidate the inference

Consider Alternate Sample (External Validity)

Causal Inference concern:

How much of the effect would have to be context specific to invalidate the causal inference?

We cannot assert cause if the effect is not constant across contexts.

Statistical Translation:

Would the inference be valid if the sample included more of some population (e.g. teachers in other states) for which the effect was not as strong?

Procedure for Robustness Indices (Frank and Min 2007)

- 1) Define a threshold for inference
- 2) Express focal statistic in components reflecting unknown
- 3) Set focal statistic equal to threshold and solve for threshold
  - make assumptions like maximizing.

→ conditions necessary to invalidate inference

1) Defining a Threshold for Inference

- Define $r^{\#}$ as the value of r that is just statistically significant:

$$r^{\#} = \frac{2.09}{\sqrt{19 - 2.09^2}} = .44$$

n is the sample size
q is the number of parameters estimated

$r^{\#}$ can also be defined in terms of effect sizes or correlation coefficients.

2) Expression of focal statistic Consider Alternate Sample (External Validity)

Define π as the proportion of the sample that is replaced with an alternate sample.

r is correlation in unobserved data

R is combined correlation for observed and unobserved data:

$$R_{xy}=(1-\pi)r_{xy} + \pi \mathbf{r}_{xy} .$$

Assumes equal variances and means in observed and observed.

3) Conditions Necessary to Invalidate Inference

Set $\mathbf{R}=r^{\#}$ and solve for $\mathbf{r}_{xy}$:

If half the sample is replaced (π=.5), original inference is invalid if $\mathbf{r}_{xy} < 2r^{\#}-r_{xy}$

Therefore, $2r^{\#}-r_{xy}$ defines the threshold for replacement: TR(π=.5)

If $\mathbf{r}_{xy}$ =0, inference is invalid if $> 1-r^{\#}/r_{xy}$. Therefore $1-r^{\#}/r_{xy}$ defines the threshold for replacement: TR($\mathbf{r}_{xy}$=0)

Assumes means and variances are constant across samples, alternative calculations available.

Application of Thresholds for Replacement for Effect of Source Entropy on School Change in Implementation

TR(r_{xy} =0)= 1-$r^{\#}/r_{xy}$ =1-(.44/.65)=.32

About one third of the schools would have to be replaced with others in which there was relation between Source Entropy and School Change to invalidate the inference

Can also be interpreted as the proportion of bias in r_{xy} necessary to invalidate the inference.

35% of the estimate must be due to bias to invalidate the inference

Flow of Knowledge within and between Subgroups

- First knowledge adapted and developed within subgroups
- Then bridgers between subgroups
- Example:
  - Diffusion of technology within a single school (Frank & Zhao, 2005)