

The Behavioralist Goes to School: Field Experiments in Education

Sally Sadoff, UC San Diego

Steven Levitt, University of Chicago and NBER

John List, University of Chicago and NBER

Susanne Neckermann, Erasmus University of Rotterdam and
ZEW

Roland Fryer, Harvard University and NBER

August 31, 2013

Given the large and increasing returns to education, why is achievement among many students so low?

- Cannot be fully explained by the standard model in which human capital investment depends on (time consistent) discounted returns net of costs

Introduction: Research Program

- Develop general framework for the education production function drawing on insights from behavioral economics
- Use framework to motivate the design of interventions to improve achievement
- Test behavioral interventions using randomized field experiments
 - Short-term student incentives
 - Teacher performance pay

Introduction: Laboratory for Urban School Reform

	Chicago Heights¹	City of Chicago¹	Urban School Districts²
Enrollment	6,534	404,589	142,520 [†]
Per Pupil Expenditure	\$12,464	\$12,880	\$13,958 [†]
% Black	49%	45%	30% [†]
% Hispanic	37%	42%	41% [†]
% White	10%	9%	19% [†]
% Low Income	83%	87%	69% ^{††}
% Meet or Exceed Elementary Standards	64%	68% ^a	62% ^{††}
Graduation Rate	47% ³	51% ²	53% ⁴

¹Illinois Report Card 2010 (a) 2009

²Institute of Education Sciences 2009 (Sample of 40(†)/20(††) large urban school districts)

³Chicago Heights Promise Working Group 2008

⁴Education Research Center 2009 (School districts in 50 largest U.S. cities)

Structure of education investment function

- Investment requires exerting effort on tasks with low returns in the near term (e.g., pay attention in class, focus on a low stakes test)
- Returns to effort occur with a long delay

Structure of education investment function

- Investment requires exerting effort on tasks with low returns in the near term (e.g., pay attention in class, focus on a low stakes test)
- Returns to effort occur with a long delay
- \Rightarrow Isolate role of effort

Student Incentives

Potential sources of low investment

- Undervalue returns to education
- High discount rates
- Limited understanding of education production function

Potential sources of low investment

- Undervalue returns to education
 - \Rightarrow Loss framing
 - \Rightarrow Non-financial rewards
- High discount rates
- Limited understanding of education production function

Potential sources of low investment

- Undervalue returns to education
 - \Rightarrow Loss framing
 - \Rightarrow Non-financial rewards
- High discount rates
 - \Rightarrow Immediate rewards
- Limited understanding of education production function

Potential sources of low investment

- Undervalue returns to education
 - \Rightarrow Loss framing
 - \Rightarrow Non-financial rewards
- High discount rates
 - \Rightarrow Immediate rewards
- Limited understanding of education production function
 - \Rightarrow Vary size of reward

- Offer students a reward for improved score on a low stakes diagnostic test
- Announced immediately before test begins and distributed immediately after test ends (except delayed rewards)
- Collect scores from subsequent tests in order to test for crowding out of intrinsic motivation from the incentive

Experimental Design

	Bloom 10th grade	Chicago Heights 3rd - 8th grade	CPS 2nd - 8th grade
Financial Low (\$10)	◆	◆	◆
Financial High (\$20)	◆	◆	◆
Non-Financial (Trophy)		◆	◆
Financial Loss	◆		◆
Non-Financial Loss			◆
Delayed Incentives			◆
Test	Reading ~15 minutes	Math ~30 minutes	Reading & Math ~60 minutes
Level of Randomization	Class	School-Grade	School-Grade

Results: Effects of Immediate Incentives

Treatment Effects on Test Score Improvement		
	(1)	(2)
Financial Low	0.005 (0.046)	-0.001 (0.045)
Financial High	0.113*** (0.038)	0.104*** (0.032)
Non-Financial	0.045 (0.041)	0.091** (0.040)
Financial Loss	0.132*** (0.042)	0.129*** (0.036)
Non-Financial Loss	0.095** (0.043)	0.109*** (0.041)
Session	Yes	Yes
Other Covariates		Yes
Observations	6844	6844
Classes/School-Grades	226	226

The table reports OLS estimates in standard deviation units. Standard errors clustered by class (Bloom) or school-grade (CH and CPS) are reported in parentheses. Other covariates include: standardized baseline score on the tested subject (score, score squared and score cubed), past treatment, school, grade, teacher, test subject and demographics (gender, race/ethnicity, free/reduced lunch status and IEP status), where applicable. Asterisks indicate significance at *0.1, **0.05, ***0.001

Results: Effects of Delayed Incentives

Treatment Effects on Test Score Improvement		
	CPS	
Delayed Financial High	-0.078 (0.095)	-0.071 (0.110)
Delayed Non-Financial	-0.171* (0.093)	-0.121 (0.087)
Delayed Financial Loss	-0.005 (0.086)	0.154 (0.149)
Delayed Non-Financial Loss	-0.096* (0.054)	-0.088 (0.092)
Immediate Incentive Treatments	Yes	Yes
Other Covariates		Yes
Observations	1362	1362
School-Grades	81	81

The table reports OLS estimates in standard deviation units. Standard errors clustered by school-grade are reported in parentheses. Other covariates include: standardized baseline score on the tested subject (score, score squared and score cubed), school, grade, test subject and demographics (gender, race/ethnicity, free/reduced lunch status and IEP status). Asterisks indicate significance at *0.1, **0.05, ***0.001

Results: Effects of Age

Treatment Effects on Test Score Improvement: By Grade

	<i>All Students</i>	Elementary	Middle/ Secondary	p-value
Financial Low	0.002 (0.044)	0.145 (0.115)	-0.106** (0.050)	0.042
Financial High	0.103*** (0.032)	0.126** (0.053)	0.059 (0.036)	0.275
Non-Financial	0.091** (0.040)	0.146** (0.069)	0.093 (0.068)	0.578
Financial Loss	0.127*** (0.036)	0.122** (0.054)	0.155*** (0.039)	0.616
Non-Financial Loss	0.108*** (0.041)	0.235*** (0.052)	-0.035 (0.053)	0.000
Session	Yes	Yes	Yes	
Other Covariates	Yes	Yes	Yes	
Observations	6844	3462	3382	
Classes/School-Grades	226	118	122	

The table reports OLS estimates in standard deviation units. Standard errors clustered by class (Bloom) or school-grade (CH and CPS) are reported in parentheses. Other covariates include: standardized baseline score on the tested subject (score, score squared and score cubed), past treatment, school, grade, teacher, test subject and demographics (gender, race/ethnicity, free/reduced lunch status and IEP status), where applicable. Asterisks indicate significance at *0.1, **0.05, ***0.001

Results: Crowding Out

Treatment Effects on Future Test Score Improvement

	Subsequent Subject Same Test Session		Same Subject Future Test Session	
Financial Low	0.035 (0.074)	-0.107 (0.083)	-0.136 (0.084)	-0.044 (0.075)
Financial High	0.049 (0.060)	-0.011 (0.045)	-0.024 (0.061)	-0.007 (0.057)
Non-Financial	-0.028 (0.073)	-0.033 (0.050)	-0.056 (0.068)	-0.011 (0.044)
Financial Loss	0.036 (0.069)	0.052 (0.046)	0.069 (0.053)	0.026 (0.047)
Non-Financial Loss	0.009 (0.055)	0.100** (0.047)	0.037 (0.067)	0.086* (0.047)
Session	Yes	Yes	Yes	Yes
Subsequent Treatment	Yes	Yes		
Other Covariates		Yes		Yes
Observations	4238	4238	5502	5502
Classes/School-Grades	166	166	180	180

The table reports OLS estimates in standard deviation units. Standard errors clustered by class (Bloom) or school-grade (CH and CPS) are reported in parentheses. Asterisks indicate significance at *0.1, **0.05, ***0.001

Summary of Results

- Short term incentives can significantly improve performance solely through increasing effort
- Effectiveness does not depend on the framing of rewards
- Non-financial incentives are considerably more cost-effective for younger students but not effective for older students
- Consistent with hyperbolic discounting, delayed incentives are not effective
- Students understand how to respond to the incentive but require sufficient motivation
- Little evidence of crowding out

Implications

- In the absence of incentives many students may exert low effort on standardized tests, which may bias estimates of achievement gaps.
- Identifying effort gaps vs. ability gaps can help shape interventions
- The delay of educational returns may be leading to underinvestment in low stakes tasks that lead to human capital accumulation
- Short term rewards can help address low motivation, habit formation, planning failures and limited understanding of the production function

Teacher Incentives

- Good teachers matter
 - 1 s.d. improvement in teacher quality
 - → 0.15-0.24 s.d. improvement in test performance
 - → Increases lifetime earnings by about 1% per year
 - Comparable to reducing class size by about one-third
- Difficult to improve teacher quality
 - Screening, training, etc. are largely ineffective
- Mixed evidence on effectiveness of performance pay
 - Large effects (0.12-0.27 s.d.) in developing countries where there is a low degree of professionalism
 - No significant effects in U.S.

Teacher Incentives

- Design high powered loss incentives
- Power of framing widely demonstrated in the lab; limited experimental evidence on loss aversion in the field.
 - Our student incentive study was run concurrently
 - Only previous field study finds framing bonuses as losses improves productivity of teams in a Chinese factory (Hossain and List 2009)
- As far as we know, no previous field study has tested loss aversion with actual endowments

Experimental Design: Program Details

- Chicago Heights K-8 schools (2010-2011 school year, no overlap with student incentives experiment)
- Classroom teachers and interventionists eligible (90% sign-up)
- Incentivized teachers receive rewards based on their students' improvement on a low stakes diagnostic test
- Based on pay for percentile measure (Barlevy & Neal 2011)
 - \$80 per percentile \Rightarrow Expected value = \$4,000
 - \sim 8% of average teacher salary
- Collect scores from unincentivized state standardized test to examine potential gaming

Experimental Design: Treatments

- “Gain” Treatment
 - Teacher receives her total reward at the end of the year
- “Loss” Treatment
 - Teacher receives \$4,000 at the *beginning* of the year
 - If her students’ end of the year performance is below average, teacher agrees to return the difference between \$4,000 and her final reward
 - If her final reward is above \$4,000, we issue an additional check for the difference

- Within “Gain” and “Loss”, we compare individual & team incentives
 - Individual: Reward based on individual performance
 - Team: Reward based on average of team performance (teachers paired by school, grade and subject)

Experimental Design: Randomization

- Randomize at the teacher level
- For treated teachers who rotate classes, randomize which classes are incentivized
 - Allows us to measure spillovers
- In reading, students are often exposed to multiple teachers and thus multiple treatments, confounding the analysis
 - We therefore focus on the math results

Experimental Design: Overview

Pooled Treatment	Treatment Arm	Timing	Basis for Rewards
<i>Pooled Loss</i>	Individual Loss	Receive \$4,000 at the <i>beginning</i> of the year. Must pay back difference at the end of the year.	Teacher's own students
	Team Loss		Teacher's & teammate's students
<i>Pooled Gain</i>	Individual Gain	Receive full reward at The end of the year.	Teacher's own students
	Team Gain		Teacher & teammate's students
<i>Pooled Control</i>	Pure Control	Teacher is not incentivized on any of her classes	
	Spillover Control	Teacher is incentivized on another class	

Results: Treatment Effects on Incentivized Test Scores

	Percentile Rank		Scaled Score	
	(1)	(2)	(3)	(4)
Pooled Loss	6.866** (2.677)	6.840*** (2.554)	0.222*** (0.070)	0.220*** (0.065)
Pooled Gain	1.263 (2.888)	1.884 (2.834)	0.078 (0.072)	0.092 (0.070)
Controls	No	Yes	No	Yes
Pr(Gain = Loss)	0.019	0.031	0.019	0.027
Observations	2311	2311	2311	2311
Homerooms	141	141	141	141

The table reports OLS estimates in standard deviation or percentile units. Standard errors clustered by homeroom are reported in parentheses. All regressions include dummy variables for school, grade and baseline test scores. Additional controls include: race/ethnicity, gender, age, free-lunch status, English proficiency and special education status. Asterisks indicate significance at *0.1, **0.05, ***0.001

Results: Treatment Effects on Unincentivized Test Scores

	Percentile Rank		Scaled Score	
	(1)	(2)	(3)	(4)
Pooled Loss	6.867** (3.269)	6.680** (3.194)	0.213*** (0.078)	0.398*** (0.076)
Pooled Gain	0.228 (3.402)	1.012 (3.351)	-0.013 (0.085)	-0.003 (0.084)
Controls	No	Yes	No	Yes
Pr(Gain = Loss)	0.017	0.042	0.003	0.006
Observations	2144	2144	2144	2144
Homerooms	140	140	140	140

The table reports OLS estimates in standard deviation or percentile units. Standard errors clustered by homeroom are reported in parentheses. All regressions include dummy variables for school, grade and baseline test scores. Additional controls include: race/ethnicity, gender, age, free-lunch status, English proficiency and special education status. Asterisks indicate significance at *0.1, **0.05, ***0.001

Alternative mechanisms to loss aversion

- Gaming
 - No evidence of attrition (e.g., of weaker students)
 - Equivalent effects on unincentivized test
- Changing the production function (e.g., easing credit constraints)
 - Loss and Gain teachers do not report different amounts of spending on students
 - At the end of March, 69% of Loss teachers report that they have spent none of the upfront payment

Implications

- The impact of introducing the Loss frame is roughly equivalent to increasing teacher quality by 1 s.d.
- We replicate null findings for Gain bonuses, implying that our results are due to the addition of framing rather than differences in setting, reward structure or implementation.
- The loss treatment seems to work largely through psychological effects rather than changing the production function.
- As in the study of student incentives, we demonstrate the potential for insights from behavioral economics to increase the efficacy of interventions.

Where We're Headed Next: Individualizing Interventions

Randomized experiments typically only recover the average treatment effect in a population.

We know that there is some distribution of treatment effects and that we would increase the (cost) effectiveness of our interventions if they could be targeted at individuals with the largest treatment effects.

Where We're Headed Next: Individualizing Interventions

- Develop methods for identifying who has the largest treatment effects
- Design interventions that allow for greater individualization
- Ongoing research
 - Teacher screening
 - Blended learning
 - High and low cost parent programs

Results: Balance

Summary Statistics by Treatment Group: Bloom

	Control	Financial Low	Financial High	Financial Loss	F-Test p-value
Observations	240	143	288	125	
Baseline Test Score	0.133 (0.965)	0.132 (0.893)	-0.074 (0.972)	0.289 (1.035)	0.543
Female	0.533 (0.500)	0.517 (0.501)	0.472 (0.500)	0.464 (0.501)	0.659
Black	0.585 (0.494)	0.413 (0.494)	0.586 (0.493)	0.449 (0.500)	0.029
Hispanic	0.314 (0.465)	0.448 (0.499)	0.331 (0.471)	0.390 (0.490)	0.284
Free or Reduced Price Lunch	0.729 (0.446)	0.720 (0.450)	0.737 (0.441)	0.822 (0.384)	0.097

The table reports group means. Standard deviations are reported in parentheses. Asterisks indicate a difference of means (compared to control with standard errors clustered by class) significant at *0.1, **0.05, ***0.001

Results: Balance

Summary Statistics by Treatment Group: Chicago Heights

	Control	Financial Low	Financial High	Non- Financial	F-Test p-value
Observations	160	165	29	69	
Baseline Test Score	-0.457 (0.772)	-0.510 (0.781)	-0.421 (1.078)	-0.682 (0.775)	0.025
Grade	6.200 (2.070)	5.133 (1.446)	5.414 (1.402)	5.072 (1.229)	0.803
Female	0.513 (0.501)	0.497 (0.502)	0.448 (0.506)	0.449 (0.501)	0.965
Black	0.513 (0.501)	0.461 (0.500)	0.310 (0.471)	0.290* (0.457)	0.368
Hispanic	0.363 (0.482)	0.461 (0.500)	0.621 (0.494)	0.623** (0.488)	0.038
Free or Reduced Price Lunch	0.863 (0.345)	0.891 (0.313)	0.897 (0.310)	0.928 (0.261)	0.204
Individualized Education Plan (IEP)	0.074 (0.263)	0.069 (0.255)	0.034 (0.186)	0.101 (0.304)	0.857

The table reports group means. Standard deviations are reported in parentheses. Asterisks indicate a difference of

Results: Balance

Summary Statistics by Treatment Group: CPS (Immediate Incentives)

	Control	Financial Low	Financial High	Non-Financial	Financial Loss	Non-Financial Loss	F-Test p-value
Observations	2910	105	644	545	715	706	
Baseline Test Score	-0.007 (0.882)	0.058 (0.871)	0.098 (0.852)	0.148* (0.898)	-0.026 (0.865)	0.051 (0.854)	0.616
Grade	5.109 (1.858)	5.619 (1.361)	5.351 (1.714)	5.209 (1.948)	4.752 (1.937)	5.021 (1.896)	0.977
Subject – Math	0.307 (0.461)	0.238 (0.428)	0.253 (0.435)	0.178 (0.383)	0.274 (0.446)	0.358 (0.480)	0.740
Female	0.508 (0.500)	0.686*** (0.466)	0.547 (0.498)	0.508 (0.500)	0.585*** (0.493)	0.482 (0.500)	0.000
Black	0.983 (0.130)	0.990 (0.098)	0.978 (0.146)	0.993** (0.085)	0.989 (0.105)	0.983 (0.129)	0.312
Free or Reduced Price Lunch	0.988 (0.108)	1.000*** (0.000)	0.989 (0.105)	0.984 (0.125)	0.992 (0.089)	0.988 (0.110)	0.932
Individualized Education Plan (IEP)	0.092 (0.289)	0.089 (0.286)	0.100 (0.300)	0.089 (0.285)	0.113 (0.317)	0.101 (0.301)	0.595

The table reports group means. Standard deviations are reported in parentheses. Asterisks indicate a difference of means (compared to control with standard errors clustered by school-grade) significant at *0.1, **0.05, ***0.001

Results: Balance

Summary Statistics by Treatment Group: CPS (Delayed Incentives)

	Control	Financial High	Non-Financial	Financial Loss	Non-Financial Loss	F-Test p-value
Observations	2910	85	82	30	74	
Baseline Test Score	-0.007 (0.882)	0.036 (0.816)	-0.036 (0.859)	-0.173 (0.862)	0.152 (0.950)	0.616
Grade	5.109 (1.858)	5.235 (0.811)	4.793 (1.577)	4.833 (1.599)	4.838 (2.054)	0.977
Subject – Math	0.307 (0.461)	0.353 (0.481)	0.171 (0.379)	0.000*** (0.000)	0.000*** (0.000)	0.740
Female	0.508 (0.500)	0.518 (0.503)	0.432* (0.498)	0.367 (0.490)	0.541 (0.502)	0.000
Black	0.983 (0.130)	0.976 (0.152)	0.988 (0.111)	0.933 (0.254)	1.000*** (0.000)	0.312
Free or Reduced Price Lunch	0.988 (0.108)	1.000*** (0.000)	1.000*** (0.000)	1.000*** (0.000)	1.000*** (0.000)	0.932
Individualized Education Plan (IEP)	0.092 (0.289)	0.050 (0.219)	0.123 (0.331)	0.069 (0.258)	0.048 (0.216)	0.595

The table reports group means. Standard deviations are reported in parentheses. Asterisks indicate a difference of means (compared to control with standard errors clustered by school-grade) significant at *0.1, **0.05, ***0.001

Results: Treatment Effects by Subject

Treatment Effects on Test Score Improvement: By Subject

	<i>All Students</i>	Math	Reading	p-value
Financial Low	0.002 (0.044)	0.145** (0.066)	-0.145*** (0.047)	0.000
Financial High	0.103*** (0.032)	0.279*** (0.069)	0.054 (0.034)	0.002
Non-Financial	0.091** (0.040)	0.082 (0.089)	0.076 (0.051)	0.942
Financial Loss	0.127*** (0.036)	0.272** (0.118)	0.076** (0.035)	0.099
Non-Financial Loss	0.108*** (0.041)	0.097 (0.060)	0.089* (0.050)	0.909
Other Covariates	Yes	Yes	Yes	
Observations	6844	3462	3382	
Classes/School-Grades	226	118	122	

The table reports OLS estimates in standard deviation units. Standard errors clustered by class (Bloom) or school-grade (CH and CPS) are reported in parentheses. Other covariates include: standardized baseline score on the tested subject (score, score squared and score cubed), past treatment, school, grade, teacher, test subject and demographics (gender, race/ethnicity, free/reduced lunch status and IEP status), where applicable. Asterisks indicate significance at *0.1, **0.05, ***0.001

Results: Treatment Effects by Gender

Treatment Effects on Test Score Improvement: By Gender

	<i>All Students</i>	Male	Female	p-value
Financial Low	0.002 (0.044)	0.066 (0.058)	-0.056 (0.055)	0.095
Financial High	0.103*** (0.032)	0.108** (0.044)	0.091*** (0.035)	0.731
Non-Financial	0.091** (0.040)	0.129** (0.050)	0.053 (0.049)	0.164
Financial Loss	0.127*** (0.036)	0.157*** (0.052)	0.097** (0.042)	0.309
Non-Financial Loss	0.108*** (0.041)	0.155** (0.060)	0.062 (0.039)	0.121
Session	Yes	Yes	Yes	
Other Covariates	Yes	Yes	Yes	
Observations	6844	3462	3382	
Classes/School-Grades	226	118	122	

The table reports OLS estimates in standard deviation units. Standard errors clustered by class (Bloom) or school-grade (CH and CPS) are reported in parentheses. Other covariates include: standardized baseline score on the tested subject (score, score squared and score cubed), past treatment, school, grade, teacher, test subject and demographics (gender, race/ethnicity, free/reduced lunch status and IEP status), where applicable. Asterisks indicate significance at *0.1, **0.05, ***0.001

Results: Balance

Summary Statistics by Treatment Arm

	Control	Gain	Loss	<i>p-value</i>
Female	0.501	0.489	0.484	0.791
White	0.043	0.048	0.044	0.957
Black	0.357	0.359	0.365	0.978
Hispanic	0.598	0.573	0.576	0.953
Free or Reduced Price Lunch	0.988	0.981	0.972	0.172
Limited English Proficiency	0.168	0.151	0.121	0.866
Individualized Education Plan	0.121	0.096	0.121	0.440
Baseline Thinklink Math Score	-0.069	0.059	0.037	0.451
Baseline Thinklink Reading Score	-0.040	0.040	0.021	0.722
Observations	656	981	982	
Homerooms	38	48	59	

The last column reports the p-value from a test of equal means in the three groups, with standard errors clustered at the homeroom level. Asterisks indicate significance at *0.1, **0.05, ***0.001

Results: Treatment Effects on Incentivized Test Scores

	Percentile Rank		Scaled Score	
	(1)	(2)	(3)	(4)
Pooled Loss	6.866** (2.677)	6.840*** (2.554)	0.222*** (0.070)	0.220*** (0.065)
Individual Loss	5.827* (3.090)	6.284** (3.002)	0.212** (0.081)	0.220*** (0.078)
Team Loss	8.201*** (3.111)	7.471** (2.889)	0.238*** (0.084)	0.221*** (0.076)
Pooled Gain	1.263 (2.888)	1.884 (2.834)	0.078 (0.072)	0.092 (0.070)
Individual Gain	0.942 (3.299)	1.939 (3.293)	0.066 (0.083)	0.092 (0.081)
Team Gain	1.777 (3.440)	1.846 (3.343)	0.095 (0.085)	0.093 (0.082)
Controls	No	Yes	No	Yes
Pr(Gain = Loss)	0.019	0.031	0.019	0.027
Observations	2311	2311	2311	2311
Homerooms	141	141	141	141

The table reports OLS estimates in standard deviation or percentile units. Standard errors clustered by homeroom are reported in parentheses. Asterisks indicate significance at *0.1, **0.05, ***0.001

Results: Treatment Effects on Unincentivized Test Scores

	Percentile Rank		Scaled Score	
	(1)	(2)	(3)	(4)
Pooled Loss	6.867** (3.269)	6.680** (3.194)	0.213*** (0.078)	0.398*** (0.076)
Individual Loss	2.915 (3.918)	3.240 (3.917)	0.157* (0.090)	0.149 (0.092)
Team Loss	11.686*** (3.472)	10.803** (3.350)	0.292*** (0.092)	0.271*** (0.085)
Pooled Gain	0.228 (3.402)	1.012 (3.351)	-0.013 (0.085)	-0.003 (0.084)
Individual Gain	0.205 (3.818)	0.991 (3.770)	-0.054 (0.098)	-0.036 (0.096)
Team Gain	0.666 (4.283)	1.207 (4.251)	0.042 (0.115)	0.039 (0.113)
Controls	No	Yes	No	Yes
Pr(Gain = Loss)	0.017	0.042	0.003	0.006
Observations	2144	2144	2144	2144
Homerooms	140	140	140	140

The table reports OLS estimates in standard deviation or percentile units. Standard errors clustered by homeroom are reported in parentheses. Asterisks indicate significance at *0.1, **0.05, ***0.001

Results: Spillovers

	ThinkLink		ISAT/ITBS	
	(1)	(2)	(3)	(4)
Pooled Loss	6.840*** (2.554)	9.559*** (3.585)	6.680** (3.194)	10.838** (5.470)
Individual Loss	6.284** (3.002)	9.128** (4.114)	3.240 (3.917)	7.333 (6.311)
Team Loss	7.471** (2.889)	10.011*** (3.690)	10.803*** (3.350)	14.025** (5.564)
Pooled Gain	1.884 (2.834)	3.901 (3.508)	1.012 (3.351)	4.556 (5.198)
Individual Gain	1.939 (3.293)	3.706 (3.889)	0.991 (3.770)	3.883 (5.346)
Team Gain	1.846 (3.343)	4.070 (3.891)	1.207 (4.251)	4.569 (6.007)
Controls	Yes	Yes	Yes	Yes
Spillovers Removed	No	Yes	No	Yes
Pr(Gain = Loss)	0.031	0.015	0.006	0.004
Observations	2311	1986	2144	1828
Homerooms	141	125	140	124

The table reports OLS estimates in standard deviation or percentile units. Standard errors clustered by homeroom

Results: Treatment Effects by Grade

	<i>Whole</i>	<i>Grade Level</i>		p-val
	<i>Sample</i>	K-2	3-8	
Pooled Loss	6.840*** (2.554)	17.745*** (6.126)	5.791** (2.555)	0.065
Individual Loss	6.284** (3.002)	17.437** (6.922)	6.603** (2.871)	0.137
Team Loss	7.471** (2.889)	17.328** (6.931)	5.690* (3.157)	0.118
Pooled Gain	1.884 (2.843)	15.515*** (5.359)	-0.740 (3.266)	0.008
Individual Gain	1.939 (3.293)	15.515*** (5.247)	-0.740 (4.244)	0.004
Team Gain	1.846 (3.343)	14.487** (6.403)	2.544 (3.538)	0.093
Controls	Yes	Yes	Yes	
Observations	2311	638	1673	
Homerooms	141	45	97	

The table reports OLS estimates in standard deviation or percentile units. Standard errors clustered by homeroom are reported in parentheses. Asterisks indicate significance at *0.1, **0.05, ***0.001