WHAT STARTS HERE CHANGES THE WORLD
THE UNIVERSITY OF TEXAS AT AUSTIN
COLLEGE of LIBERAL ARTS

A Treatment is a Treatment is a Treatment?

The Importance of Context in the Evaluation of Educational Interventions

Fang He
(US Government Accountability Office)
Leigh L. Linden
(The University of Texas at Austin)
Margaret MacLeod
(US Department of State)

Quality of Education in Developing Countries

- The quality of education in developing countries is poor
- Particularly important as enrollment levels increase
- To ameliorate this problem, we need to understand how to change the educational context to make it more productive
  - Economics: education production function
- Impact evaluations of individual interventions directed at identifying viable improvements.
  - RCTs being the “gold standard”

Effects of CAL/CAI Programs on Primary Subject

Positive	None	Negative
0.17, Barrow et al. (2009)	-0.24, Angrist and Lavy (2002)	-0.57, Linden (2008)
0.28, Linden (2008)	-0.08, Angrist and Lavy (2002)	
0.28, He et al. (2008)	0.02, Campuzano et al. (2009)	
0.36 He et al. (2008)	0.03, Campuzano et al. (2009)	
0.47, Bannerjee et al. (2007)	0.07, Kreuger and Rouse (2004)	
Machin et al. (2007)	0.08, Barrera and Linden (2009)	

Instruction in a Single Subject

Simple typology of components of an educational context:

- Students: Who is going to learn?
  - Intelligence, educational experiences, demographics, etc.
- Content: What information should students learn?
  - Number identification, single digit addition, etc.
- Pedagogy: How will the information be taught?
  - Experience of examples, rote memorization, etc.
- Resources: What is available to do the teaching?
  - Teacher, computers, desks, classroom, etc.
- Organization: How are the resources organized to teach students information using the pedagogical strategy?
  - Teaching schedule, length of school day, class size, etc

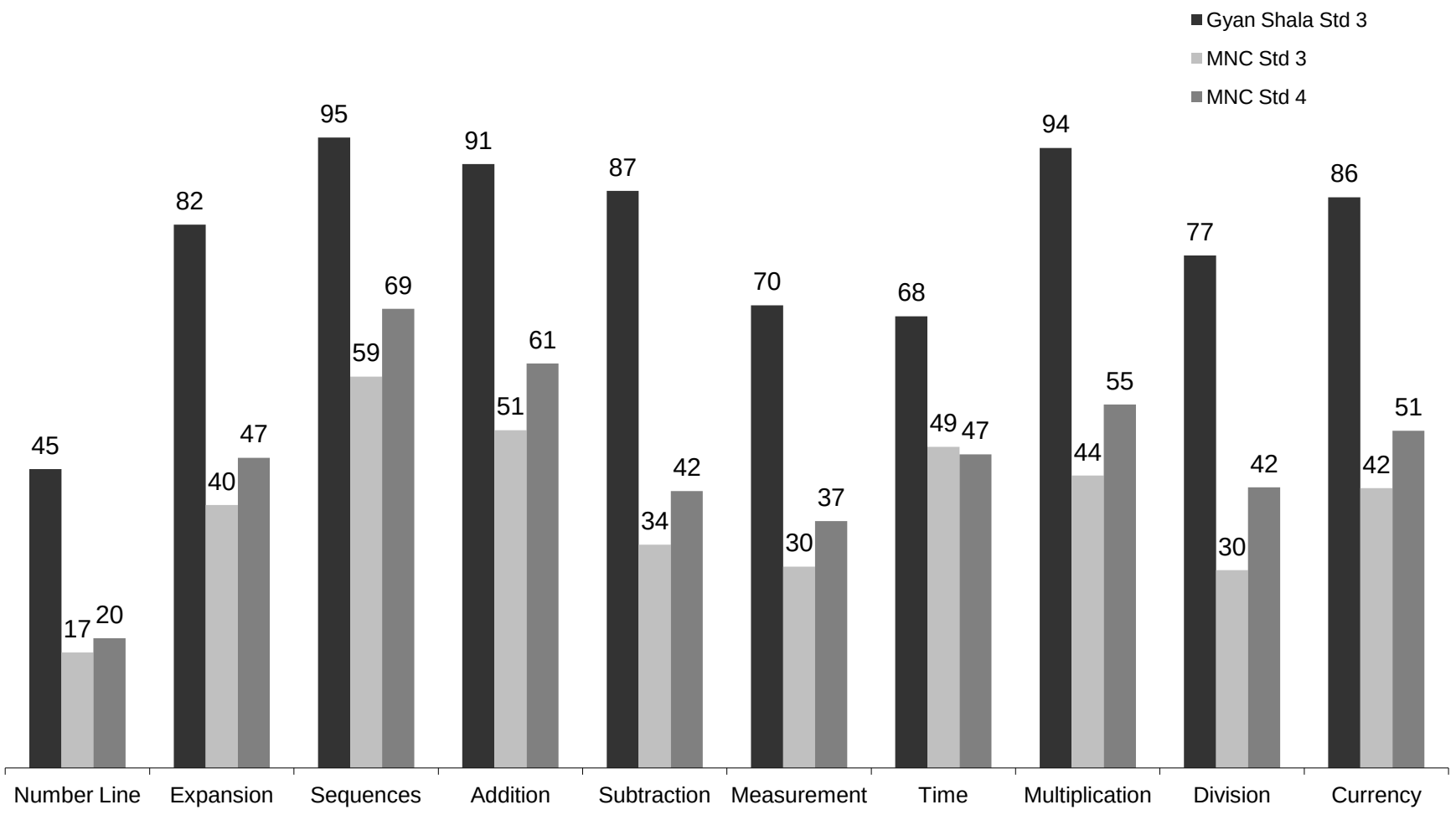
What then is a Treatment?

- Educational “Treatments” then embody manipulations of multiple components.
  - Some of these changes are integral to the treatment.
    - Computers have technological constraints
  - Some changes are optional.
    - Pull-out versus add-on model
  - Some changes are facilitated by the treatment.
    - More targeted teaching methods with ability grouping
- Claim: These individual components can significantly alter the efficacy of an intervention.
- If the changes to various components matter:
  - Difficult to consider “Treatments” as monolithic changes
  - Raises difficult questions about evaluation practice
    - Impractical to test every variation of the various components
    - Generalizability of results

Study Overview

- Demonstrate that these components matter significantly
- Present results from two studies
  - Each varies the treatment components
  - Different arrangements have different effects
- Studies:
  - Gyan Shala CAL Program
  - Pratham PicTalk Program

Gyan Shala


Gyan Shala Std 3
MNC Std 3
MNC Std 4
45
17
20
82
40
47
95
59
69
91
51
61
87
34
42
70
30
37
68
49
47
94
44
55
77
30
42
86
42
51
Number Line
Expansion
Sequences
Addition
Subtraction
Measurement
Time
Multiplication
Division
Currency

Gyan Shala CAL

- Focused on grades 2 and 3
- Designed to reinforce math concepts
- Carefully integrated into highly structured curriculum
  - Students complete practice worksheets
- One hour of computer time a day
  - Pull-Out Model
  - Out-of-School Time Model

Gyan Shala: Sample

Unit of Analisys	Treatment	Control	Total
Schools in In-School Program Experiment (Year 1)	11	12	23
Schools in Out-of-School Program Experiment (Year 2)	19	18	37
Total Villages	**30**	**30**	**60**
Children in In-School Program Experiment (Year 1)	392	387	779
Children in Out-of-School Program Experiment (Year 2)	682	695	1377
Total Children	**1074**	**1082**	**2156**

Gyan Shala: Math Scores

	Control Average	Difference w/ Controls
Combined Programs	-0.001	-0.06
	(0.035)	(0.154)
Pull-Out Program (Year 1)	0.006	-0.656**
	(0.059)	(0.264)
Out-of-School Program (Year 2)	-0.005	0.332**
	(0.044)	(0.162)

Gyan Shala: Math Scores

	Pull-Out Program (Y1)		OST Program (Y2)	
Sub-Sample	**Control Average**	**Difference w/ Controls**	**Control Average**	**Difference w/ Controls**
Lowest Tercile	-0.661	-0.847**	-0.514	0.472**
	(0.136)	(0.368)	(0.087)	(0.194)
Middle Tercile	-0.018	-0.622**	-0.102	0.223
	(0.087)	(0.246)	(0.067)	(0.197)
Top Tercile	0.419	-0.426	0.434	0.122
	(0.072)	(0.254)	(0.063)	(0.193)

Pratham English Language Programs

- English language instruction in India is particularly poor
  - Teachers have limited proficiency
  - 61 pct of 1st graders do not recognize capital letters
  - 28 pct of 5th graders can read a simple sentence (ASER, 2007)
- Change in pedagogy:
  - De-emphasize rote memorization
  - Teach grammar, vocabulary, and pronunciation simultaneously.
- Two implementation technologies:
  - PicTalk machine
    - Based on the Interactive Paper Technology developed by LeapFrog Enterprises
  - Activity cards
    - Set of 440 flashcards with teachers manual
    - Guides teacher through interactive drills and games

PicTalk Machine

Pratham: Research Design

- External Implementation (2005-06), Thane
  - Grades 2 and 3
  - Both interventions provided on alternating days
  - Implemented by externally hired and managed assistants
  - Randomized within school by grade, stratified by test score
- Teacher Implementation (2006-07), Mangaon
  - Grades 1-5
  - Teachers attended a 5-day training
  - Had access to Pratham monitors
  - Three interventions:
    - PicTalk and Activities, same as Thane
    - PicTalk only
    - Activities only
  - Randomized by school, stratified by enrollment

Pratham: Research Design

- Variations in the Treatment
  - Combined Treatment:
    - Vary the management of the intervention
    - External vs. Teacher Implementation
  - Delivery Mechanism:
    - Individual, self-paced machine
    - Group, teacher directed activities

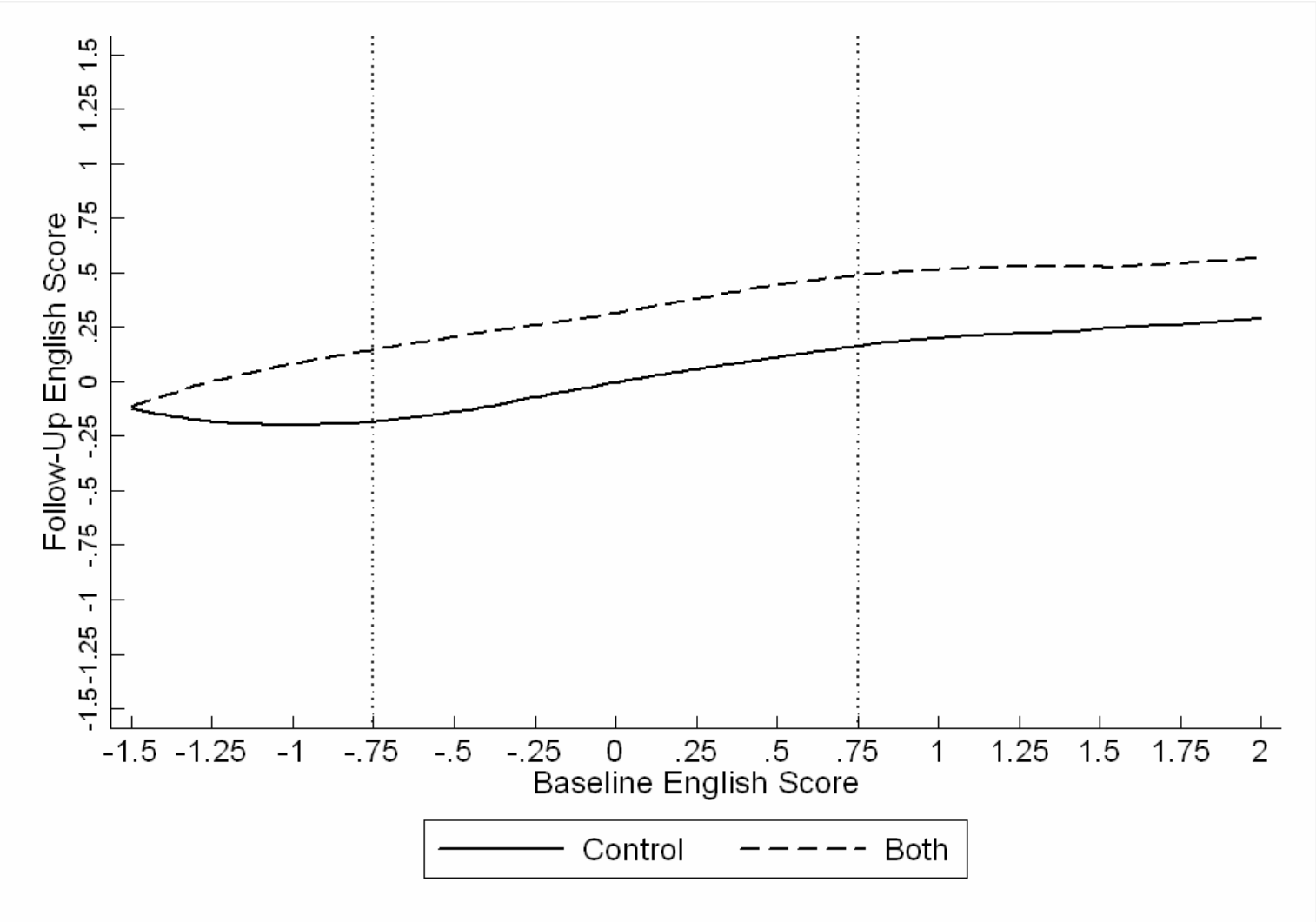
Pratham: Research Sample

	External (Year 1)		Teacher Implementation (Year 2)			
Sample Description	**Control Group**	**Machines and Activities Group**	**Control Group**	**Machines and Activities Group**	**Machines Group**	**Activities Group**
Number of Schools[†]			61	61	61	60
Number of Classes	97	97	253	254	254	246
Number of Students	2618	2699	2458	2514	2449	2324

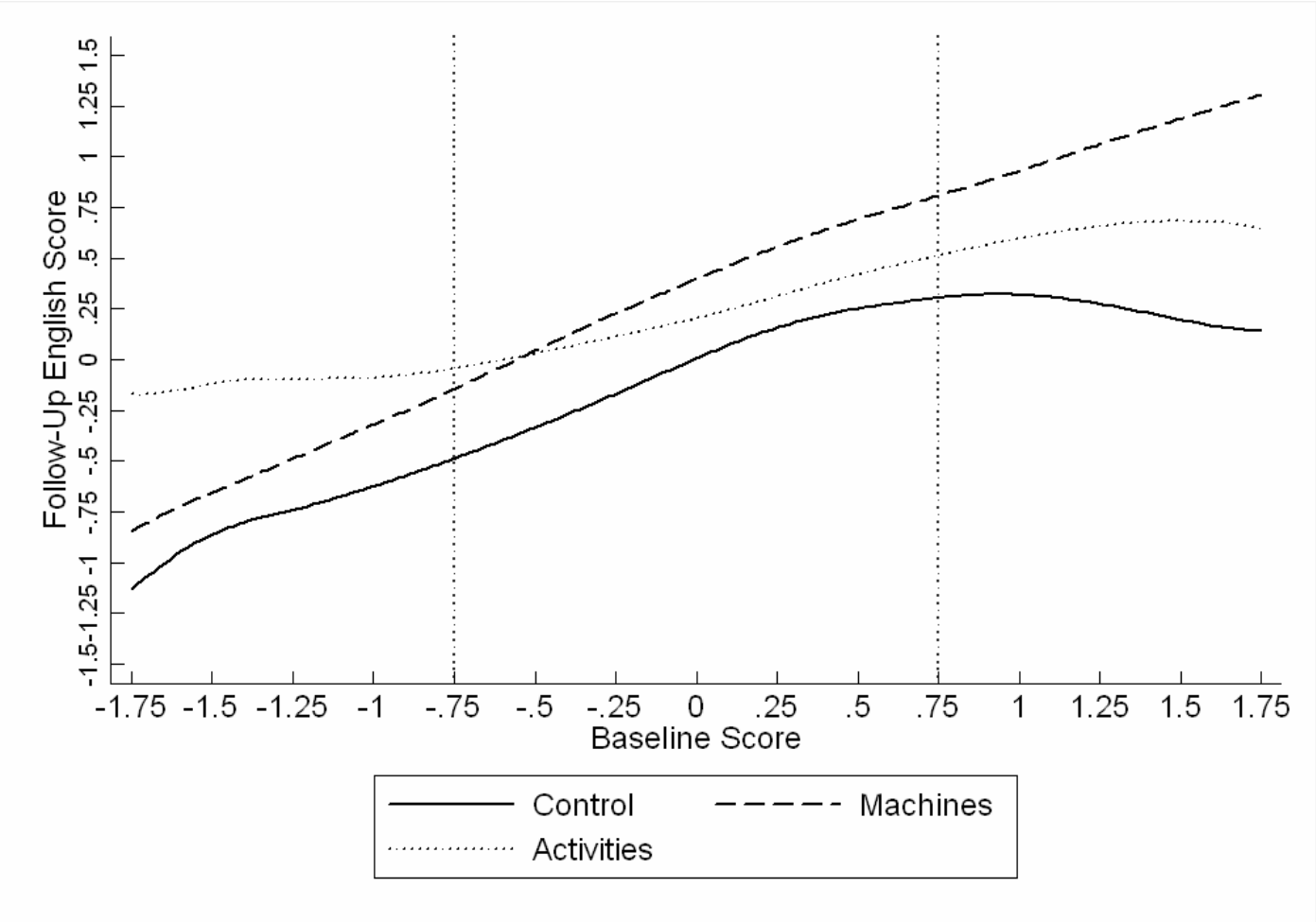
Pratham: English Scores

Competency	All Treat-Control	Both - Control	Machines-Control	Activities-Control
Panel A: External Implementation (Year 1)				
Section 1 Total		0.297***		
		(0.090)		
Section 2 Total		0.229**		
		(0.092)		
English Total		**0.278*** **		
		(0.093)		
Panel B: Teacher Implementation (Year 2)				
Section 1 Total	0.316**	0.269*	0.323**	0.309**
	(0.125)	(0.142)	(0.139)	(0.147)
Section 2 Total	0.305**	0.304*	0.284*	0.25
	(0.150)	(0.168)	(0.148)	(0.154)
English Total	**0.355****	**0.328****	**0.345****	**0.320****
	(0.149)	**(0.165)**	**(0.152)**	**(0.160)**

Heterogeneity: External Implementation



Heterogeneity: Teacher Implementation



Heterogeneity: Teacher Implementation

Characteristics	External	Teacher Implementaion		
	Both - Control	Both - Control	Machine-Control	Activities-Control
Entire Sample	0.287***	0.328**	0.345**	0.320**
	-0.092	(0.165)	(0.152)	(0.160)
Baseline < -0.75	0.241*	0.529***	0.096	0.521***
	-0.126	(0.152)	(0.174)	(0.185)
-0.75 > Baseline < 0.75	0.317***	0.347**	0.247	0.211
	-0.096	(0.162)	(0.153)	(0.146)
Baseline > 0.75	0.302**	0.072	0.646***	0.215
	-0.132	(0.204)	(0.193)	(0.225)

Heterogeneity: Teacher Implementation

Activities Only Relative to Machine Only

	Activities Only	
	(3)	(4)
Treat*(Baseline Eng < -0.75)	0.566***	
	(0.190)	
Treat*(-0.75 < Baseline Eng < 0.75)	-0.109	
	(0.151)	
Treat*(Baseline Eng > 0.75)	-0.401**	
	(0.181)	
Treat		-0.026
		(0.137)
Treat*Baseline Eng		-0.358***
		(0.110)

Heterogeneity

- All students benefit from each intervention, but…
  - Stronger students benefit more from machines
  - Weaker students benefit more from activities
- Mirrors differences in the underlying learning process
  - Machines: Individualized instruction
    - Students who can figure things out on their own can move at their own pace.
    - Weaker students who require help need to wait on individual assistance from the teacher.
  - Activities: Group instruction
    - Stronger students held back because lesson is not tailored to their pace.
    - Weaker students may benefit from teachers ability to tailor interactions to their level of understanding.
- Note that machine effects differ from Gyan Shala
  - Remedial program vs. self-paced program

Indirect Effects on Other Subjects

- External implementation was carefully controlled
- No control over the teacher implementation
  - Monitored activities by asking children about lesson on the previous day of instruction
  - Compromise to avoid directly monitoring teachers
  - Two challenges
    - These children are young, response bias
    - Difficult to distinguish Pratham activities from regular activities
- Implementation closely track treatment assignment
- Implemented less frequently than in external implementation
  - Did teachers efficiently switch to teaching other subjects?

Indirect Effects: Math Scores

	All Treat-Control	Both - Control	Machines-Control	Activities-Control
Panel A: Without Controls				
External Implementation		0.038		
		(0.075)		
Teacher Implementation	0.301***	0.274**	0.388***	0.250*
	(0.106)	(0.135)	(0.142)	(0.137)
Panel B: With Controls				
External Implementation		0.052		
		(0.071)		
Teacher Implementation	0.341***	0.391***	0.284**	0.319***
	(0.087)	(0.103)	(0.117)	(0.102)

Indirect Effects: Utilization

English Activity	Control Group	Both Treatments	Machine Group	Activities Group
Average Utilization (Treatment Group Relative to Control)				
English Classes	0.945	-0.006	-0.007	0.002
		(0.019)	(0.018)	(0.018)
Textbook	0.978	0	-0.011	-0.006
		(0.010)	(0.013)	(0.013)
Chalkboard	0.969	0.006	-0.018	-0.021
		(0.012)	(0.017)	(0.017)
Machine	0.01	0.710***	0.635***	0.019
		(0.031)	(0.026)	(0.014)
Activities	0.165	0.564***	-0.005	0.546***
		(0.076)	(0.077)	(0.077)

Conclusions

- Components of individual treatment matter
- Important to note:
  - Different from just a question of generalizability (but related)
  - Not a limitation of any particular evaluation method
- Implications for evaluation practice
  - Can we deem an intervention effective with only one style of implementation?
  - How many and which variations to test?
- Possible avenues forward
  - Thoroughly describe evaluations.
    - Nature of the treatment
    - Nature of context
    - Specific changes to the context
  - Develop more general theory of education