Diffusion and use of Evidence Based Practices in Education: Implications for Criminal Justice

Kenneth A. Frank (kenfrank@msu.edu)
College of Education and Fisheries and Wildlife
Michigan State University
https://www.msu.edu/~kenfrank/
Presented at the Panel on Criminal Justice
National Academy of Sciences
Jan 31 2014
Overview

• The process of implementation
  – Theory of complex production in education
  – Implications for local networks to locally adapt

• Implications for discourse about causal inferences in other contexts
  – Sensitivity analysis, bounding, confidence

• Implications for national networks for diffusion

• Application to Criminal Justice
Factors Affecting Implementation

• Organizational control
  – Hierarchy: can direct implementation form above
  – Collegial: collaborate and coordinate to implement

• Production process
  – Simple and uniform: can adopt practices from other contexts with high fidelity
  – Complex: requiring local adaptation
The Production Process in Education
Teaching is Complex

• Student composition
• Draw on previous training/experience
• Curriculum
• Pedagogy
• Administrators
• Assessments
• Local coordination

Teaching as Complex Production

- Previous Training
- Curriculum
- pedagogy
- Assessments
- Coordination
- Difficult to control through hierarchy

Student Composition
Knowledge is Key

• Local knowledge for how to implement new behavior in context
• Adapt to existing features
  – Previous curriculum
  – Student composition
• Coordinate with others’ practices
• Respond to external expectations
Question 1

• How are criminal justice organizations controlled?

• What is the production process?
  – How much must it be locally adapted and coordinated?

• Answers will affect capacity to implement evidenced based practices with “high fidelity.”
Local adaptation Creates Need for Local Sources of Knowledge

• Pre-service training
  – Adequate?
• Externally provided professional development
  – Localized?
• Opportunities to experiment
  – Possible, but inefficient
• On the job knowledge sharing
  – Local networks are important
The Bases for Informational Flows

• If local knowledge is important, then teachers should access knowledge from colleagues

• But why would a colleague provide help? It takes time, might be distracting
  – Gain status?
  – Direct exchange?
  – Social capital exchange:
    help ↔ conformity
Social Capital Exchange
Knowledge for Compliance to Norms

Social Capital and the Diffusion of Innovations within Organizations: Application to the Implementation of Computer Technology in Schools."
Interpretation

• Teachers’ practices are influenced by the previous practices of those they nominate as close colleagues or sources for advice
  – Knowledge to support
  – Norms to constrain
  – Small to moderate effects

• Causality:
  – models are based on change in behavior over time: controls for tendency to seek others who have similar behaviors to own
  – will return to issues of causal inference
• Social Factors and Implementation of Innovations in Organizations: https://www.msu.edu/~kenfrank/research.htm#social

• Integrated presentation
• Instruments and related resources


Question 2: Importance of Networks for Local Knowledge

• How important are personal networks for obtaining local knowledge about law enforcement?

• Answer will affect
  – who can change practices
  – Persistence of norms
  – How people gain status
In a sharp break with New York’s crime-fighting policies of the last decade, Police Commissioner William J. Bratton has told his top chiefs that he intends to fundamentally alter a program that sent waves of rookie officers into crime-ridden neighborhoods, but also inflamed tensions in minority communities.

At a closed-door meeting in Police Headquarters in January, Mr. Bratton said that rather than thrusting inexperienced officers into dangerous or highly charged situations, part of a program known as Operation Impact, he envisioned a return to a more traditional approach where rookies would first be placed in local precincts.

“I want to change the dynamic of kids coming out of the academy and immediately being put into Operation Impact assignments, where they really have an almost single-minded focus and really don’t get a full flavor of the job,” he said, according to a recording of the meeting reviewed by The New York Times.

“I think they would benefit from it, working with officers in traditional precinct assignments,” he added.

The more experienced officers could serve as mentors, Mr. Bratton said, helping the rookies “during that first critical year that they come into the department.”

His comments would amount to a major shift in the department’s
What do I expect?

- Rookies will have more access to local knowledge, more effective
- Mentors will gain in status, more able to exact conformity of rookies
- More difficult to change precinct cultures (for better or worse)
- Factions may emerge within precincts if cliques or subgroups are already present
Implications of Network Models for Intra-Organizational Diffusion

- “Network influence” is not necessarily positive
  - Those with network members engaged in preferred practices increase preferred behaviors
  - Those with network members engaged in non-preferred practices decrease preferred practices

- If there are cliques of teachers who interact a lot, then implementation process can create bifurcation or factions as people conform to the norms within their cliques.
Pre-existing Cliques and the Organizational Response to NCLB

• The No Child Left Behind Act of 2001 changed the institutional environment of schooling
  – Sanctions for schools failing to meet achievement targets for all subgroups of students (“tightening” coupling)
  – Requirement that schools and districts adopt evidence-based programs and practices
  – In reading, a focusing of resources on phonics-based instruction that built decoding skills of early readers (reducing heterogeneity of environment)

• A core assumption of NCLB is that school actors will adapt to the changed environment because they are motivated by the threat of sanctions and promise of resources and rewards

NCLB Pressures

Institutional Environment

Sanctions
Resources (Programs, PD)

School
NCLB Pressures: Varying Initial Practices

Institutional Environment

Sanctions

Resources (Programs, PD)

School
NCLB Pressures: Varying Initial Practices and Subgroups

Institutional Environment

(microfoundations)

School
Normative Pressure

- Pressure result from having a collegial tie (direct effect) with someone or from being part of the same subgroup (indirect effect)
- Individual teachers may be particularly responsive to pressure from subgroup members to the extent that:
  - They share a common context for teaching (Smylie, 1989; Kennedy, 2005)
  - High levels of trust exist among subgroup members (see Ingersoll, 2003)
NCLB Pressures
Sanctions
Resources (Programs, PD)

Institutional Environment
TIME 2

School
Summary of Findings

- % of variance between subgroups increased from 2007-2008 by 20%.
- % of variance within subgroups decreased from 2007-2008 by 30%.
- Small variance between schools

Implications

• Subgroups, conformity to subgroup norm (for knowledge exchange)
  → increased variation between subgroups in organizational response
• Uncoordinated effort
• Stratification: which kids/families can compensate?
• Schism affecting next implementation of next innovation
• Changing organizations, not people
  – Changing schools not teachers
  – Changing police departments, not officers?
Issues of Causal Inference: Interpreting Research results

• I said: Teachers’ practices are influenced by the previous practices of those they nominate as close colleagues or as sources for advice.

• How can I be sure I have identified the correct causal mechanism?

• It could be that there is a variable omitted from my model that explains choice of network members as well future behavior.
Alternative Explanation for Network Influence

Prior behavior of network members

Prior behavior (teaching basic skills at time 1)

Selection

Influence

Outcome behavior
(teaching basic skills at time 2)
Accounting for Selection of Network Members

• Control for prior behavior in our models
• Control for other factors that may affect choice of network members
  – Prior training
• In educational research, controlling for prior behaviors works fairly well
Observational studies with controls for pretests work better than you think
Shadish et al (JASA 2008)

- Quantify how much biased removed by statistical control using pretests in a given setting
  - Sample: Volunteer undergraduates
  - Outcome: Math and vocabulary tests
  - Treatment:
    - basic didactic,
    - showing transparencies
    - defining math concepts

OLS Regression with pretests removes 84% to 94% of bias relative to RCT!!
Propensity by strata not quite as good

See also Concato et al., (2000) in medical research
See Kane, T., & Staiger, D. (2008) in economics
Criminal Justice?
Example of Randomized Experiment in Education: Effect of Open Court Curriculum on Reading Achievement

- Open Court “scripted” curriculum versus business as usual
  - Not locally adapted
- 917 elementary students in 49 classrooms
- Comparisons within grade and school
- Outcome Measure: Terra Nova comprehensive reading score

Differences between Open Court and Business as Usual

• Difference across grades: about 10 units
  – 7.95 using statistical model
• “statistically significant” unlikely (probability < 5%) to have occurred by chance alone if there were really no differences in the population
• But is the Inference about Open Court valid in other contexts?
  – How much of the estimate must be due to sampling bias (with volunteer sample) to invalidate the inference?
Inferences from Randomized Studies: The Fundamental Problem of External Validity

• Before a randomized experiment:
  – People believe they do not “know” what generally works
  – People choose treatments based on idiosyncratic conditions -- what they believe will work for them (Heckman, Urzua and Vytlacil, 2006)

• After a randomized experiment:
  – People believe they know what generally works
  – People are more inclined to choose a treatment shown to generally work in a study because they believe “it works”

• The population is fundamentally changed by the experimenter (Ben-David; Kuhn)

• The fundamental problem of external validity
  – the more influential a study the more different the pre and post populations, the less the results apply to the post experimental population
  – All the more so if it is due to the design (Burtless, 1995)
% Exceeding Threshold for Open Court Estimated Effect

\[
\hat{\delta} = 7.95
\]

54 % above threshold = 1 - 3.68/7.95 = 0.54

\[\delta^# = 3.68\]  Threshold for statistical significance
Interpretation

• To invalidate the inference:
  – 54% of the estimate must be due to sampling bias to invalidate Borman et al.’s inference

• Interpretation in terms of replacement of cases
  – We expect you would have to replace 54% of the cases (about 30 classes) with cases in which Open Court had no effect to invalidate the inference

• Quantifies discourse about the concern of validity

• Can help people talk about implications of a study
Example Replacement of Cases from Non-Volunteer Schools to Invalidate Inference of an Effect of the Open Court Curriculum

**Business as Usual**

**Open Court**

- **Black**: Original volunteer cases that were not replaced
- **Gray**: Replacement cases from non-volunteer schools with no treatment effect
- **Dashed**: Original distribution for all volunteer cases
Why do I Cling to Observational Studies?

• It is difficult to assign people to network members
  – Jim Carey in Truman

• If network dynamics are important within organizations,
  – then individuals within organizations are not independent
  – must randomly assign organizations to treatments, not individuals

• need degrees of freedom at the organization level (e.g., 40 organizations)
  – Bloom et al (2007); Raudenbush & Liu (2001); Raudenbush et al 2007

• Very expensive/study ($10 million / study – Slavin, 2008)

• Narrows the questions that can be asked

• Limits the number of people who can ask questions

• Generally: lesson learned from education:
  – Randomized experiments are valuable if you can do them, but there are $, logistical, and ethical challenges
Back to the Inference from the Observational Study

• I said: Teachers’ practices are influenced by the previous practices of those they nominate as close colleagues or as sources for advice.
  – controlled for prior behavior. That helps

• But can I make the inference that the network has an effect?

• Of course I can. I just might be wrong.
  – It’s causal inference. Not causal proof or certainty.

• But what would it take for me to be wrong?
  – Changing the nature of discourse about causal inferences
% Exceeding Threshold for Network Effect on Teachers’ use of Computers

Controlling for prior use of computers and expertise reduced the estimated network effect by 10%; Omitted variable(s) would have to be twice as strong as prior behavior to change the inference.

\( \{ \text{22\% above threshold (for statistical significance)} \} \)
Interpret Inferences in Terms of Replacing Cases

• What % of the cases must be replaced to invalidate inference?
• For randomized experiment: replacement cases from non-sampled population
• For observational study: replacement cases from counterfactual – exact pairing with those in study, but under opposite condition experienced in the study
• Assume no treatment effect in replacement cases
Different Thresholds Relative to Transaction Costs

Definition: Threshold in policy-oriented research: the point at which evidence from a study would make one indifferent to policy choices

Thresh olds

1. Changing beliefs, without a corresponding change in action.
2. Changing action for an individual (or family)
3. Increasing investments in an existing program.
4. Initial investment in a pilot program where none exists.
5. Dismantling an existing program and replacing it with a new program.

• In research: statistical significance
  • Depends on sample size ...
Other ways to talk about Inferences: All Good

• Bounding (e.g., Altonji et, Elder & Tabor, 2005; Imbens 2003; Manski & Nagin, 1998)
  – *lower bound*: “how small could effect be?”
    • Focus on estimate
  – *% robustness*: “how strong would unobserved factors have to be to invalidate inference?”
    • Focus on inference, policy & behavior

• External validity based on propensity to be in a study (Hedges and O’Muircheartaigh)
  – focus on estimate
  – I focus on comparison with a threshold

• Other sensitivity (e.g., Rosenbaum or Robins or Frank, 2000)
  – Characteristics of variables needed to change inference
  – I focus on how sample must change.
    • Can be applied to observational study or RCT
    • Use propensity scores to differentiate, then apply indices

• Confidence intervals
  – What does it mean to say a confidence interval is close to zero? (see next slide)
Relationship between the Confidence Interval and % Bias Necessary to Invalidate the Inference of an Effect of Open Court on Comprehensive Reading Score

Borman et al
Threshold to invalidate
\( \hat{\delta} - \delta^# \)

Confidence Interval
\( \delta^# \)
\( \hat{\delta} \)

Lower bound of confidence interval “close to 0” → estimate exceeds threshold by small amount

Weaker inference
\( \hat{\delta} - \delta^# \)
\( \delta^# \)
\( \hat{\delta} \)

Confidence Interval
\( \delta^# \)
\( \hat{\delta} \)
Big Picture

• Think about implementation process
  – Many practices must be locally adapted and coordinated
    → local networks matter
    → changing organizations, not people
    → randomized experiments difficult
      studies should be at organizational level
difficult to manipulate networks

• When making inferences, quantify what it must take
to change the inference
  • Improves discourse between research and application
Lesson Learned from Education: The Role of Randomized Experiments

• Many weak designs → push for “rigorous designs”: randomized experiments
• “What Works” clearinghouse, National Center for Educational Research
• Limited value: Few results hold up with randomized experiments.
• Why?
  – More resources for treatment than control, not fair
  – Equal resources: teachers adapt to equal effectiveness
Applications to Criminal Justice

• Which practices must be locally adapted and coordinated?
  – Place based policing?

• Which do not require local adaptation and coordination?
  – Scripted curriculum in education
  – Universal Protocol in medical practice

• What is the nature of current discourse about inferences?

• Great Stuff at: Center for Evidence-Based Crime Policy (CEBCP)
Diffusion Between Organizations

- Venues (e.g., conferences, documents, phone calls) for exposure
- Repeated participation
- Creates structural positions
- Constrains or affords action
Network Location and Policy-Oriented Behavior: An Analysis of Two-Mode Networks of Coauthored Documents Concerning Climate Change in the Great Lakes Region*

Ken Frank, I-Chien Chen, Youngmi Lee, Scott Kalafatis, Tingqiao Chen, Yun-Jia Lo, and Maria Carmen Lemos

This study explores how a scientist’s location in science-based policy networks can affect her policy-oriented behaviors. In particular, we hypothesize that those scientists who fill structural holes in their networks will be more likely than others to engage in policy-oriented behaviors. The network data are defined by scientists’ coauthorship on policy documents regarding climate change in the Great Lakes. We employ a two-mode network analysis to identify clusters of scientists who coauthored similar documents, and relative to those clusters, we identify those who fill structural holes by bridging between clusters. We find that those scientists who bridged between clusters were more likely to engage in policy-oriented behaviors of policy advocacy and advising than were others in the network. This is an example of a link between network location and policy-oriented behavior indicative of the broader phenomenon of how individuals exert agency, given structural constraints.

KEY WORDS: policy behavior, networks, scientists, climate change
Clusters of People who Co-authored Documents about Climate Change in the Great Lakes Region
Technical Appendix: Documents by Cluster

• Cluster 1
  – 20040: Preparing for a Changing Climate: The Potential Consequences of Climate Variability and Change in the Great Lakes Region
  – 20002: Adapting to Climate Change and Variability in the Great Lakes-St. Lawrence Basin
• 20038: Preface to the Potential Impacts of Climate Change in the Great Lakes Region
• 20023: From Impacts to Adaptation: Canada in a Changing Climate 2007

• Cluster 2
  • 20015: Confronting Climate Change in the Great Lakes Region: Impacts on Our Communities and Ecosystems
  • 20016: Confronting Climate Change in the Great Lakes Region
  • 20011: Climate Change in the Great Lakes Region: Starting a Public Discussion
  • 20021: Ecological Impacts of Climate Change
  • 20025: Global Climate Change Impacts in the US: A State of Knowledge Report from the U.S. Global Change Research Program
  • 20031: Informing Decisions in a Changing Climate: Panel on Strategies and Methods for Climate-Related Decision Support
  • 20032: Introduction: Assessing the effects of climate change on Chicago and the Great Lakes
  • 20037: Potential Impacts of Climate Change on U.S. Transportation
  • 20039: Chicago Climate Action Plan
  • 20044: Scientific Assessment of the Effects of Global Climate on the United States
  • 20059: Economic Impacts of Climate Change on Pennsylvania

• Cluster 3
  • 20009: Climate Change Impacts and Adaptation: A Canadian Perspective
  • 20003: Adapting to Climate Change in Ontario: Towards the Design and Implementation of a Strategy and Action Plan (Report of the Expert Panel on Climate Change Adaptation)
  • 20033: IPCC 4th Assessment Report, Working Group II Report "Impacts, Adaptation and Vulnerability" North America, Chapter 14:
Interpretation

- 3 positions (or clusters)
- Statistically significant (rejecting null of no clustering)
  - Not really causal inference
- Each group a mixture of
  - academic and government
- Define bridging role relative to clusters
- Relate bridging role to outcomes
  - Policy advocacy and activism
Measures of Policy Oriented Behaviors

- **Political Advocacy**: Extent to which an actor engages in activities with an intention to influence policy and behavior.
  - participation in meetings,
  - media campaigns regarding climate change issues,
  - participating in conferences and workshops that engage decision-makers,
  - participating in interviews, press conferences, writing articles or blogs to increase awareness of climate change and advocate climate change-related action.

- **Policy advising**: Attendance at policy-related or governmental meetings, in the role of directly informing policies or plans (e.g. contributing solutions, participating in policy design) with research about climate change and expert knowledge.

- **Scale for Both**: 0 to 4 (5 scales).
  - 0: no evidence that the actor was involved in policy advocacy activity
  - 1): the actor’s reports or publications were aimed at being policy-relevant (i.e. expressed the intention or claim that the document could inform policy)
  - 2: actor’s activities were related to policy advocacy, but it was not their primary activity
  - 3: policy advocacy was a primary activity they were given score of at least 3.
  - 4: consistently involved in policy advocacy over time
Bridgers more Engaged in Policy Advocacy and Advising

Those who bridge between clusters of actors were more involved in *policy advocacy* than others in their social system

Bridgers more likely to be engaged in political advising

- 2.15 for bridgers versus .7 for others (on our scale from 0 to 4). ($p \leq .0001$).

- Controlling for differences among groups and sector, the bridgers were more likely to be policy advocates
  - estimated difference of 1.56, standard error of .34, $p \leq .0001$.

Those who bridge between clusters of actors were more involved in policy advising than others in their social system.

- 3.6 for bridger’s versus 2.4 for others (scale of 0-4; $p \leq .002$).

- Controlling for differences among clusters and sector
  - difference of 1.30 (standard error of 0.34, $p \leq 0.001$).

- Key: Inferences might be wrong, but what would it take to be wrong?
  - Comparable to inference about CO2 on temperature

- Not that sensitive to tentatively placed actors
In the last ten years or so it’s become obvious that we need to engage more with other groups, especially scientists. Our organization especially was too insular ten years ago. The issue of climate change has been one of the drivers of realizing that and making an effort to change it.

Structural constraint makes it difficult to bridge

Actors don’t really know effect of social structure on behavior:

“I can’t really tell you what interactions have pushed future involvement and what haven’t.”
Policy implications

• Change agents
  – create venues which affect which social structures can emerge
  – Can influence participation/attendance venues

• Enhanced serendipity
  – Always changing
  – Find gaps and support
  – Encourage people to pursue own links
End Here
Inference from study B more robust than from study A, all else equal.
<table>
<thead>
<tr>
<th>Randomized Block</th>
<th>State</th>
<th>Grade</th>
<th>Class Size</th>
<th>% Free Lunch</th>
<th>% ESL</th>
<th>% Special Education</th>
<th>% Minority</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ID</td>
<td>1</td>
<td>19</td>
<td>78</td>
<td>44</td>
<td>17</td>
<td>50</td>
</tr>
<tr>
<td>2</td>
<td>ID</td>
<td>2</td>
<td>22</td>
<td>70</td>
<td>22</td>
<td>0</td>
<td>35</td>
</tr>
<tr>
<td>3</td>
<td>ID</td>
<td>3</td>
<td>22</td>
<td>58</td>
<td>8</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>ID</td>
<td>4</td>
<td>21</td>
<td>53</td>
<td>16</td>
<td>5</td>
<td>32</td>
</tr>
<tr>
<td>5</td>
<td>FL</td>
<td>1</td>
<td>18</td>
<td>48</td>
<td>0</td>
<td>17</td>
<td>39</td>
</tr>
<tr>
<td>6</td>
<td>FL</td>
<td>2</td>
<td>28</td>
<td>19</td>
<td>0</td>
<td>35</td>
<td>19</td>
</tr>
<tr>
<td>7</td>
<td>FL</td>
<td>3</td>
<td>21</td>
<td>40</td>
<td>0</td>
<td>10</td>
<td>35</td>
</tr>
<tr>
<td>8</td>
<td>NC</td>
<td>2</td>
<td>15</td>
<td>100</td>
<td>7</td>
<td>7</td>
<td>100</td>
</tr>
<tr>
<td>9</td>
<td>NC</td>
<td>5</td>
<td>16</td>
<td>71</td>
<td>0</td>
<td>29</td>
<td>88</td>
</tr>
<tr>
<td>10</td>
<td>TX</td>
<td>1</td>
<td>15</td>
<td>100</td>
<td>6</td>
<td>6</td>
<td>100</td>
</tr>
<tr>
<td>11</td>
<td>TX</td>
<td>2</td>
<td>17</td>
<td>100</td>
<td>0</td>
<td>11</td>
<td>100</td>
</tr>
<tr>
<td>12</td>
<td>TX</td>
<td>3</td>
<td>18</td>
<td>100</td>
<td>0</td>
<td>17</td>
<td>94</td>
</tr>
<tr>
<td>13</td>
<td>TX</td>
<td>4</td>
<td>18</td>
<td>100</td>
<td>38</td>
<td>5</td>
<td>100</td>
</tr>
<tr>
<td>14</td>
<td>TX</td>
<td>5</td>
<td>20</td>
<td>100</td>
<td>0</td>
<td>19</td>
<td>95</td>
</tr>
<tr>
<td>15</td>
<td>IN</td>
<td>1</td>
<td>23</td>
<td>9</td>
<td>0</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>IN</td>
<td>1</td>
<td>22</td>
<td>22</td>
<td>9</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>IN</td>
<td>1</td>
<td>22</td>
<td>0</td>
<td>4</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td><strong>Means</strong></td>
<td></td>
<td></td>
<td><strong>19.48</strong></td>
<td><strong>77</strong></td>
<td><strong>13</strong></td>
<td><strong>9</strong></td>
<td><strong>71</strong></td>
</tr>
</tbody>
</table>

**Control (n = 22)**

<table>
<thead>
<tr>
<th>Randomized Block</th>
<th>State</th>
<th>Grade</th>
<th>Class Size</th>
<th>% Free Lunch</th>
<th>% ESL</th>
<th>% Special Education</th>
<th>% Minority</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ID</td>
<td>1</td>
<td>24</td>
<td>72</td>
<td>33</td>
<td>17</td>
<td>56</td>
</tr>
<tr>
<td>2</td>
<td>ID</td>
<td>2</td>
<td>21</td>
<td>70</td>
<td>26</td>
<td>0</td>
<td>48</td>
</tr>
<tr>
<td>3</td>
<td>ID</td>
<td>3</td>
<td>22</td>
<td>72</td>
<td>0</td>
<td>4</td>
<td>68</td>
</tr>
<tr>
<td>4</td>
<td>ID</td>
<td>4</td>
<td>19</td>
<td>71</td>
<td>33</td>
<td>0</td>
<td>48</td>
</tr>
<tr>
<td>5</td>
<td>FL</td>
<td>1</td>
<td>21</td>
<td>28</td>
<td>4</td>
<td>16</td>
<td>28</td>
</tr>
<tr>
<td>6</td>
<td>FL</td>
<td>2</td>
<td>16</td>
<td>68</td>
<td>0</td>
<td>16</td>
<td>74</td>
</tr>
<tr>
<td>7</td>
<td>FL</td>
<td>3</td>
<td>18</td>
<td>33</td>
<td>0</td>
<td>24</td>
<td>17</td>
</tr>
<tr>
<td>8</td>
<td>NC</td>
<td>2</td>
<td>13</td>
<td>80</td>
<td>13</td>
<td>7</td>
<td>87</td>
</tr>
<tr>
<td>9</td>
<td>NC</td>
<td>5</td>
<td>18</td>
<td>61</td>
<td>6</td>
<td>6</td>
<td>89</td>
</tr>
<tr>
<td>10</td>
<td>TX</td>
<td>1</td>
<td>15</td>
<td>100</td>
<td>63</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>11</td>
<td>TX</td>
<td>2</td>
<td>18</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>12</td>
<td>TX</td>
<td>3</td>
<td>17</td>
<td>100</td>
<td>0</td>
<td>5</td>
<td>95</td>
</tr>
<tr>
<td>13</td>
<td>TX</td>
<td>4</td>
<td>19</td>
<td>100</td>
<td>0</td>
<td>14</td>
<td>100</td>
</tr>
<tr>
<td>14</td>
<td>TX</td>
<td>5</td>
<td>18</td>
<td>100</td>
<td>0</td>
<td>10</td>
<td>95</td>
</tr>
<tr>
<td>15</td>
<td>IN</td>
<td>1</td>
<td>23</td>
<td>8</td>
<td>0</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>IN</td>
<td>1</td>
<td>24</td>
<td>13</td>
<td>9</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>IN</td>
<td>1</td>
<td>22</td>
<td>4</td>
<td>4</td>
<td>9</td>
<td>74</td>
</tr>
</tbody>
</table>
## TABLE 2
Comparison of Baseline Characteristics for Open Court Treatment Classrooms and Control Classrooms

<table>
<thead>
<tr>
<th>Variable</th>
<th>Condition</th>
<th>$N$</th>
<th>$M$</th>
<th>$SD$</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
<th>$t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading Comprehension</td>
<td>Open Court</td>
<td>27</td>
<td>591.11</td>
<td>41.72</td>
<td>-21.34</td>
<td>25.12</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>22</td>
<td>589.22</td>
<td>38.25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading Vocabulary</td>
<td>Open Court</td>
<td>27</td>
<td>562.95</td>
<td>53.83</td>
<td>-28.84</td>
<td>31.49</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>22</td>
<td>561.63</td>
<td>50.12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading Composite</td>
<td>Open Court</td>
<td>27</td>
<td>577.31</td>
<td>47.01</td>
<td>-24.79</td>
<td>27.86</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>22</td>
<td>575.78</td>
<td>43.71</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Minority</td>
<td>Open Court</td>
<td>27</td>
<td>0.71</td>
<td>0.38</td>
<td>-0.24</td>
<td>0.17</td>
<td>-0.33</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>22</td>
<td>0.74</td>
<td>0.32</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% English as a Second Language</td>
<td>Open Court</td>
<td>27</td>
<td>0.13</td>
<td>0.26</td>
<td>-0.09</td>
<td>0.17</td>
<td>0.61</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>22</td>
<td>0.09</td>
<td>0.16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Special Education</td>
<td>Open Court</td>
<td>27</td>
<td>0.09</td>
<td>0.09</td>
<td>-0.02</td>
<td>0.07</td>
<td>1.05</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>22</td>
<td>0.06</td>
<td>0.07</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Free Lunch</td>
<td>Open Court</td>
<td>27</td>
<td>0.77</td>
<td>0.33</td>
<td>-0.18</td>
<td>0.19</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>22</td>
<td>0.76</td>
<td>0.30</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Value of Randomization

- Few differences between groups
- But done at classroom level
- Must prevent teachers from talking to each other
<table>
<thead>
<tr>
<th>Grade</th>
<th>Posttest</th>
<th>Open Court</th>
<th></th>
<th>Control</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>CTBS/5 Scale Score</td>
<td>N</td>
<td>CTBS/5 Scale Score</td>
</tr>
<tr>
<td></td>
<td>Classes</td>
<td>Students</td>
<td>( M )</td>
<td>Classes</td>
<td>Students</td>
</tr>
<tr>
<td>1</td>
<td>Reading Composite</td>
<td>9</td>
<td>165</td>
<td>575.79</td>
<td>37.19</td>
</tr>
<tr>
<td></td>
<td>Vocabulary</td>
<td>9</td>
<td>165</td>
<td>563.59</td>
<td>46.01</td>
</tr>
<tr>
<td></td>
<td>Reading Comprehension</td>
<td>9</td>
<td>165</td>
<td>587.46</td>
<td>36.15</td>
</tr>
<tr>
<td>2</td>
<td>Reading Composite</td>
<td>6</td>
<td>121</td>
<td>610.01</td>
<td>37.50</td>
</tr>
<tr>
<td></td>
<td>Vocabulary</td>
<td>6</td>
<td>121</td>
<td>596.74</td>
<td>43.41</td>
</tr>
<tr>
<td></td>
<td>Reading Comprehension</td>
<td>6</td>
<td>121</td>
<td>622.74</td>
<td>38.51</td>
</tr>
<tr>
<td>3</td>
<td>Reading Composite</td>
<td>5</td>
<td>93</td>
<td>642.44</td>
<td>45.35</td>
</tr>
<tr>
<td></td>
<td>Vocabulary</td>
<td>5</td>
<td>93</td>
<td>633.18</td>
<td>48.64</td>
</tr>
<tr>
<td></td>
<td>Reading Comprehension</td>
<td>5</td>
<td>93</td>
<td>651.17</td>
<td>47.14</td>
</tr>
<tr>
<td>4</td>
<td>Reading Composite</td>
<td>4</td>
<td>77</td>
<td>637.23</td>
<td>39.33</td>
</tr>
<tr>
<td></td>
<td>Vocabulary</td>
<td>4</td>
<td>77</td>
<td>631.08</td>
<td>41.98</td>
</tr>
<tr>
<td></td>
<td>Reading Comprehension</td>
<td>4</td>
<td>81</td>
<td>641.42</td>
<td>41.79</td>
</tr>
<tr>
<td>5</td>
<td>Reading Composite</td>
<td>3</td>
<td>51</td>
<td>661.65</td>
<td>33.19</td>
</tr>
<tr>
<td></td>
<td>Vocabulary</td>
<td>3</td>
<td>51</td>
<td>660.18</td>
<td>40.25</td>
</tr>
<tr>
<td></td>
<td>Reading Comprehension</td>
<td>3</td>
<td>51</td>
<td>662.53</td>
<td>33.04</td>
</tr>
<tr>
<td>All Grades</td>
<td>Reading Composite</td>
<td>27</td>
<td>507</td>
<td>614.15</td>
<td>49.36</td>
</tr>
<tr>
<td></td>
<td>Vocabulary</td>
<td>27</td>
<td>507</td>
<td>604.23</td>
<td>55.85</td>
</tr>
<tr>
<td></td>
<td>Reading Comprehension</td>
<td>27</td>
<td>511</td>
<td>623.46</td>
<td>48.11</td>
</tr>
</tbody>
</table>
Table 4
Multilevel Models Predicting Student, Classroom-Level, and Block-Level Literacy Outcomes for the Experimental Sample

<table>
<thead>
<tr>
<th>Fixed Effect</th>
<th>Reading Composite</th>
<th>Literacy Outcomes</th>
<th>Reading Comprehension</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Effect</td>
<td>SE</td>
<td>t</td>
</tr>
<tr>
<td>Average classroom Mean</td>
<td>610.82</td>
<td>1.74</td>
<td>351.08***</td>
</tr>
<tr>
<td>Average classroom OCR impact</td>
<td>7.95</td>
<td>1.83</td>
<td>4.34*</td>
</tr>
<tr>
<td>Average classroom pretest covariate effect</td>
<td>0.72</td>
<td>0.03</td>
<td>22.86***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Random Effect</th>
<th>Estimate</th>
<th>$\chi^2$</th>
<th>df</th>
<th>Estimate</th>
<th>$\chi^2$</th>
<th>df</th>
<th>Estimate</th>
<th>$\chi^2$</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students (Level 1)</td>
<td>1327.93</td>
<td></td>
<td></td>
<td>1798.77</td>
<td></td>
<td></td>
<td>1420.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Classrooms (Level 2)</td>
<td>3.23</td>
<td>49.83</td>
<td>18</td>
<td>24.32</td>
<td>54.56</td>
<td>18</td>
<td>18.37</td>
<td>59.11</td>
<td>18</td>
</tr>
<tr>
<td>Blocks (Level 3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept ($u_{000}$)</td>
<td>21.82</td>
<td>29.45</td>
<td>14</td>
<td>38.96</td>
<td>26.59</td>
<td>14</td>
<td>12.18</td>
<td>22.26</td>
<td>14</td>
</tr>
<tr>
<td>OCR impact ($u_{020}$)</td>
<td>7.10</td>
<td>12.05</td>
<td>14</td>
<td>97.71</td>
<td>21.99</td>
<td>14</td>
<td>0.82</td>
<td>10.16</td>
<td>14</td>
</tr>
</tbody>
</table>

Note: * $p < .05$; ** $p < .01$; *** $p < .001$. 
Everett Rogers’ Diffusion of Innovations (1995)
Diffusion: Beneath the Surface


NPR Science Friday
Diffusion: Beneath the Surface: Entering the System
Penetrating the Boundary
Absorbed by the System
System Adaptation
Internal System Reaction
Counteraction
How Does the Social Organization of the School Create a Complex System?

Social organization of the school (beneath the surface) → complex response