

BÁO CÁO NGHIÊN CỨU NEWARK

Bản dịch tiếng Việt từ “Newark Study White Paper”

CÁC ĐIỂM CHÍNH

- Báo cáo chuyên đề này trình bày kết quả của một nghiên cứu dọc kéo dài ba năm đối với học sinh tiểu học và trung học cơ sở tại một học khu đô thị có tỷ lệ nghèo cao ở Hoa Kỳ. Mục tiêu của nghiên cứu là đánh giá tác động nhân quả của Khan Academy đối với thành tích học tập của học sinh.
- Những học sinh tăng mức độ sử dụng Khan Academy qua từng năm cũng ghi nhận mức cải thiện tương ứng trong tốc độ tiến bộ học tập.
- Xu hướng này nhất quán giữa các phân nhóm nhân khẩu học của học sinh, như chủng tộc/sắc tộc, giới tính và điều kiện đủ tiêu chuẩn tham gia chương trình bữa trưa miễn phí/giảm giá.

Giới thiệu

Newark North Ward (sau đây gọi là học khu) là một phần của Học khu Công lập Newark. Trong năm học 2023-2024, 13 trường tại North Ward giảng dạy cho hơn 6.000 học sinh từ lớp 3 đến lớp 8. Phần lớn học sinh là người Hispanic/Latino và người Da đen hoặc người Mỹ gốc Phi. Hơn 90% học sinh tham gia chương trình Bữa trưa Miễn phí và Giảm giá.

Vào tháng 11 của năm học 2022-2023, học khu đã hợp tác với Khan Academy để sử dụng giải pháp Khan Academy Districts (KAD) như một công cụ hỗ trợ cho chương trình giảng dạy. Trước đó, một số trường trong học khu đã sử dụng phiên bản Khan Academy miễn phí trong các lớp Toán. Trong năm học 2023-2024, toàn bộ 13 trường trong học khu triển khai KAD trong suốt năm học.

Phân tích được trình bày trong báo cáo này đánh giá tác động của việc sử dụng KAD đối với kết quả trong kỳ Đánh giá Học tập Học sinh New Jersey (New Jersey Student Learning Assessments - NJSLA). NJSLA là bài kiểm tra trên máy tính được tổ chức hằng năm trên toàn bang cho học sinh từ lớp 3 đến lớp 11 tại New Jersey nhằm đo lường tiến bộ của các em ở các môn Ngữ văn Anh (ELA), Toán và Khoa học theo chuẩn học tập của bang. Báo cáo này tập trung vào điểm Toán NJSLA.

Ở các lớp và khóa học khác nhau, điểm theo thang đo Toán NJSLA nằm trong khoảng 650-850 trên một thang điểm được chuẩn hóa theo chiều dọc. Do đó, một điểm thang đo thể hiện cùng một mức năng lực tiềm ẩn bất kể lớp học hay dạng bài kiểm tra; ví dụ, 750 tương ứng với mức “Đạt kỳ vọng” ở mọi khối lớp.

Các mức thành tích được xác định như sau: Mức 1 (<700, “Chưa đạt”); Mức 2 (700-724, “Đạt một phần”); Mức 3 (725-749, “Tiệm cận”); Mức 4 (750-786, “Đạt”); và Mức 5 (≥ 787 , “Vượt”).

Khoảng thời gian phân tích bao gồm ba năm học, từ 2021-2022 đến 2023-2024, và sử dụng các phương pháp bán thực nghiệm để ước tính tác động nhân quả của KAD - được vận hành hóa bằng số giờ học và số kỹ năng được nâng lên mức thành thạo hoặc cao hơn (STP+) - đối với kết quả NJSLA ở thời điểm sau. Phần 2 trình bày chi tiết mẫu phân tích năm 2023 - 2024; Phần 3 mô tả cách xây dựng dữ liệu bảng dọc và phương pháp thống kê.

Mẫu nghiên cứu

Bảng 1 dưới đây cho thấy phân bố học sinh theo khối lớp và trường trong năm học 2023 - 2024. Phần lớn (87%) trong gần 6.700 học sinh từ lớp 3 đến lớp 8 thuộc diện được hưởng bữa trưa miễn phí hoặc giảm giá. 48% học sinh là nam, 77% là người Hispanic hoặc Latino, 13% là người Da đen hoặc người Mỹ gốc Phi. Khoảng 30% học sinh là người học tiếng Anh (ELL).

Bảng 1. Quy mô mẫu tổng thể

Trường	n	Lớp 3	Lớp 4	Lớp 5	Lớp 6	Lớp 7	Lớp 8
Abington Avenue	561	81	92	91	102	100	95
Dr. E. Alma Flagg	326	44	51	50	65	56	60
Dr. William H. Horton	453	55	65	62	91	86	94
Elliott Street	522	68	84	81	96	99	94
First Avenue	688	118	110	114	120	114	112
Franklin	469	62	73	104	84	76	70
Luis Munoz Marin	404	62	65	57	71	79	70
McKinley	387	51	71	63	62	69	71
Park	468	63	78	77	92	77	81
Rafael Hernandez	312	45	60	58	39	52	58
Ridge Street	360	52	57	62	59	58	72
Roberto Clemente	460	67	75	79	73	88	78
Salome Urena	271	44	54	50	49	48	26

Bảng 2 tổng hợp thành tích của học sinh theo mức kết quả Toán NJSLA trong năm học 2023-2024. Xét trên dữ liệu thực tế, mức tăng trung bình qua từng năm ở cấp học khu trên bài đánh giá Toán NJSLA theo thang điểm dọc là 2–4 điểm thang đo ($\mu \approx 3,4$; $\sigma \approx 21$) ; chưa đủ để phần lớn học sinh vượt qua các ranh giới mức thành tích cách nhau 25 điểm. Do đó, trạng thái thành thạo theo nhóm tương đối ổn định qua các năm (xem Bảng 3).

Bảng 2. Kết quả NJSLA của toàn bộ mẫu

Năm học	Điểm TB	Mức 1	Mức 2	Mức 3	Mức 4*
2021–2022	715,4	30%	35%	23%	12%
2022–2023	719,0	27%	32%	25%	16%
2023–2024	721,7	25%	32%	24%	18%

Lưu ý: Mức 4 và Mức 5 được gộp để trình bày ngắn gọn.

Bảng 3 cho thấy tỷ lệ học sinh tăng hoặc giảm ít nhất một mức thành tích NJSLA trong năm học 2023–2024 so với năm học 2022–2023.

Bảng 3. Thay đổi mức NJSLA theo khối lớp

Khối	n	Δ NJSLA	Tăng	Giảm
Tổng thể	4.241	2,7	1.126 (27%)	867 (20%)
4	816	3,4	222 (27%)	139 (17%)
5	835	3,6	213 (26%)	163 (20%)
6	879	0,2	195 (22%)	202 (23%)
7	905	9,2	315 (35%)	97 (11%)
8	806	-3,4	181 (22%)	266 (33%)

Lưu ý: “Tăng/Giảm” chỉ những học sinh tăng hoặc giảm một mức NJSLA so với năm trước. Khối 3 không được đưa vào vì đây là năm đầu tiên các em làm bài kiểm tra.

Mức độ sử dụng Khan Academy

Học khu sử dụng các khóa học Toán Illustrative Mathematics của Khan Academy như một công cụ hỗ trợ trong thời gian dạy Toán cốt lõi. Trong hai năm gần nhất, thời gian trung bình dành cho luyện tập trên Khan Academy tăng từ 20 giờ lên 26 giờ trong một năm học, tương đương lần lượt khoảng 25 và 40 phút mỗi tuần. Dựa trên các nghiên cứu hiệu quả trước đây, Khan Academy khuyến nghị ít nhất 30 phút luyện tập mỗi tuần. Trong năm học 2022-2023, khoảng 50% học sinh đạt mức sử dụng này; sang năm học 2023-2024, con số tăng lên hơn 70%.

Mặc dù thời gian học là một chỉ báo hữu ích về mức độ tương tác, việc làm chủ các kỹ năng trên Khan Academy thường là thước đo chính xác hơn về việc học trên nền tảng. Đạt mức làm chủ đối với các kỹ năng cụ thể trên Khan Academy cho thấy học sinh thực sự đã làm các bài luyện tập thay vì chỉ xem video, chẳng hạn và giải đúng các bài đó. Nếu sử dụng thời gian hiệu quả, học sinh có thể nâng từ 2 kỹ năng trở lên mỗi tuần lên mức thành thạo hoặc làm chủ, tương đương khoảng 60 kỹ năng trở lên trong một năm học. Trong năm học 2022-2023, khoảng 30% học sinh nâng được từ 60 kỹ năng trở lên; sang năm học 2023-2024, hơn 50% học sinh đạt mốc này.

Đặc điểm của mẫu học sinh trong năm học 2023-2024 nhất quán với những gì quan sát được trong các năm 2022-2023 và 2021-2022.

Phương pháp nghiên cứu

Trong bối cảnh nghiên cứu quan sát, việc tiếp nhận can thiệp mang tính nội sinh: những học sinh có thành tích cao hơn hoặc có động lực hơn, cùng giáo viên của các em, có xu hướng sử dụng các công cụ hỗ trợ như Khan Academy nhiều hơn. Vì vậy, việc so sánh đơn giản người dùng với người không dùng sẽ bị sai lệch bởi cả các hiệp biến quan sát được và không quan sát được (ví dụ: động lực, chất lượng giáo viên, sự hỗ trợ của gia đình).

Các chiến lược ghép cặp truyền thống (ví dụ: ghép cặp chính xác, láng giềng gần nhất hoặc ghép cặp theo điểm xu hướng) chỉ giảm sai lệch trong phạm vi tất cả các hiệp biến liên quan đều được quan sát. Những đặc điểm không đo lường được vẫn không được kiểm soát, kích thước mẫu sau ghép cặp bị thu hẹp, và các trọng số xu hướng cực đoan có thể làm tăng phương sai của ước lượng. Thông thường chỉ có điểm kiểm tra ban đầu, khối lớp, trường và giáo viên được quan sát và dùng để ghép cặp, nhưng những yếu tố này hiếm khi đủ để làm cho nhóm dùng và không dùng trở nên tương đương.

Một cách tiếp cận thay thế là sử dụng thiết kế dữ liệu bảng, trong đó học sinh được quan sát tại nhiều thời điểm. Việc nhận diện tác động nhân quả của can thiệp xuất phát từ sự thay đổi mức độ can thiệp ở cấp học sinh từ thời điểm t sang $t + 1$. Ví dụ, một học sinh sử dụng Khan Academy 15 phút mỗi tuần ở năm 1 rồi tăng thêm trung bình 10 phút mỗi tuần ở năm 2 được kỳ vọng sẽ tăng điểm kiểm tra ở năm 2 so với năm 1.

Trong thiết kế này có hai nguồn gây nhiễu: các yếu tố gây nhiễu không thay đổi theo thời gian và các yếu tố gây nhiễu thay đổi theo thời gian. Nhóm thứ nhất bao gồm những đặc điểm tương đối cố định của học sinh và giáo viên, như năng lực tổng thể của học sinh và sở

thích/phong cách lớp học của giáo viên. Nhóm thứ hai đến từ những yếu tố thay đổi theo thời gian cùng với thay đổi ở mức độ can thiệp. Ví dụ, nếu một trường áp dụng chương trình học mới ở năm 2, đồng thời tăng mức độ can thiệp và triển khai các biện pháp khác nhằm cải thiện thành tích Toán, thì tác động của can thiệp sẽ bị nhiễu bởi tác động của những biện pháp khác.

Hồi quy tác động cố định với tác động cố định của học sinh và giáo viên

Trong nghiên cứu này, chúng tôi xây dựng một bộ dữ liệu bằng cách quan sát học sinh. Bộ dữ liệu giới hạn ở những học sinh được quan sát trong ít nhất hai năm bất kỳ của giai đoạn ba năm. Học sinh được đưa vào phân tích không bị giới hạn theo mức độ sử dụng; tuy nhiên, theo cấu trúc của kỹ thuật phân tích, những học sinh không sử dụng Khan Academy trong tất cả các năm mà các em được quan sát sẽ bị loại khỏi phân tích. Bộ dữ liệu bằng gồm 5.200 học sinh. Chúng tôi sử dụng hồi quy bình phương tối thiểu với tác động cố định để xử lý các yếu tố gây nhiễu không thay đổi theo thời gian đã mô tả ở trên.

Cách tiếp cận này hoạt động bằng cách so sánh mỗi học sinh với chính bản thân em đó qua nhiều thời điểm, sau khi trừ giá trị trung bình tổng thể của mức độ can thiệp và kết quả của từng học sinh khỏi mỗi quan sát tức tác động cố định ở cấp học sinh. Quá trình này, được gọi là “demeaning”, về thực chất loại bỏ các yếu tố không thay đổi theo thời gian. Tương tự, nếu giáo viên được quan sát qua nhiều năm, có thể loại bỏ giá trị trung bình của giáo viên về điểm kiểm tra hoặc mức sử dụng can thiệp bằng cùng quy trình.

Cách tiếp cận này ước tính tác động của thay đổi trong mức độ can thiệp của cùng một học sinh (mức sử dụng Khan Academy) đối với thay đổi trong kết quả của chính học sinh đó, qua đó kiểm soát hiệu quả mọi đặc điểm ổn định của học sinh và giáo viên. Tuy nhiên, nó không xử lý được hoàn toàn các yếu tố gây nhiễu thay đổi theo thời gian. Mô hình của chúng tôi xử lý những cú sốc theo thời gian ảnh hưởng chung đến tất cả học sinh bằng cách đưa tác động cố định theo năm vào mô hình, nhưng không thể loại bỏ các cú sốc có tương quan với cả mức sử dụng và kết quả ở cấp trường hoặc lớp học (ví dụ: đồng thời áp dụng một chương trình Toán mới trong năm t và chương trình đó cũng làm tăng mức sử dụng Khan). Khi dữ liệu cho phép, chúng tôi kiểm tra độ vững bằng cách bổ sung tác động cố định trường $\times$ năm và giáo viên $\times$ năm, qua đó có thể loại bỏ các yếu tố gây nhiễu thay đổi theo thời gian do chính sách cấp trường hoặc thay đổi ở cấp lớp học; tuy nhiên, hiệu quả của cách tiếp cận này phụ thuộc vào mức độ sẵn có của dữ liệu.

Phân tích và kết quả

Phần lớn học sinh tăng mức sử dụng từ 1-20 giờ trong năm học. Khoảng 10% tăng mức sử dụng từ 35 giờ trở lên, và khoảng 25% học sinh giảm mức sử dụng.

Tác động của việc sử dụng Khan Academy đối với thành tích học tập của học sinh được ước tính bằng mô hình tác động cố định hai chiều sau:

$$Y_{it} = \beta T_{it} + \alpha_i + \lambda_t + \delta g + \gamma t + \varepsilon_{it} \quad (1)$$

Trong đó α_i là tác động cố định của học sinh, λ_t là tác động cố định của giáo viên, δg là tác động cố định của khối lớp, γt là tác động cố định của năm và ε_{it} là sai số. Hệ số β biểu thị mức thay đổi trung bình trong điểm NJSLA chuẩn hóa của cùng một học sinh tương ứng với một đơn vị tăng trong mức độ sử dụng Khan Academy (được đo bằng một STP+ hoặc một giờ học), với điều kiện các tác động cố định đã nêu được đưa vào mô hình. Tất cả các ước lượng hệ số được báo cáo với khoảng tin cậy 95% bền vững với phương sai thay đổi.

Bảng 4 cho thấy tác động ước tính của việc sử dụng Khan Academy đối với điểm NJSLA. Sai số chuẩn được phân cụm ở cấp học sinh và giáo viên; khoảng tin cậy 95% (CI) được thể hiện trong ngoặc vuông.

Bảng 4. Kết quả hồi quy tác động cố định

Biến	(1)	(2)
Số giờ học	0,006*** [0,005; 0,008]	
STP+		0,003*** [0,003; 0,004]
Số quan sát	12.942	12.942
R ²	0,883	0,885
R ² hiệu chỉnh	0,794	0,796
R ² within	0,013	0,022
R ² within hiệu chỉnh	0,013	0,022
RMSE	0,34	0,34
Sai số chuẩn	theo id & teacherId	theo id & teacherId
FE: id	X	X
FE: teacherId	X	X
FE: grade	X	X
FE: year	X	X

* p < 0,05; ** p < 0,01; *** p < 0,001.

Tác động điều trị trung bình

Sau khi kiểm soát tác động cố định theo năm, khối lớp, học sinh và giáo viên, chúng tôi ước tính rằng mỗi giờ luyện tập Khan Academy bổ sung làm tăng điểm kiểm tra NJSLA thêm +0,006 SD [0,005; 0,008], và mỗi kỹ năng bổ sung được nâng lên mức thành thạo làm tăng điểm NJSLA thêm +0,003 SD [0,003; 0,004]. Trong 32 tuần của một năm học, điều này tương ứng với mức tăng trung bình +0,10 SD [0,008; 0,123] khi tăng thời gian luyện tập Khan Academy thêm 30 phút mỗi tuần, và mức tăng trung bình +0,19 SD [0,16; 0,22] khi tăng thêm 60 kỹ năng đạt mức thành thạo.

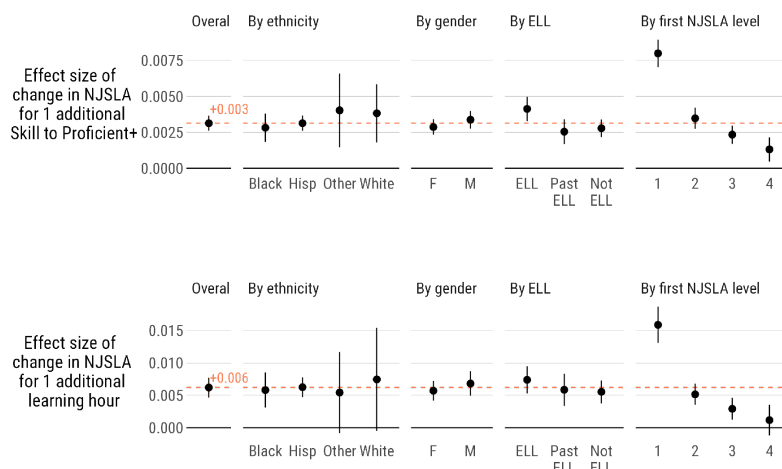
Xét về ý nghĩa thực tiễn, các tác động này tương ứng với khoảng 2- 4 điểm NJSLA bổ sung khi tăng thời gian luyện tập 30 phút mỗi tuần (18 giờ trong cả năm học), hoặc 5–8 điểm NJSLA bổ sung khi tăng mức độ làm chủ kỹ năng thêm 2 kỹ năng mỗi tuần (tương đương 60 kỹ năng trong năm học).

Ghi chú công thức trong tài liệu gốc: ¹ $0,0062 \times 16$ [0,0047 $\times$ 16; 0,0077 $\times$ 16]. ² $0,0031 \times 60$ [0,0026 $\times$ 60; 0,0031 $\times$ 60].

Tác động khác biệt giữa các nhóm

Chúng tôi xem xét tác động không đồng nhất theo các phân nhóm nhân khẩu học và điểm kiểm tra NJSLA ban đầu, tức điểm NJSLA đầu tiên có sẵn tùy theo năm mà học sinh lần đầu được quan sát. Hình 1 trình bày tác động ước tính cho từng phân nhóm sau khi kiểm soát cùng các tác động cố định như trong mô hình ban đầu. Trong khi tác động nhất quán giữa các nhóm nhân khẩu học như giới tính, sắc tộc và tình trạng ELL, tác động đối với những học sinh có điểm ban đầu dưới mức kỳ vọng (<700 trên thang NJSLA) lớn hơn một cách có ý nghĩa thống kê. Điều này gợi ý rằng những học sinh yếu hơn có thể hưởng lợi nhiều hơn.

Các yếu tố gây nhiễu thay đổi theo thời gian có tương quan đồng thời với mức sử dụng và tư cách thành viên nhóm (ví dụ: các biện pháp phụ đạo hướng tới học sinh điểm thấp) có thể làm thiên lệch tăng mối tương tác giữa năng lực ban đầu và mức sử dụng. Ví dụ, nếu học sinh điểm thấp tập trung ở những trường vừa khuyến khích sử dụng Khan vừa triển khai các biện pháp can thiệp khác, tác động này có thể bị thổi phồng một cách giả tạo.



Hình 1. Tóm tắt các tác động điều trị không đồng nhất.

Dịch các nhãn chính trong hình: Overall = Tổng thể; By ethnicity = Theo sắc tộc; By gender = Theo giới tính; By ELL = Theo tình trạng ELL; By first NJSLA level = Theo mức NJSLA ban đầu. Biểu đồ trên: mức thay đổi NJSLA tương ứng với 1 kỹ năng bổ sung đạt Proficient+; biểu đồ dưới: mức thay đổi NJSLA tương ứng với 1 giờ học bổ sung.

Để kiểm tra khả năng này, chúng tôi ước lượng lại mô hình với tác động cố định trường $\times$ năm và giáo viên $\times$ năm. Các thành phần này sẽ hấp thụ mọi chính sách hoặc chương trình áp dụng chung cho tất cả học sinh trong cùng một trường trong cùng năm học (ví dụ: chương trình phụ đạo toàn trường dành cho học sinh điểm thấp), hoặc các nỗ lực của từng giáo viên nhằm cải thiện điểm kiểm tra trong lớp.

Sau khi bổ sung các tác động cố định này, tương tác giữa năng lực ban đầu và mức sử dụng vẫn có ý nghĩa thống kê. Có thể có một số cách giải thích: (1) giáo viên có thể tổ chức các buổi hỗ trợ riêng cho học sinh điểm thấp, hoặc phụ đạo cá nhân đang diễn ra và có tương quan với cả điểm kiểm tra lẫn mức sử dụng Khan; (2) học sinh điểm thấp thực sự có thể hưởng lợi nhiều hơn từ một giờ luyện tập hoặc một kỹ năng; (3) học sinh điểm thấp có thể sử dụng thời gian theo cách có lợi hơn, chẳng hạn tập trung vào các kiến thức tiên quyết. Hai cách giải thích sau có lẽ ít khả năng hơn so với việc học sinh được phụ đạo có mục tiêu.

Tuy nhiên, đối với học sinh khá/giỏi (mức ban đầu 4), tác động của thời gian học gần như bằng 0 trong khi tác động của một STP thì không, điều này gợi ý rằng có thể tồn tại cách sử dụng thời gian trên Khan hiệu quả hơn.

Chúng tôi thực hiện thêm hai kiểm tra đối với mô hình. Thứ nhất, chúng tôi đưa mức sử dụng trễ $X_{i,t-1}$ vào để tiến hành kiểm định tác động trễ, nhằm loại trừ khả năng hệ số β của can thiệp đang phản ánh tác động suy giảm của mức sử dụng trước đó.

Thứ hai, chúng tôi thực hiện kiểm định giả được bằng cách đưa mức sử dụng của năm sau $X_{i,t+1}$ vào mô hình và ước tính tác động của nó đối với điểm NJSLA của năm hiện tại. Nếu hệ số của mức sử dụng năm sau có ý nghĩa thống kê, điều đó sẽ hàm ý rằng thành tích của năm hiện tại tác động đến quyết định sử dụng trong năm tương lai, từ đó vi phạm giả định ngoại sinh. Để thực hiện các kiểm định này, chúng tôi giới hạn mẫu ở những học sinh được quan sát trong cả ba năm ($n = 2.924$). Bảng 5 trình bày các hệ số từ hai mô hình này.

Tác động trễ có ý nghĩa thống kê, cho thấy có một phần tác động được duy trì từ năm trước. Vì tác động này chỉ bằng khoảng một phần ba tác động điều trị và được ước tính sau khi đã loại ảnh hưởng của tác động chính, nó không chi phối kết quả chính; thay vào đó, chỉ cho thấy học sinh duy trì được một phần lợi ích nhỏ từ việc luyện tập của năm trước. Hệ số của mức sử dụng năm sau không có ý nghĩa thống kê, do đó giả định ngoại sinh nghiêm ngặt chưa bị bác bỏ.

Bảng 5. Kết quả kiểm định độ vững

Biến	(1)	(2)
Số giờ học	0,008*** [0,005; 0,010]	0,007*** [0,004; 0,009]
Số giờ học (trễ)	0,003* [0,001; 0,005]	
Số giờ học (năm sau)		-0,002 [-0,005; 0,001]
Số quan sát	5.701	5.468
R^2	0,911	0,916
R^2 hiệu chỉnh	0,803	0,802
R^2 within	0,014	0,016
R^2 within hiệu chỉnh	0,013	0,015
RMSE	0,30	0,28
Sai số chuẩn	theo id & teacherId	theo id & teacherId
FE: id	X	X

FE: teacherId	X	X
FE: grade	X	X
FE: year	X	X

* $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$.

Nhìn chung, các kết quả ủng hộ một tác động nhân quả ở mức vừa phải nhưng có ý nghĩa về mặt giáo dục của việc duy trì tương tác với Khan Academy đối với thành tích Toán NJSLA, với lợi ích lớn hơn đối với những học sinh bước vào dữ liệu bảng khi còn ở mức thấp đáng kể so với chuẩn thành thạo.

Kết luận

Dựa trên dữ liệu bảng tác động cố định hai chiều ở cấp học sinh và giáo viên trong ba năm của 5.500 học sinh Newark North Ward, chúng tôi nhận thấy việc tăng mức độ tương tác với Khan Academy có liên quan đến mức tăng điểm Toán NJSLA có ý nghĩa cả về mặt thống kê lẫn thực tiễn.

Cụ thể, tác động của mỗi giờ luyện tập bổ sung là mức tăng 0,006 SD trong điểm thang đo; 30 phút mỗi tuần (18 giờ mỗi năm) tương ứng với 0,10 SD (4 điểm). Tác động của việc làm chủ mỗi kỹ năng được nâng cấp là 0,003 SD; đạt 60 kỹ năng làm chủ trong một năm học tạo ra mức tăng 0,18 SD (7 điểm). Lợi ích tương đối đồng đều giữa giới tính, sắc tộc và tình trạng ELL, nhưng lớn hơn đáng kể đối với những học sinh bắt đầu dưới ngưỡng thành thạo (<650 , theo nguyên văn tài liệu).

Các kết quả này hàm ý rằng việc sử dụng Khan Academy có thể làm tăng gấp đôi mức tăng điểm NJSLA hằng năm điển hình của học khu.

Tuy nhiên, vẫn còn các hạn chế sau: không thể loại trừ hoàn toàn các yếu tố gây nhiễu còn sót lại và thay đổi theo thời gian ở cấp lớp học và học sinh; những học sinh luôn có mức sử dụng bằng 0 bị loại khỏi mẫu phân tích theo thiết kế; và tác động chỉ được đo đối với môn Toán.

Ngay cả với những lưu ý này, bằng chứng cho thấy việc sử dụng Khan Academy một cách bền bỉ và tập trung vào làm chủ kỹ năng có thể là một thành tố hiệu quả của dạy học Toán ở các khối trung học cơ sở tại những học khu đô thị có tỷ lệ nghèo cao.

Ghi chú của bản dịch

Bản dịch giữ nguyên các số liệu và cách diễn đạt định lượng của tài liệu gốc. Tài liệu nguồn có một số điểm không nhất quán nội bộ: phần phương pháp nêu bộ dữ liệu bảng gồm 5.200 học sinh, trong khi phần kết luận nêu 5.500 học sinh; phần phân tích xác định nhóm dưới kỳ vọng là



<700 trên thang NJSLA, trong khi phần kết luận ghi ngưỡng <650. Các chi tiết này được giữ nguyên, không tự ý hiệu chỉnh.