Evaluation Report: Investing in Innovation Pathways to Success

Prepared for:
McREL International
4601 DTC Boulevard, Suite 500
Denver, CO 80237-2596

Submitted by:
Mark H. C. Lai
University of Southern California
3620 S McClintock Ave
Los Angeles, 90089

2021

Table of Contents

Overview of the IBM Journey Program

IBM Journey is a development grant funded by the U.S. Department of Education's Investing in Innovation (i3) Fund (Award Number U411C150011). Oyserman and the USC team used this funding to translate the identity-based motivation (IBM) in vivo program, *Pathways to Success (Pathways)*, into a digital platform to deliver IBM to middle and high school students. To create this translation, Oyserman and the USC team took the 12 half-hour sessions of *Pathways* and translated them into 12 15-minute digital platform activities. Just as the in vivo *Pathways*, the digital Pathways program was planned as a twice-weekly intervention to take place over the first six to eight weeks of the academic school year (though as detailed below, this plan was not attained). The goal of the *Pathways* program is to improve academic outcomes by changing the three elements of identity-based motivation. These elements are operationalized in *Pathways* as having school-focused possible identities with strategies to attain them, experiencing one's adult future as close and connected to the present, and productive interpretations of experienced difficulty. The specific non-cognitive factors targeted by the *Pathways* program are feelings of connection to the future, strategies for action, and productive interpretation of experienced difficulty. Improvements in these non-cognitive factors are expected to lead to improvements in students' school engagement (teacher report) and academic outcomes (school report).

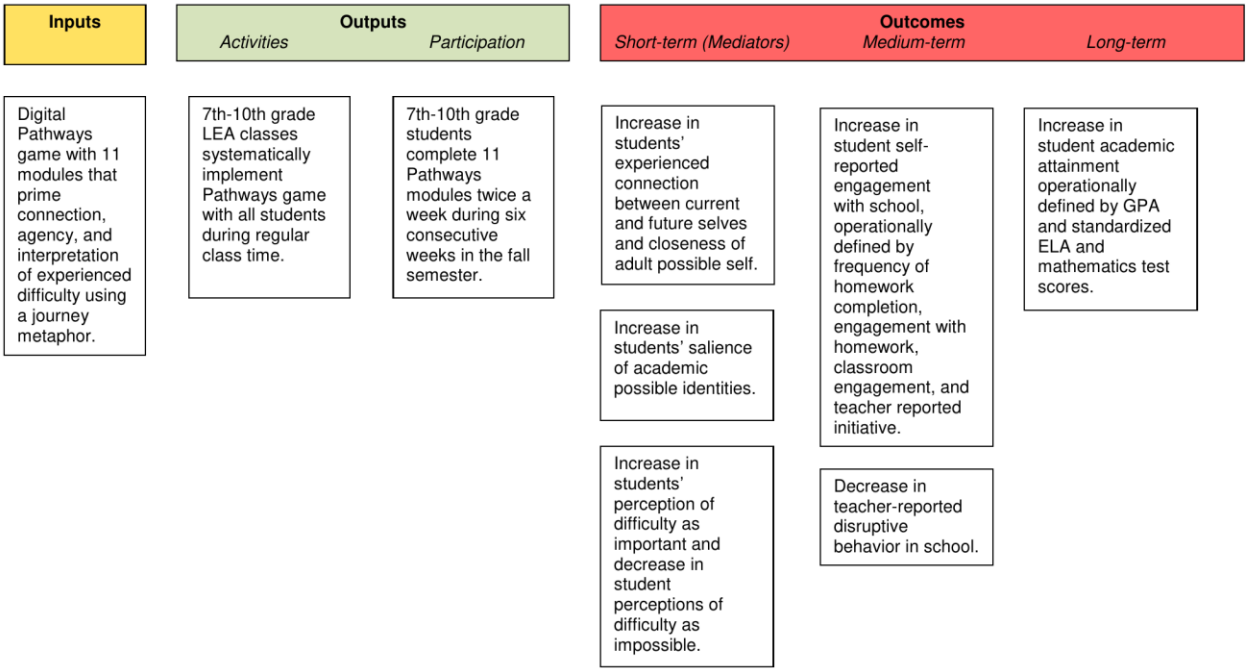
Digital Pathways

The initial months of grant activity (May 2016 until August 2017) focused on the development of digital *Pathways*, Daphna Oyserman, and the USC team working with a software development team led this effort. Oyserman translated the in vivo activities and core take-home points to digital activities. She devised a way to activate social norms and to promote active engagement ("saying is believing"). A logic model is shown in Figure 1.

Planned Implementation

To attain proper dosage, participating schools and teachers were asked to allow students to engage with the digital platform individually in their classrooms twice a week for six consecutive weeks in the fall semester of 2017. Students mostly engaged in the digital learning experience in their science classes, but because principals could choose the class period in which implementation would take place, a small proportion of students engaged in the experience in their language arts classes. Prior to implementation, students were to complete a baseline survey. Students were instructed either to complete engagement in six weeks (twice each week) or in seven weeks (one session in week 1, two sessions each in weeks 2, 3, 4, 5, and 6, and one session in week 7). Teachers who engaged with the digital platform for either six or seven weeks were considered to be in compliance with our expectations for implementation. Teachers were asked to allow their students to make up any missed sessions as soon as possible.

Figure 1: Logic Model of Identity-Based Motivational Journey to Academic Success



Dosage was operationalized at the classroom level by teachers who a) provided class time for 12 sessions in one semester and b) implemented games over 6 consecutive weeks, 2 times per week, to complete all 12 modules. The registered fidelity matrix can be found in Table 1. Table 2 provides a summary of the actual dosage scored using the What Works Clearinghouse coding scheme. As can be seen in Table 2, with regard to classroom-level dosage, only 2.7% of classrooms met both dosage requirements: 63.5% of classrooms provided class time for 12 sessions during the fall semester. Based on the login data only 2.7% of teachers followed the implementation schedule, implying that the treatment was not delivered as intended.

The a priori operationalization of dosage at the student level focused on session completion. Full dosage would include a) completion of all 12 sessions in the fall semester whether in the treatment or control group and b) in the treatment condition, that students shared or clicked on at least one trailblazer during each of the 12 sessions. In addition, because for the in vivo Pathways, anyone receiving at least half of Pathways topic sessions was considered to have received a sufficient treatment dosage, we also looked at theoretically sufficient dosage fidelity as students having a) completed the first seven sessions in the fall semester and b) shared or clicked on at least one trailblazer during each of the seven sessions. However, teachers did not necessarily know if students had completed a session, so we also calculated the percentages by the number of sessions students attempted. Table 3 provides a summary of session completion for the treatment group, 31.7% of students completed all 12 sessions and 52.3% completed at least the first 7 sessions.

Table 1. Registered Fidelity Matrix for Implementation

Identity-Based Motivational Journey										
Indicators	**Definition**	**Unit of implem-entation**	**Data Source(s)**	**Data Collection (who, when)**	**Score for levels of implementation at unit level**	**Threshold for adequate implementation at unit level**	**Roll-up to program level**	**Threshold for adequate implementation at sample level)**	**Expected sample for fidelity measure**	**Expected years of fidelity measurement**
Key Component 1:School Implementation of IBM Journey Game										
School delivers Pathways game to students during regular class time	Teacher provides class time for students to complete all 12 modules in fall semester	Class	Digital platform data/computer logs (specifically, date, time, and # modules completed)	Extracted from digital platform by evaluator	0= Teacher provides time for students to complete fewer than 12 modules in one semester 1 = Teacher provides classroom time for students to complete all 12 modules in one semester	Adequate implementation at teacher level = score of "1"	0 = less than 50% of teachers receive a score of "1" 1 = 51-85% of teachers receive a score of "1" 2 = 86-100% of teachers receive a score of "1"		All teachers assigned to implement Pathways game	2017-2018
Pathways game implemented in twice-weekly sessions in 6 consecutive weeks	Teacher implements games over 6 consecutive weeks, 2 times per week, to complete all 12 modules	Class	Digital platform data/ computer logs (specifically, date collected from student login)	Extracted from digital platform by evaluator	0= Teacher does not provide class time for students to complete modules in 6 consecutive weeks 1 = Teacher provides class time for students to complete modules in 6 consecutive weeks	Adequate implementation at teacher level = score of "1"	0 = less than 50% of teachers receive a score of "1" 1 = 51-85% of teachers receive a score of "1" 2 = 86-100% of teachers receive a score of "1"		All teachers assigned to implement Pathways game	2017-2018
Both indicators		class	Digital platform data (specifically, date, time, and	Extracted from digital platform by evaluator			Range for component score = 0 - 4	Fidelity of implementation at program level = at least 86% of	All teachers assigned to implement Pathways game	2017-2018

Identity-Based Motivational Journey										
			# modules completed)					teachers receive score of 4 (score of 2 for full implementation of game [86-100% of teachers implement all 12 modules] + score of 2 on timing of modules [86-100% of teachers follow 6-week schedule])		
Key Component 2: Student Participation in IBM Journey Game										
Students complete the Pathways game	Student completes all 12 modules in fall semester	Student	Computer logs	Extracted from digital platform by evaluator	0 = student completes fewer than 12 modules 1 = student completes all 12 modules	Adequate implementation at student level = score of "1"	0 = less than 50% of students receive a score of "1" 1 = 51-85% of students receive a score of "1" 2 = 86-100% of students receive a score of "1"		All students assigned to treatment group	2017-2018
Students click on trailblazer for each session/ module	Student writes or clicks on a trailblazer during a session	Student	Computer logs	Extracted from digital platform by evaluator	0 = student never writes or clicks on a trailblazer across completed module 1 = student writes or clicks on a trailblazer at least once for	Adequate quality of delivery at student level = score of "1"	0 = less than 50% of students receive a score of "1" 1 = 51-85% of students receive a score of "1"		All students assigned to treatment group	2017-2018

Identity-Based Motivational Journey										
					each session completed		2 = 86-100% of students receive a score of "1"			
Both indicators		Student	Computer logs	Extracted from digital platform by evaluator			Range for component score = 0 - 4	Fidelity of implementation at program level = at least 86% of students receive score of 4 (score of 2 on completion of all 12 modules in 6-week period [86-100% of students] and score of 2 on use of at least 1 trailblazer on each module [86-100% of students])	All students assigned to treatment group	2017-2018

Table 2. Fidelity of Implementation of Digital Pathways (2017-2018, Component 1: School Implementation, *N* = 74 classrooms)

Fidelity Indicator	**Number and Percentage of Total Classrooms Meeting Each Threshold**		**Score**
	n	**%**	
School staff delivered the Pathways game to students during regular class time	47	63.5%	1
The Pathways game was implemented in twice-weekly sessions in 6 consecutive weeks	2	2.7%	0
Both indicators were met	2	2.7%	1

Note. Fidelity Scores follow the Department of Education, *What Works Clearinghouse.™ Procedures and Standards Handbook. Version 3.0* NEi3 review protocol. For single indicators, 0 = < 50%; 1=51% to 85%; 2= 86% to 100% of students met the indicator criterion. For the combined (both indicators) score, 0 = < 50% on both indicators; 1= 51% to 85% on one of the two indicators; 2=51% to 85% on both indicators; 3= 51% to 85% on one indicator and 86-100% on the other indicator; 4=86-100% on both indicators.

Table 3. Implementation Fidelity of Digital Pathways (2017-2018, Component 2: Student Participation, *N* = 859 students)

Fidelity Indicator	**Number and Percentage of Total Students Meeting Thresholds**		**Score**
	n	**%**	
Students completed the Pathways game	272	31.7%	0
Students clicked on trailblazer for each session/ module	272	31.7%	0
Both indicators were met	272	31.7%	0

Note. Fidelity Scores follow the Department of Education, *What Works Clearinghouse.™ Procedures and Standards Handbook. Version 3.0* NEi3 review protocol, for single indicators, 0 = less than 50% of students; 1=51-85% of students; 2= 86=100% of students; for the combined indicator, 0 = less than 50% on both indicators; 1= 51-85% on one of the two indicators; 2=51-85% on both indicators; 3= 51-85% on one indicator and 86-100% on two indicators; 4=86-100% on both indicators. 12-session (full Pathways Program), 7-session (a priori indicator of sufficient content)

Control Condition

The software development partner of this project shared the use of *Motion Force*---a science-learning edu-game, for students in the control condition who engaged Pathways in their science class, and the *Sleep Furiously* edu-game for students in the control condition who engaged

Pathways in their language arts class. The delivery of the control condition was planned to be identical to that of the *Pathways* condition, however the delivery differed because of technical issues (e.g., the treatment software saved individual student responses, and the control software did not save individual student responses). This meant that control students could not continuously engage in the experience because they had to start each new session at the point where the prior computer user left off.

Participants and Settings

During the fall semester of 2017, the digital *Pathways* program was provided to middle and high school students in 10 participating schools from five school districts in Colorado. The school-level eligibility criterion was a school leader's willingness to allow students to engage with the digital learning experience programs twice a week for six consecutive weeks. The teacher-level eligibility criterion was a teacher's willingness to complete brief baseline and year-end student engagement surveys. At the student level, all 7th-to-10th grade students in the classrooms of teachers who agreed to participate were randomly assigned to treatment (digital Pathways) or active control (Motion Force or Sleep Furiously). No exclusion criteria regarding special education, individualized educational plan, or English language learner status were set.

Table 4 shows the free- and reduced-price lunch information both at the school- and student-level. At the student level, the initial randomization sample consisted of 1,708 students (n = 859 treatment, n = 849 control) in 74 classrooms. After excluding students whose parent/guardians opted them out, students who were not allowed to use computers, and students who withdrew from their schools prior to the start of the intervention, there were 228 students from 7th grade, 295 from 8th grade, 520 from 9th grade, 383 from 10th grade, in the Fall of 2017; in addition, 2 students were from 11th or 12th grade and were registered in the classes randomized to either Pathways or the control condition. In terms of demographics, 42.1% of the students were eligible for free- or reduced-price lunch; 53% were male; 58.6% were non-Latino white, 29.1% were Latino or Hispanic heritage (any race), 8.9% were African American or Native American, and 3.3% identified themselves as other categories of ethnicity. The median age was 14.

Table 4: Percentage of Students Receiving Free or Reduced Lunch (School-level and in our sample), and School-level Special Education, and English-language Learner Services

School	**% Free/Reduced Price Lunch**		**% SPED**	**%ELL**
	School-Level	**In the sample**		
Manzanola Jr Sr High	68.5%	No admin data provided	N/A	N/A
Sterling Middle School	55.2%	50.8%	14%	7%
Liberty Point IB	50.1%	61.2%	14%	6%
Caliche Jr-Sr High	46.0%	52.0%	15%	N/A

Delta High School	44.3%	55.6%	11%	5%
Craver Middle School	41.3%	41.2%	14%	N/A
Sterling High School	35.6%	43.5%	13%	4%
Rye High School	34.7%	37.2%	16%	N/A
Pueblo West High School	29.2%	36.2%	11%	2%
Skyview Middle School	23.6%	31.0%	12%	4%

Note. We had access to school-level but not sample-level data on special education and English Language Learning. Jr-Sr = Junior and Senior, SpED=Special Education, ELL=English Language Learners, N/A=Not available because data were suppressed by the Colorado Department of Education to protect student privacy.

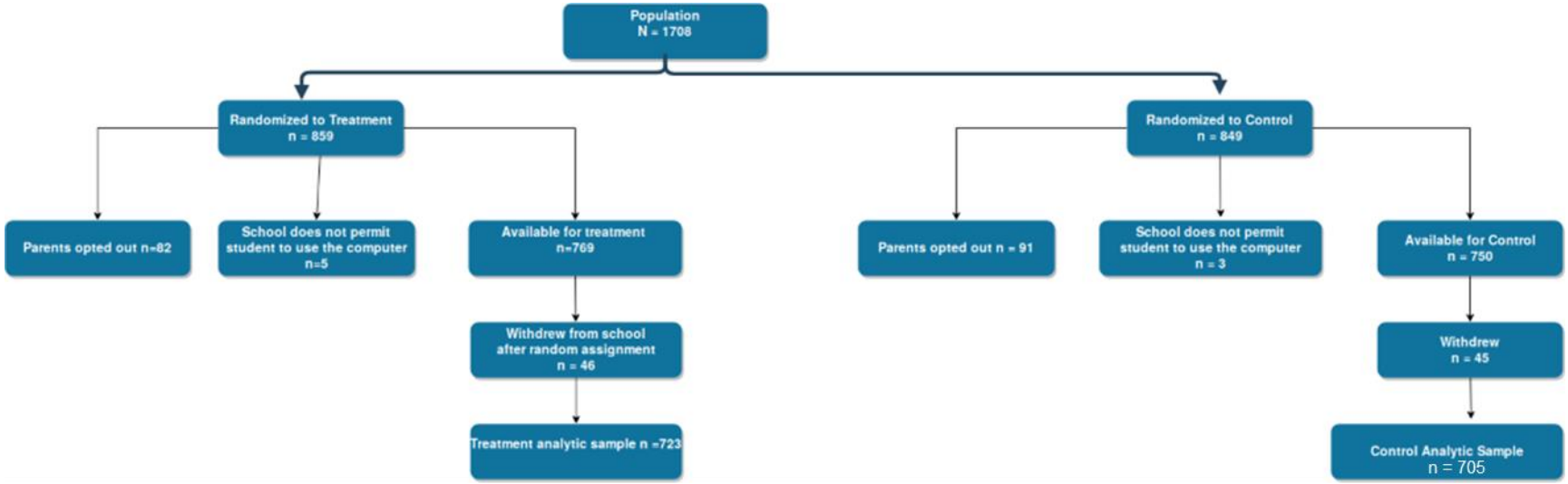
Study Design

Random Assignment

The original evaluator (Dr. Kristin Bub) conducted the randomization based on a list of student identification numbers for each participating classroom provided by the schools. The original evaluator had no other information about the students at the time of the random assignment. Hence, no other factors were involved in the random assignment. She randomized half of the full roster of students in each participating classroom to treatment and half to control using the Stata randomize software. The original evaluator repeated this procedure for each of the 38 participating teachers and 74 classrooms. Therefore, students were randomly assigned to *Pathways* or control conditions, with the classroom as a blocking variable.

Post randomization, a total of 280 students were excluded. Figure 2 shows a breakdown.

Figure 2: From Population (Students Randomly Assigned) to Analytic Sample



Measures

The following measures were collected for academic years 2016-2017 (pre-intervention), 2017-2018 (6 months post-intervention), and 2018-2019 (18 months post-intervention).

Overall Achievement

Cumulative grade point average (GPA)

GPA data for each student was collected from school records. GPA belongs to the general achievement domain.

English Language Achievement (ELA) and Math Achievement

In 2016-2017, researchers collected data from the English language arts and literacy subtests and the mathematics subtests from the statewide PARCC assessments in Colorado for students up to the 10th grade. In 2017-2018, the state of Colorado adopted the Colorado Measures of Academic Success (CMAS) assessments, which were comparable to the PARCC, and the ELA and math subtests were similarly obtained for students who were in their 7th and 8th grades in 2017-2018. However, CMAS was only administered for students in grades three through eight, so data from the Preliminary Scholastic Assessment Test (PSAT) test were obtained for 9th grade students. Similarly, in 2018-2019, CMAS data were obtained for students who were in 7th grade in 2017-2018, and PSAT data were obtained for other students. The standardized test data (ELA and math) were obtained from the school districts. These two assessments are from the ELA achievement and the Mathematics achievement domains.

Student-Reported School Engagement

The student-reported school engagement measure from the student reported school engagement domain, was adapted from Finn, Pannozzo, and Voelkl (1995) and from personal communication with J. Finn (October 14, 1998). There are 14 items on a 5-point scale (1 = *never* to 5 = *always*). Seven of them were reverse-coded. An example item is “How often do you pay attention in class?” The composite score was computed using the average of 14 items. Cronbach’s α reliability ranged from .75 to .79.

Interpretation of Experienced Difficulties

Difficulty-as-impossibility

This 6-item 5-point response scale student-survey collective measure comes from Oyserman et al. (2015). This scale belongs to the interpretation of difficulty as impossible domain. An example item is “When I feel stuck on a school task, it’s a sign that my effort is better spent elsewhere.” Responses from all 6 items were averaged into a composite score. Previous studies with similar measures have established internal consistency and convergent validity (Oyserman et al., 2015).

Difficulty-as-importance

This 6-item 5-point response scale student-survey collective measure comes from Oyserman et al. (2015). This scale belongs to the interpretation of difficulty as important domain. An example item is "When I feel stuck on a school task, it's a sign that my effort is better spent elsewhere." Responses from all 6 items were averaged into a composite score. Previous studies with similar measures have established internal consistency and convergent validity (Oyserman et al., 2015).

Academic Possible Selves

Experienced Closeness and Connection to Future Self

Two single-item measures (a graphic display of circles labeled current me and future me) revised from Nurra & Oyserman (2018) were used to assess student experience of their future self as being close to and connected with their current self. Responses from these two items were averaged into a composite score.

Table 5 shows the Cronbach α reliability coefficients for student-reported engagement and student-reported difficulty-as-impossibility and difficulty-as-importance for each time point for students randomly assigned to the intervention and control conditions. For the single item connection and closeness to adult future-self measures, inter-item reliability cannot be calculated; instead, Spearman rho statistics (test-retest reliability) are provided in Table 6. Positive correlation implies that students who experienced more connection to their future selves at one time point also experience more connection to that self at other time points.

Table 5. Student-Reported Intermediate Outcomes: Cronbach Alpha Reliability by Condition

		Year 16-17		**Year 17-18**		**Year 18-19**	
Scale	**Condition**	***α***	**n[a]**	***α***	**n[a]**	***α***	**n[a]**
Student-reported School Engagement (1-5)	Treatment	.77	640	.77	599	.75	306
	Control	.78	634	.79	577	.77	298
Difficulty-as-Impossibility (1-5)	Treatment	.81	639	.84	598	.85	311
	Control	.85	632	.87	579	.88	299
Difficulty-as-Importance (1-5)	Treatment	.86	640	.89	597	.88	311
	Control	.88	639	.91	580	.89	300

[a]n=Sample size reported as the smallest sample size of all items.

Table 6. Student-Reported Intermediate Outcomes: Test-Retest Reliability by Condition

		Year 16-17		**Year 17-18**		**Year 18-19**	
Scale		**Reliability**	**n[a]**	**Reliability**	**n[a]**	**Reliability**	**n[a]**
Connection to Adult	Year 16-17	—	—	0.18	540	0.27	284
Possible Self (1-4)	Year 17-18	0.28	547	—	—	0.26	284

	Year 18-19	0.22	286	0.24	295	—	—

Note. Reliability for the two single-item measures cannot be computed, instead, test-retest reliability (Spearman's rho statistics) across time are provided. Reliability coefficients for the treatment group are reported in the lower diagonal, and the control group's reliability coefficients are reported in the upper diagonal.

Teacher-Reported Student Engagement

Initiative-taking

This 5-item 5-point response scale teacher-survey collected measure come from Finn, Pannozzo, and Voelkl (1995). This scale is in the domain of teacher reported student engagement. An example item is "How often does this student do more than the work assigned?" The composite score was computed using the average of 5 items. The validity of the measure was established in Oyserman et al. (2006).

Disruptive Behaviors

This 4-item 5-point response scale teacher-survey collected measure come from Finn, Pannozzo, and Voelkl (1995). This scale is in the domain of teacher reported student engagement. An example item is "How often does this student annoys peers or interferes with peers' work?" The composite score was computed using the average of 5 items. Validity was established in Oyserman et al. (2006).

Table 7 shows the Cronbach α reliability coefficients for teacher-reported initiative-taking and disruptive behaviors.

Table 7. Teacher-Reported Intermediate Outcomes: Cronbach Alpha Reliability by Condition

		Year 16-17		**Year 17-18**		**Year 18-19**	
scale	**Condition**	***α***	**n[a]**	***α***	**n[a]**	***α***	**n[a]**
Student Initiative-taking (0-4)	Treatment	.78	595	.82	598	.90	182
	Control	.78	608	.81	574	.84	192
Student Disruptive Behavior (0-4)	Treatment	.77	597	.75	590	.79	186
	Control	.76	598	.72	574	.78	195

[a]n=Sample size reported as the smallest sample size of all items.

Analytic Approach

Baseline Equivalence

We conducted a series of baseline equivalence t-tests to investigate the effectiveness of random assignment. Following the What Works Clearinghouse Standards Handbook, we calculated standardized effect size (Hedges's *g*) for differences, excluding students who opted out. As shown in the results, as the treatment and control groups showed a difference of > 0.05 SDs but < 0.25 SD on some of the baseline measures, consequently, in the subsequent impact analyses we adjusted for the baseline measures.

Statistical Model

Given that students are nested within classrooms, teachers, and schools, a 4-level multilevel model was conducted to examine the impact of random assignment to the treatment condition on each outcome during the 2017-2018 academic year and the 2018-2019 academic year, respectively, with the binary treatment assignment indicator (0 = Control, 1 = Treatment) at the student-level, to conduct an intent-to-treat analysis. For all outcomes, the corresponding baseline measures were included as covariates; specifically, for PSAT ELA (or Math), we used the baseline measures of PARCC ELA (or Math) as covariates. Random treatment effects were included at classroom, teacher, and school levels to avoid bias induced due to the omission of random effects. Restricted maximum likelihood estimation was used, with *p* values obtained using the *F* test with the Satterthwaite method for denominator degrees of freedom.

The statistical model can be described, in equation form as

Level-1 (student-level):

$$Y_{ijkl} = \beta^{c}_{0jkl} + \beta^{c}_{1jkl}\text{TREATMENT}_{ijkl} + \beta^{c}_{2jkl}\text{BASELINE}_{ijkl} + e_{ijkl}$$

Level-2 (classroom-level):

$$\begin{aligned}\beta^{c}_{0jkl} &= \beta^{t}_{00kl} + u^{c}_{0jkl}\\ \beta^{c}_{1jkl} &= \beta^{t}_{10kl} + u^{c}_{1jkl}\\ \beta^{c}_{2jkl} &= \beta^{t}_{20kl}\end{aligned}$$

Level-3 (teacher-level):

$$\begin{aligned}\beta^{t}_{00kl} &= \beta^{s}_{000l} + u^{t}_{00kl}\\ \beta^{t}_{10kl} &= \beta^{s}_{100l} + u^{t}_{10kl}\\ \beta^{t}_{20kl} &= \beta^{s}_{200l}\end{aligned}$$

Level-4 (school-level):

$$\beta^s_{000l} = \gamma_{0000} + u^s_{000l}$$
$$\beta^s_{100l} = \gamma_{1000} + u^s_{100l}$$
$$\beta^s_{200l} = \gamma_{2000}$$

We used complete case analysis because attrition did not differ for the outcome variables.

Results

Impact

Table 8 shows the estimated impact during the 2017-2018 academic year (6-month follow-up). No significant difference between students randomized to the Treatment and the Control conditions were found at .01 significance-level for the intermediate outcomes: student-reported classroom behavior; student-reported difficulty mindsets; student-reported distance and connection between current and adult selves; and teacher-reported initiative-taking and disruptive behaviors. No significant difference between students randomized to the Treatment and the Control conditions were found at .01 significance level for the long-term outcomes: cumulative grade point average and state-administered ELA and mathematics scores (PARCC for students at 10th grade or below; PSAT for students at 11th grade or above). The effect sizes ranged between -0.06 to 0.12.

Table 8: Impact Analysis for Outcomes Measures During the 17-18 Academic Year

	Control				**Treatment**				**Estimated Impact**			
Outcome	**N2**	**N1**	**M**	**SD**	**N2**	**N1**	**M**	**SD**	**M Diff T - C**	**SE**	**p**	**ES**
GPA	71	636	2.93	0.87	69	651	2.91	0.85	0.00	0.03	0.90	0.00
ELA (PARCC)	25	196	747.30	30.00	25	206	743.23	31.30	-3.80	2.89	0.25	-0.12
Math (PARCC)	25	197	739.58	30.95	25	206	737.08	26.49	-0.79	2.67	0.78	-0.03
ELA (PSAT)	46	321	451.31	75.39	45	315	449.24	80.82	2.35	4.88	0.64	0.03
Math (PSAT)	46	321	443.30	73.61	45	315	438.89	72.92	1.60	3.98	0.69	0.02
Student-reported School Engagement	68	579	3.60	0.54	68	600	3.59	0.52	-0.02	0.03	0.43	-0.04
Difficulty-as-Impossibility	68	581	2.47	0.80	68	601	2.50	0.78	0.06	0.07	0.42	0.08
Difficulty-as-Importance	68	581	3.10	0.82	68	601	3.16	0.80	0.04	0.05	0.45	0.05
Connection to Adult Possible Self	68	582	2.43	0.68	68	601	2.41	0.70	-0.05	0.04	0.26	-0.07
Student Initiative-taking (Teacher Report)	65	577	2.67	0.87	65	598	2.65	0.87	-0.05	0.04	0.31	-0.05
Student Disruptive Behavior (Teacher Report)	65	574	1.35	0.53	65	590	1.43	0.61	0.06	0.03	0.07	0.11

Note. N2 = number of classrooms, N1 = number of students. Baseline measures were adjusted in the analysis.

Table 9 shows the estimated impact during the 2018-2019 academic year (18-month follow-up). No significant difference between students randomized to the Treatment and Control conditions was found at .01 significance level for the intermediate outcomes: student-reported classroom

behavior; student-reported difficulty mindsets; student-reported distance and connection between current and adult selves; and teacher-reported student initiative-taking and disruptive classroom behavior. No significant difference between students randomly assigned to the Treatment and the Control conditions were found at .01 significance level for the long-term outcomes: cumulative grade point average and state-administered ELA and mathematics scores (PARCC for students at 10th grade or below; PSAT for students at 11th grade or above). The effect sizes ranged between -0.14 to 0.08.

Table 9: Impact Analysis for Outcomes Measures During the 18-19 Academic Year

	Control				**Treatment**				**Estimated Impact**			
Outcome	**N2**	**N1**	**M**	**SD**	**N2**	**N1**	**M**	**SD**	**M Diff T - C**	**SE**	**p**	**ES**
GPA	71	610	2.86	0.92	71	632	2.81	0.94	-0.02	0.05	0.68	-0.02
ELA (PARCC)	11	85	742.20	37.63	11	102	743.39	32.44	1.34	9.21	0.90	0.04
Math (PARCC)	11	85	728.56	32.06	11	102	730.01	30.85	-5.21	3.95	0.27	-0.17
ELA (PSAT)	59	470	477.06	85.38	59	465	469.23	86.89	-4.91	3.66	0.18	-0.06
Math (PSAT)	59	470	468.83	82.37	59	465	459.55	82.05	-2.31	3.78	0.54	-0.03
Student-reported School Engagement	51	299	3.59	0.51	54	307	3.58	0.48	-0.02	0.04	0.68	-0.04
Difficulty-as-Impossibility	51	301	2.62	0.73	54	313	2.57	0.72	-0.08	0.06	0.24	-0.11
Difficulty-as-Importance	51	301	3.15	0.69	54	312	3.11	0.72	-0.04	0.08	0.65	-0.06
Connection to Adult Possible Self	51	300	2.52	0.70	54	311	2.47	0.64	-0.07	0.06	0.28	-0.10
Student Initiative-taking (Teacher Report)	45	196	3.10	0.92	41	184	3.04	0.97	-0.10	0.16	0.59	-0.10
Student Disruptive Behavior (Teacher Report)	45	195	1.46	0.62	42	186	1.49	0.63	0.08	0.08	0.33	0.12

Note. N2 = number of classrooms, N1 = number of students. Baseline measures were adjusted in the analysis.

Attrition

A total of 1,708 students (859 in the Treatment condition, 849 in the Control condition) were randomly assigned to treatment and control groups. For each measure, attrition rates were computed as (1 - number of valid responses) / number of students randomized to condition. After excluding students whose parent/guardians opted them out, students who were not allowed to use computers, and students who withdrew from their schools prior to the start of the study, the analytic sample included 1,428 students (723 in Treatment, 705 in control), so the attrition rate was 16.4% (15.8% in Pathways, 17.0% in control). Tables 10 to 12 show the attrition rates of each measure for the 2016-2017, 2017-2018, and 2018-2019 academic years. Overall, the attrition rates were similar across the treatment and the control conditions, with attrition rates between 23.4% to 30.7% in 2016-2017, between 24.2% to 39.1% in 2017-2018, and 26.43% to 78.81% in 2018-2019.

Table 10: Attrition rates in 16-17

Outcome	Control (*n* = 849)		Treatment (*n* = 859)	
	N	Attrited %	N	Attrited %
Student-reported School Engagement (1-5)	643	24.26	648	24.56
Difficulty-as-Impossibility (1-5)	641	24.50	645	24.91
Difficulty-as-Importance (1-5)	643	24.26	645	24.91
Connection to Adult Possible Self (1-4)	650	23.44	651	24.21
Student Initiative-taking (Teacher Report) (0-4)	608	28.39	595	30.73
Student Disruptive Behavior (Teacher Report) (0-4)	598	29.56	597	30.50
GPA (0-4)	603	28.98	617	28.17
ELA Standardized Scores	590	30.51	605	29.57
Math Standardized Scores	588	30.74	603	29.80

Table 11: Attrition rates in 17-18

Outcome	Control (*n* = 849)		Treatment (*n* = 859)	
	N	Attrited %	N	Attrited %
Student-reported School Engagement (1-5)	579	31.80	600	30.15
Difficulty-as-Impossibility (1-5)	580	31.68	599	30.27
Difficulty-as-Importance (1-5)	581	31.57	598	30.38
Connection to Adult Possible Self (1-4)	582	31.45	601	30.03
Student Initiative-taking (Teacher Report) (0-4)	574	32.39	598	30.38
Student Disruptive Behavior (Teacher Report) (0-4)	574	32.39	590	31.32
GPA (0-4)	636	25.09	651	24.21
ELA Standardized Scores	517	39.10	521	39.35
Math Standardized Scores	518	38.99	521	39.35

Table 12: Attrition rates in 18-19

	Control (*n* = 849)	Treatment (*n* = 859)

Outcome	N	Attrited %	N	Attrited %
Student-reported School Engagement (1-5)	299	64.78	307	64.26
Difficulty-as-Impossibility (1-5)	301	64.55	313	63.56
Difficulty-as-Importance (1-5)	301	64.55	312	63.68
Connection to Adult Possible Self (1-4)	300	64.66	311	63.80
Number of School-focused Possible Selves with at least one Strategy (0-8)	283	66.67	315	63.33
Student Initiative-taking (Teacher Report) (0-4)	192	77.39	182	78.81
Student Disruptive Behavior (Teacher Report) (0-4)	195	77.03	186	78.35
GPA (0-4)	610	28.15	632	26.43
ELA Standardized Scores	555	34.63	567	33.99
Math Standardized Scores	555	34.63	567	33.99

Baseline Equivalence

Intermediate outcomes

We found no statistically significant differences between treatment and control on any of our intermediate outcomes: Student-reported classroom behavior; student-reported difficulty mindsets; student-reported distance and connection between current and adult selves; and teacher-reported student initiative-taking and disruptive classroom behavior. Table 13a presents results for all grades (7th to 12th), Table 14a presents results for middle school, and Table 15a presents results for high school. Treatment group and control group youth were equally engaged with homework and in the classroom, felt similarly distant from and connected to their adult selves, and scored similarly on difficulty mindsets.

Long-term outcomes

We found no statistically significant differences between treatment and control on any of our long-term academic outcomes: cumulative grade point average, state-administered ELA and mathematics scores provided by schools and districts for the 2016-2017 academic year. Results are summarized in Table 13b (full sample) Table 14b (middle school) and Table 15b (high school).

Table 13a: Baseline Equivalence by Treatment Status (7th -12th Graders Combined) Intermediate Outcomes

	Control			Treatment			Mean Difference (C-T)			
Outcome (Measure Min-Max Range)	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***t*-test**	***p***	**Hedges *g***

Student-reported School Engagement (1-5)	3.74	0.51	643	3.75	0.48	648	-0.01	-0.41	0.684	-0.02
Difficulty-as-Impossibility (1-5)	2.49	0.83	641	2.43	0.76	645	0.06	1.30	0.195	0.07
Difficulty-as-Importance (1-5)	3.32	0.82	643	3.31	0.77	645	0.02	0.36	0.716	0.02
Connection to Adult Possible Self (1-4)	2.42	0.70	650	2.47	0.70	651	-0.05	-1.05	0.294	0.06
Student Initiative-taking (Teacher Report; 0-4)	2.71	0.80	611	2.75	0.82	611	-0.04	-0.77	0.442	0.04
Student Disruptive Behavior (Teacher Report; 0-4)	1.33	0.58	598	1.37	0.60	597	0.04	-1.11	0.268	0.06

C=Control, T=Treatment, C-T=Mean Difference Treatment from Control

Table 13b: Baseline Equivalence by Treatment Status (7th-12th Graders Combined) on Long-Term Outcomes

	Control			**Treatment**			**Mean Difference (C-T)**			
Outcome (Measure Min-Max Range)	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***t*-test**	***p***	**Hedges *g***
ELA Standardized Scores (650-836)	743.24	30.34	590	743.38	29.18	605	-0.14	-0.08	.933	0.00
Math Standardized Scores (650-808)	731.89	25.70	588	730.24	26.14	603	1.65	1.10	.272	0.06
Cumulative GPA (0-4)	2.95	0.83	603	2.90	0.88	617	0.05	1.01	.314	0.06

C=Control, T=Treatment, C-T=Mean Difference Treatment from Control

Table 14a: Baseline Equivalence by Treatment Status for Middle Schoolers (7th -8th Graders) Intermediate Outcomes

	Control			**Treatment**			**Mean Difference (C-T)**			
Outcome (Measure Min-Max Range)	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	**t-test**	***p***	**Hedges *g***
Student-reported School Engagement (1-5)	3.77	0.51	235	3.78	0.50	249	-0.01	-0.15	0.883	-0.01
Difficulty-as-Impossibility (1-5)	2.47	0.86	228	2.37	0.84	244	0.10	1.27	0.205	0.12
Difficulty-as-Importance (1-5)	3.38	0.80	230	3.33	0.84	244	0.05	0.68	0.496	0.06
Connection to Adult Possible Self (1-4)	2.34	0.73	237	2.40	0.74	250	-0.06	-0.91	0.366	0.08
Student Initiative-taking (Teacher Report; 0-4)	2.87	0.75	236	2.91	0.74	249	-0.04	-0.60	0.551	0.05
Student Disruptive Behavior (Teacher Report; 0-4)	1.49	0.64	229	1.48	0.59	240	0.01	0.19	0.844	-0.02

C=Control, T=Treatment, C-T=Mean Difference Treatment from Control

Table 14b: Baseline Equivalence by Treatment Status for Middle School Students (7 and 8th Graders) on Long-Term Outcomes

Outcome (Measure Min-Max Range)	Control			Treatment			Mean Difference (C-T)			
	M	*SD*	*n*	*M*	*SD*	*n*	*M*	t-test	*p*	Hedges *g*
ELA Standardized Scores (650-827)	743.92	29.70	213	746.36	28.31	229	-2.45	-0.89	.376	-0.08
Math Standardized Scores (650-808)	732.85	25.25	213	733.33	23.05	229	-0.47	-0.21	.838	-0.02
Cumulative GPA (1-4)	3.29	0.63	171	3.28	0.65	184	0.02	0.24	.814	0.02

C=Control, T=Treatment, C-T=Mean Difference Treatment from Control

Table 15a: Baseline Equivalence by Treatment Status for High School (9th-10th Grade) Intermediate Outcomes

Outcome (Measure Min-Max Range)	Control			Treatment			Mean Difference (C-T)			
	M	*SD*	*n*	*M*	*SD*	*n*	*M*	t-test	*p*	Hedges *g*
Student-reported School Engagement (1-5)	3.72	0.50	408	3.73	0.47	399	-0.01	-0.36	0.719	-0.03
Difficulty-as-Impossibility (1-5)	2.50	0.81	413	2.47	0.71	401	0.03	0.58	0.562	0.04
Difficulty-as-Importance (1-5)	3.29	0.83	413	3.29	0.73	401	-0.00	-0.04	0.967	-0.00
Connection to Adult Possible Self (1-4)	2.47	0.68	413	2.51	0.67	401	-0.04	-0.69	0.488	0.05
Student Initiative-taking (Teacher Report; 0-4)	2.62	0.81	375	2.64	0.86	362	-0.02	-0.37	0.707	0.03
Student Disruptive Behavior (Teacher Report; 0-4)	1.23	0.51	369	1.29	0.60	357	-0.06	-1.51	0.131	0.11

C=Control, T=Treatment, C-T=Mean Difference Treatment from Control

Table 15b: Baseline Equivalence by Treatment Status for High School Students (9th to 12th Graders) Long-Term Outcomes

Outcome (Measure Min-Max Range)	Control			Treatment			Mean Difference (C-T)			
	M	*SD*	*n*	*M*	*SD*	*n*	*M*	t-test	*p*	Hedges *g*
ELA Standardized Scores (650-836)	742.85	30.74	377	741.56	29.58	376	1.29	0.59	.558	0.04
Math Standardized Scores (650-808)	731.34	25.97	375	728.34	27.72	374	3.00	1.53	.127	0.11
Cumulative GPA (0-4)	2.82	0.87	432	2.74	0.92	433	0.07	1.21	.225	0.08

C=Control, T=Treatment, C-T=Mean Difference Treatment from Control

References

Finn, J. D., Pannozzo, G. M., & Voelkl, K. E. (1995). Disruptive and Inattentive-Withdrawn Behavior and Achievement among Fourth Graders. *The Elementary School Journal*, *95*(5), 421–434. https://doi.org/10.1086/461853

Nurra, C., & Oyserman, D. (2018). From future self to current action: An identity-based motivation perspective. *Self and Identity*, *17*(3), 343–364. https://doi.org/10.1080/15298868.2017.1375003

Oyserman, D., Bybee, D., & Terry, K. (2006). Possible selves and academic outcomes: How and when possible selves impel action. *Journal of Personality and Social Psychology*, *91*(1), 188–204. https://doi.org/10.1037/0022-3514.91.1.188

Oyserman, D., Destin, M., & Novin, S. (2015). The Context-Sensitive Future Self: Possible Selves Motivate in Context, Not Otherwise. *Self and Identity*, *14*(2), 173–188. https://doi.org/10.1080/15298868.2014.965733