New Types of Data and Their Applications in Educational Research

The emergence of new types of data and methods exhibit an exciting opportunity for education researchers to examine phenomena that were hard to measure before. Detailed data collected from social media, written or oral communication, and even medical fields such as human genotype, open up new possibilities for education research. At the same time, such data are often unstructured and high-dimensional, posing challenges with regard to how to use them effectively and rigorously. This symposium convenes leading researchers who work in big data and their application to education to demonstrate the value and challenges of such data and its implication for education research. Our symposium aims to provide an opportunity for researchers from a wide range of research areas to spark conversations on how to use new types of data in educational research.

The first two papers apply text-as-data methods. New technologies have made large volumes of textual data increasingly available to researchers. Text information from education policy making, classroom processes, written materials from students and teachers, and many other sources provide an unprecedented opportunity for education researchers to construct new metrics, look deeper into the process of schooling, and answer important or new research questions. One paper focuses on school reform. It uses a large volume of text data from school planning and implementation reports generated by underperforming schools in the state of Washington. This paper conducts thorough validity checks on the measures generated from topic modeling and also examines which practices can have high-leveraging effects on student outcomes. This paper contributes to the knowledge of effective school management practices and policies. The second paper studies freshman seminars organized by interest groups (FIG) on student graduation and first-year retention. Using propensity score matching, this paper first documents that FIG students had higher graduation and re-enrollment rates than non-FIG counterparts. Then, to study how exactly freshmen found FIG helpful, they used Topic Modeling to code over 12,000 FIG students' open-ended survey responses to derive measures on students' non-academic traits.

The third paper uses social network data to study teachers' resource acquisition and sharing in social media. The authors apply both interrupted time series design and piecewise regression model to examine how teachers' engagement with social media changes after the enactment of Common Core State Standards. The last paper concerns genotype data from the National Longitudinal Study of Adolescent to Adult Health (Add Health) and focuses on how teens' smoking behavior is affected by peers' genetics. The authors leverage the as-if-random distribution of grade-mates conditional on school-level variation in a nationally representative sample. Specifically, they evaluate whether the genetic propensity to smoke of one's peers affects one's own smoking behavior net of one's own genotype.

Promises and Challenges of Using Text as Data in Educational Research: Using the Study of School Improvement Strategies to Illustrate

Background: Quantitative program evaluation that employs rigorous quasi-experimental or experimental designs, although can readily inform whether to continue or terminate a given program, has often had limited influence on the theory of change employed by schools and districts (Singer, 2018). This limited influence is mainly due to the fact that our rigorous research is less able to reveal the increasing complexity of the interventions and the mechanisms by which these interventions achieve their effects (Hedges, 2018). This paper aims to use novel data sources—school planning and implementation reports—and text analysis methods to extract measures on school improvement processes, aiming to inform the questions of what processes made the program work, and how.

Purpose/Research Question: In this paper, we apply textual analyses, particularly Latent Dirichlet Allocation (LDA), to identify the key, fine-grid measures on school improvement strategies at a large scale under the era of NCLB waivers and School Improvement Grants (SIG) in Washington State. Specifically, we ask:

- What are the key school improvement strategies?
- How do they vary by schools and over time?
- How valid are these measures? How do these measures predict student outcomes?

Setting/Intervention: We focus on school reform efforts by the state-identified lowest performing schools in the state of Washington (WA) from 2010 to 2016. WA used three school improvement designations: Focus schools (identified as the lowest 10% for subgroups of students), Priority schools (the bottom 5%), and SIG (who were priority schools and implemented federal prescribed school turnaround models). Once a school is identified, typically the designation for this school remains for the three years. Schools would receive funding on top of their regular school budgets: SIG schools were funded primarily through federal grants, while Priority and Focus schools were funded primarily through state funds.

Participants/Sample: Table 1 summarizes the number of schools identified as SIG, priority and focus each year from 2010-11 to 2015-16. A total yields a unique of 318 schools and 613 school-year observations. Table 2 summarizes school contextual and performance of these identified schools, in contrast to other schools in the state. Overall, identified schools, on average, serve larger proportions of historically underserved students of color, low-income, and homeless students than non-identified schools. The students in the identified schools are also relatively low-achieving and are more likely to be chronically absent.

[Table 1 & 2 Here]

Data Sources: To generate more detailed knowledge on the mechanisms of change in schools, this project uses text data from the Comprehensive School Improvement Planning and Implementation Reports (CSIPIRs) by the Office of Superintendent of Public Instruction (OSPI) at WA. CSIPIRs are reported by schools through a web-based platform called Indistar. When using the Indistar system to build a school improvement plan, schools are required to lay out specific tasks needed to achieve each goal, including individual(s) responsible, the target completion date, and the frequency of the task (see an example in Figure 1). A school is allowed to plan as many tasks as needed to achieve a goal. The school is also required to update its report periodically to mark the completion date for completed tasks, add comments on implementation, and explain how they plan to sustain the task.

Among the CSIPIR data that we received and cleaned (2011–12 to 2015–16), about 55.2% of all unique tasks proposed were marked as completed with specific completion dates. *Our subsequent analyses use only these completed tasks because they represent the completion of resource allocation and schools' committed actions.* We have 25,486 completed unique tasks.

We then linked these reform process measures with school contextual and student outcome measures from state administrative database. We used both student *absenteeism* and *achievement on state standardized tests* to measure school improvement outcomes.

Data Analysis: We first condensed thousands of diverse CSIPIR text entries into a limited number of discrete and sensible categories, or topics, using LDA (Blei et al., 2003). We then conducted extensive validation of the LDA results: we used several model diagnostic statistics (such as topic coherence, exclusivity, and residual) that pointed to either the 15-topic model or the 20-topic model as the best fit; we also asked human coders to assess whether tasks that load highly on a given topic indicate coherent meaning (see our discussion below). The subjective evaluation led to the same conclusion that the 20-topic model was optimal. We established content validity or semantic coherence, internal structure, relations to other variables (including predictive validity) (Chan, 2014).

Findings: Our analysis of interview results from staff in 10 schools shows about 82% of the top 10 topics—based on the prevalence of topic proportion in the reports—were mentioned as top initiatives by school administrators during their interview. The high alignment offers much confidence in the text analysis results, at least, as good as interview data would provide, but on a much larger scale with relatively low cost. We also find that schools vary in the reform strategies they planned and implemented as evidenced in Table 3 of the topic means and standard deviations.

The results of regressing topic proportions on student outcomes show that several topics are significantly correlated with a lower absenteeism rate. As shown in Table 4, Topic 3 of "engaging parents about student academic and behavioral learning in schools" significantly lowers both full-day and part-day unexcused absence rates. Topic 9 of "teacher team activities (e.g., reviewing data, planning, aligning standards, developing interventions) is significantly correlated with the reduction of all five absenteeism measures. Moreover, Table 5 shows that Topic-15—"Setting goals for, recognizing, and monitoring teachers' and students' growth"— is significantly positively correlated with the increase in school-level average student achievement, particularly in math.

[Table 3, Table 4 and Table 5]

Conclusions: The text analysis methods can extract school improvement strategies that align with interview data and predict student outcomes. These reform strategies can be directly used by practitioners to develop theories of action for enacting changes in schools. This text analysis approach can be fruitfully used either to explore the underlying dimensions of the data to identify patterns in implementation, or to be integrated with experimental or quasi-experimental approaches to make credible inferences about the policy effects. However, text data themselves need extensive validation to show their credibility (Grimmer & Stewart, 2013).

[word count: 1000]

References

Hedges, L. V. (2018). Challenges in Building Usable Knowledge in Education. *Journal of Research on Educational Effectiveness, 11* (1), 1-21, DOI: 10.1080/19345747.2017.1375583.

Singer, J. D. (2018) Even More Challenges in Building Usable Knowledge in Education. *Journal of Research on Educational Effectiveness, 11* (1), 22-24, DOI: 10.1080/19345747.2017.1402397.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, *3*(Jan), 993-1022.

Chan, E. E. (2014). Standards and guidelines for validation practices: Development and evaluation of measurement instruments. In *Validity and Validation in Social, Behavioral, and Health Sciences* (pp. 9-24). Springer International Publishing.

Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis, 21*, 267–297.

Figures and Tables

Figure 1. Examples of Tasks Written in the Comprehensive School Improvement Planning and Implementation Reports (CSIPIRs)

Tasks:		
1. Math team will implement learning target assessments aligned to identified grade-level power standards. Team will collaboratively review assessment data every month and adjust instruction accordingly. Team will also work to align instructional practices and math specific vocabulary during monthly meetings and weekly interventions meetings.		
	Assigned to:	Jamila Davis ← Principal
	Added date:	09/30/2016
	Target Completion Date:	06/13/2014
	Frequency:	monthly
	Comments:	Our math team completed this objective by the end of January. We have now built agreed on power standards in each grade, built standards-based learning target assessments, and agreed on some common vocabulary related to these standards.
	Task Completed:	1/29/2014 12:00:00 AM
2. Teachers will work to align instruction and instructional vocabulary around specific literacy concepts. Friday interventions teams will include this work as a on-going agenda, while also creating and monitoring student assessments/performance pertaining to these concepts.		
	Assigned to:	Erin Rebich ← Instructional Coach
	Added date:	09/30/2016
	Target Completion Date:	06/13/2014
	Frequency:	weekly
	Comments:	We have done very much specifically in this area. We are engaged in progress of monitoring of students and on-going evaluation of support services for students below grade-level, but we have not addressed literacy strategies specifically as an on-going task. We may not be able to complete this during the 2013-14 year. By the end of May we have accomplished these goals in Math and Science.
	Task Completed:	5/28/2014 12:00:00 AM

Note: Staff names are pseudonyms.

Table 1. Number of Treatment Schools that Have Test Score Data by Reform Type and Cohort

School Year	SIG	Priority	Focus
2015-16	0	68	116
2014-15	0	56	131
2013-14	10	34	79
2012-13	24	15	66
2011-12	24	0	0
2010-11	0	0	0

Note. This is for the regressions using achievement or chronic absenteeism as the outcome. Many of them miss test scores from 2014. Many of them are missing reports. Quil Ceda Tulalip Elementary is missing prior since it is newly combined from 2015.

Table 2. Student Demographic Characteristics and Education Outcomes by Reform Type

Student Demographics and Outcomes	SIG	Priority	Focus	Non-Reform
Pct. White	0.24 (0.25)	0.34 (0.3)	0 .39 (0.25)	0.56 (0.29)
Pct. African American	0.12 (0.18)	0.07 (0.13)	0.05 (0.09)	0 .04 (0.07)
Pct. Hispanic	0.39 (0.33)	0.39 (0.32)	0.42 (0.26)	0.16 (0.19)
Pct. Asian	0.06 (0.09)	0.03 (0.07)	0.04 (0.07)	0.05 (0.08)
Pct. Pacific Islander	0.01 (0.02)	0.01 (0.02)	0.01 (0.02)	0.01 (0.02)
Pct. Native American	0.1 (0.2)	0.11 (0.22)	0.03 (0.08)	0.02 (0.08)
Pct. Multi-racial	0.05 (0.05)	0.05 (0.05)	0.05 (0.04)	0.05 (0.05)
Pct. Eligible for free and reduced priced lunch	0.79 (0.18)	0.73 (0.23)	0.69 (0.17)	0.41 (0.27)
Pct. English language learner	0.39 (0.27)	0.33 (0.29)	0.36 (0.23)	0.17 (0.19)
Pct. Homeless	0.06 (0.05)	0.06 (0.06)	0.05 (0.06)	0.03 -0.04
Pct. Special Education	0.14 (0.05)	0.14 (0.09)	0.14 (0.05)	0.16 (0.21)
Prior student academic achievement	-0.6 (0.19)	-0.58 (0.3)	-0.33 (0.23)	
Student academic achievement	-0.46 (0.26)	-0.47 (0.34)	-0.31 (0.29)	-0.05 (0.44)
Pct. chronic absenteeism - full day	0.39 (0.14)	0.34 (0.16)	0.33 (0.12)	0.24 (0.17)
N(school-year)	112	278	492	17404

Note. Standard deviations are reported in parentheses. The data range from 2010 to 2016. Non-reform schools do not have prior student outcome measures because prior is defined as pre-reform years.

Table 3. Descriptives of Topic Proportions

Reform Strategy Topics	Mean Coherence Rating	Mean (SD)
1. Interventions and supports to promote positive student behaviors	4	0.07 (0.068)
2. General parent and community outreach	4	0.045 (0.055)
3. Engaging parents about student academic and behavioral learning in schools	4	0.084 (0.078)
4. Planning, providing, and evaluating professional development for instructional improvement	4	0.068 (0.059)
5. Monitoring student progress and using data to develop interventions	4	0.057 (0.058)
6. (Low-coherence)	2	0.021 (0.054)
7. Using assessment data to identify students for targeted support	3	0.047 (0.06)
8. Extending instructional time and Aligning curriculum or assessments to standards	3	0.038 (0.074)
9. Teacher team (e.g., grade-level teams, PLCs) activities (e.g., reviewing data, planning, aligning standards, developing interventions)	4	0.061 (0.077)
10. Administering common assessments and disaggregating data to differentiate interventions	3	0.021 (0.039)
11. Leadership teams goal-setting and reviewing data for school improvement	4	0.095 (0.082)
12. Teacher instructional improvement via walkthroughs, observations, and feedback	4	0.075 (0.081)
13. (Low-coherence)	2	0.037 (0.046)
14. (Incoherent)	1	0.028 (0.042)
15. Setting goals for and recognizing teachers' and students' growth	3	0.038 (0.071)
16. (Low-coherence)	2	0.039 (0.048)
17. Extending learning time (or opportunities) for students and staff	3	0.044 (0.06)
18. Collecting, analyzing, and aligning student assessments	4	0.052 (0.051)
19. Improving special education	3.5	0.038 (0.049)
20. (Low-coherence)	2	0.043 (0.063)

Note. Topic proportions are at the school by year level.

Table 4. Regression results: topic proportions and student absenteeism

Topic	Full day	Part day	Unexcused full day	Unexcused part day	% Chronic full-day absence
Topic 1	2.404 (2.883)	-1.271 (2.646)	0.876 (2.005)	-0.522 (1.92)	0.0754 (0.081)
Topic 2	-6.101 (3.936)	-3.191 (3.636)	-3.516 (2.074)	-0.0931 (0.099)	-0.2 (0.121)
Topic 3	-2.436 (1.794)	-2.765 (3.314)	-3.849* (1.552)	-3.574* (1.791)	-0.0577 (0.064)
Topic 4	-5.109 (3.395)	-1.463 (4.694)	-2.244 (2.744)	-1.242 (3.416)	-0.129 (0.013)
Topic 5	5.494 (4.677)	1.151 (3.995)	3.145 (3.383)	1.292 (3.333)	0.155 (0.143)
Topic 7	3.767 (5.444)	1.545 (5.827)	2.257 (3.967	-0.737 (4.163)	0.131 (0.18)
Topic 8	10.06* (4.153)	2.688 (5.07)	11.8 (6.618)	3.264 (6.035)	0.181* (0.075)
Topic 9	-6.806** (2.259)	-12.91*** (3.625)	-6.059** (1.973)	-7.542** (2.407)	-0.162* (0.07)
Topic 10	-24.96*** (7.315)	-17.75 (9.905)	-10.73* (5.409)	-9.847 (8.041)	-0.674 (0.225)
Topic 11	3.270 (3.372)	5.895 (5.148)	2.727 (3.323)	4.254 (4.215)	0.1 (0.095)
Topic 12	1.375 (2.611)	6.858 (3.904)	2.318 (2.32)	4.559 (3.156)	0.0362 (0.074)
Topic 15	4.483 (5.447)	2.531 (5.633)	3.657 (4.135)	2.261 (4.175)	0.169 (0.164)
Topic 17	1.59 (5.556)	7.502 (7.715)	-0.342 (4.648)	8.644 (6.423)	-0.159 (0.0969)
Topic 18	0.204 (6.146)	2.26 (8.94)	-1.007 (5.214)	3.225 (6.923)	0.0969 (0.192)
Topic 19	1.304 (4.928)	2.702 (5.891)	1.648 (4.166)	1.801 (4.534)	-0.0265 (0.126)
N	220	220	220	220	220

Note. * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$

Table 5. Regression results: topic proportions and student achievement

Topic	Composite	Math	Reading
Topic 1	-0.274* (0.132)	-0.384* (0.17)	-0.149 (0.119)
Topic 2	-0.357* (0.145)	-0.508** (0.185)	-0.192 (0.161)
Topic 3	-0.0381 (0.12)	-0.0176 (0.137)	-0.00651 (0.118)
Topic 4	-0.183 (0.162)	-0.284 (0.201)	-0.0547 (0.156)
Topic 5	0.0818 (0.159)	-0.196 (0.173)	0.369 (0.205)
Topic 7	-0.271 (0.246)	-0.348 (0.36)	-0.231 (0.155)
Topic 8	0.219 (0.146)	0.282 (0.239)	0.17 (0.155)
Topic 9	-0.182 (0.143)	0.0276 (0.122)	-0.362 (0.233)
Topic 10	0.4 (0.249)	0.517 (0.271)	0.289 (0.265)
Topic 11	-0.027 (0.148)	-0.118 (0.167)	-0.0729 (0.145)
Topic 12	-0.086 (0.137)	0.00326 (0.162)	-0.167 (0.149)
Topic 15	0.307* (0.14)	0.439* (0.201)	0.197 (0.113)
Topic 17	-0.087 (0.226)	-0.141 (0.306)	-0.0446 (0.174)
Topic 18	0.196 (0.22)	0.0489 (0.272)	0.198 (0.247)
Topic 19	-0.015 (0.244)	0.0176 (0.271)	-0.0061 (0.296)
N	596	596	580
R-sq.	0.49	0.37	0.53

Note. * $p \leq 0.05$, ** $p \leq 0.01$

Helping Students FIG-ure It Out: A large-scale study of freshmen interest groups (FIGs) and student success

Background and Context

Freshman orientation seminars (freshman seminars) are courses dedicated to helping incoming students transition to college life, both socially and academically. The popularity and ubiquity of these courses has made them among the most studied course genre in American higher education [2, 5, 3]. That said, the existence and effectiveness of these seminars on college campuses across the U.S. continues to be called into question [4]. Additionally, these programs are still inadequately assessed, particularly with respect to practical considerations for their design [1]. Furthermore, although some prior studies have used randomized controlled trials (e.g. [8]), large scale and causally rigorous studies of seminar effectiveness using matched comparison groups are rare [7].

Purpose and Research Question

In this work, we gather data from institutional databases at a large, publicly-funded U.S. university (the University of Washington, UW) to examine the impact of freshman interest groups (FIGs, a type of freshman orientation seminar) on student outcomes - namely, graduation rates and first-year retention. Using propensity score matching on nearly 58,000 students across 12 student cohorts (and 18 years of data), we use a suite of variables on students prior to their post-secondary education to match students who enrolled in FIGs with those who did not. We then explore the differences between these groups in terms of educational outcomes while also further examining specific ethnic/racial groups (namely, Hispanic and under-represented minority students).

Data Collection and Analysis

De-identified student data was collected from the UW's data custodians in early 2017. This data included complete student transcript records (courses taken, grades, etc.), student demographic information (race, gender, ethnicity, etc.), and student entrance application information (high school GPA, entrance exam scores, etc.). Data for this project was limited to first-time, first-year students who first enrolled at the UW between 1998 to 2010. This totalled 57,979 students, of whom 32,572 enrolled in a FIG (56.2%) and 15,407 did not (43.8%).

In addition, we also used data from the College Board's enrollment planning services (EPS) data and U.S. Census data. The EPS data included information on students' major intentions in higher education, students' preferences with regards to post-secondary campuses, students' long-term educational attainment goals, parents' educational attainment, and parents' income levels was used. EPS data was only available on a high-school level and not at an individual level and, as such, the data from EPS was matched to students' high schools from their applications to the University. For the U.S. Census data, information on average income, average bachelor's degree attainment, and average high school completion for each ZIP code was pulled from data available online from the U.S. Census' American Fact Finder. The data was aligned to each individual student using the ZIP code from their application to the University.

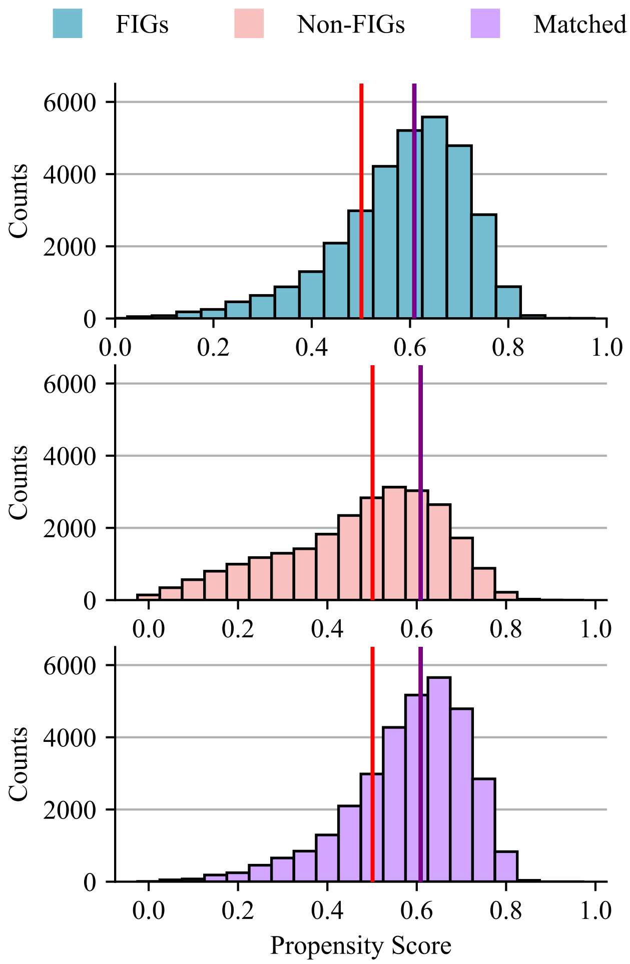


Figure 1: *Propensity score distributions for FIG students (top), non-FIG students (middle), and matched non-FIG students (bottom). The top and bottom distributions were used in the analysis. The left and right vertical lines across each distribution indicate the mean propensity score values for non-FIG students (0.50; from the middle distribution) and matched non-FIG students (0.61; from the bottom distribution), respectively. The mean of the FIG students was approximately equal to that of matched non-FIG students and, as such, the line indicating the mean for FIG students overlaps with that of matched non-FIG students.*

Because students select whether they will or will not enter a FIG (FIG and non-FIG groups, respectively), stratified propensity score matching (PSM) was used to mitigate selection bias or possible confounding variables [6]. We used students' demographic information and pre-college information from the registrar data as well information on their high schools from the EPS data and ZIP codes from the census data to calculate the propensity scores via a logit model. The model included 197 covariates. After calculating propensity scores, each student in the treatment (FIG) group was matched to students in the control (non-FIG) group using two-levels of stratification and fixed caliper widths. The students were first matched according to year of entry to the University and then by whether they were a STEM-interested student (i.e.

a student interested in science, technology, engineering, and/or math (STEM) fields). Using this dual stratification, every FIG student was paired to corresponding non-FIG students from the same entrance year and with the same (binary) indication of STEM interest, after which caliper matching was used to determine final pairings. The caliper was kept at one-tenth of the pooled standard deviation of all propensity scores and students were matched one-to-many, with replacement. The distributions of propensity scores can be seen in Figure 1.

To examine specific student groups, each Hispanic and each under-represented FIG student was matched to respective Hispanic and under-represented non-FIG students. The same strategy of dual stratification with caliper matching in a one-to-many manner was employed for this round of matching.

Results and Conclusions

Of the FIG students, 32,512 FIG students (99.8%) had at least one non-FIG student matched and 30,257 FIG students (92.9%) had at least 20 non-FIG students matched. This indicates a high level of common support for PSM. Each FIG student was matched to an average (±SD) of 66.1 (±34.0) non-FIG students. Of the 1,699 Hispanic students in FIGs, 1,546 (91.0%) were matched to at least one Hispanic non-FIG student; of the 1,578 under-represented minority students in FIGs, 1,428 (90.5%) were matched to at least one under-represented non-FIG student.

Table 1: *Graduation and re-enrollment rates for students after PSM*

Matching	Measure	FIG	non-FIG	Difference
All Students	Graduation	81.60%	74.57%	+7.03%
	Re-Enrollment	94.18%	90.49%	+3.69%
Hispanic	Graduation	77.96%	69.32%	+8.64%
	Re-Enrollment	93.06%	87.28%	+5.78%
Under-Represented	Graduation	76.81%	62.81%	+14.00%
	Re-Enrollment	94.75%	88.18%	+6.57%

Graduation and re-enrollment rates for FIG and non-FIG students after PSM are shown in Table 1. After matching, FIG students tended to have substantially higher graduation and re-enrollment rates than their matched non-FIG peers (differences of 7.0 and 3.7 percentage points, respectively). It should be noted, however, that for both groups, the re-enrollment rates are greater than 90%, thereby decreasing the potential maximum difference between the two groups in terms of percentage points.

The graduation and re-enrollment rates for Hispanic and under-represented FIG students were substantially higher than non-FIG students. This difference is much more pronounced than the estimated FIG effects for all students in Table 1. It should be noted that the graduation rates for both Hispanic and under-represented students were still below the University's average across all students. These facts could go hand-in-hand as the lower average rates across these groups allows for more differential gains to be realized for students attending FIGs. We know that the FIG curriculum at the University is not tailored to specific student groups based on race, ethnicity, or family backgrounds. Future research is needed to better explain this greater effect of FIGss on Hispanic and under-represented students.

Word Count: 1000

References

[1] B. O. Barefoot. The first-year experience. *About Campus*, 4(6):12–18, 2000.

[2] B. O. Barefoot, C. L. Warnock, M. P. Dickinson, S. E. Richardson, and M. R. Roberts. Exploring the evidence: Reporting outcomes of freshman seminars. Technical Report 2, National Resource Center for the Freshman Year Experience & Students in Transition, 1998.

[3] J. B. Cuseo. Freshman orientation seminar at community colleges: A research-based rationale for its value, content, and delivery. 1997.

[4] A. M. Griffin and J. Romm. Exploring the evidence: Reporting research on first-year seminars. Technical Report 4, National Resource Center for The First-Year Experience & Students in Transition, 2008.

[5] R. D. Padgett, J. R. Keup, and E. T. Pascarella. The impact of first-year seminars on college students life-long learning orientations. *Journal of Student Affairs Research and Practice*, 50(2):133–151, 2013.

[6] P. R. Rosenbaum and D. B. Rubin. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55, 1983.

[7] C. A. Schnell and C. D. Doetkott. First year seminars produce long-term impact. *Journal of College Student Retention: Research, Theory & Practice*, 4(4):377–391, 2003.

[8] G. Strumpf and P. Hunt. The effects of an orientation course on the retention and academic standing of entering freshmen, controlling for the volunteer effect. *Journal of The First-Year Experience & Students in Transition*, 5(1):7–14, 1993.

The Decline of Virtual Resource Seeking: The Impact of the Common Core State Standards

Background: In the field of education, social media platforms have expanded and transformed teachers' traditional social networks within the schoolhouse (Wellman, 2001). In education, virtual interactions allow teachers to direct the trajectory of their curriculum and engage with their peers around education issues (Author, 2017; Supovitz, Daly, & del Fresno, 2015).

Pinterest represents a predominate virtual space for teacher interaction, knowledge, and information diffusion. In an NSF sample of teachers, we find 90% report using Pinterest for educational purposes. We conceptualize the resources teachers pin within social media as an artifact of their practice and, as an aggregate, a window into their conceptualization of instruction.

Purpose/Research Question: Though social media engagement continues to grow, little is known about how teacher's online resource access contributes to teachers' professional growth, impacts classroom teaching, and how federal and state level policies affect teacher's online resource acquisition (Author, 2011). We seek to understand how teachers respond to educational reforms within their professional practices and examine teachers' resource acquisition and sharing on Pinterest after the implementation of the Common Core State Standards (CCSS), explicitly define content standards and national expectations for mathematics and English Language Arts. We ask,

- How do teachers respond to federal curricular reform as evidenced by resource access and sharing within Pinterest?

Sample/ Data Sources: We consider state differences across Illinois and Indiana. Indiana, initially committed to the CCSS, pulled out after significant political pressure. Sampling 278 teachers across Indiana and Illinois for over 59 months (Jan 2012 - Nov 2016), we conduct analyses (189 Indiana and 89 Illinois teachers, respectively).

To understand changes in states impacted and not impacted by the CCSS reform, we examine impacts to educational resource access across Texas—in which CCSS was never a contended reform—and Indiana (189 and 100 teachers, respectively).

Data Analysis: This paper employs an interrupted (ITS) and comparative interrupted times series (CITS) design and piecewise regression model with Indiana, Illinois, and Texas sampled teachers to identify patterns in educational resource access over time. The CITS design allows us to hold constant the Pinterest platform developmental trajectory and identify CCSS impact on participant and non-participant states.

Model 1. Illinois CCSS implementation effect on teachers' Pinterest use

Teachers Pinterest use_t=

$\beta_0+\beta_1 CCSS_t+\beta_2 Month_t+\beta_3 Month_t \times CCSS_t+\beta_4 April_t+\beta_5 June_t+\beta_6 July_t$

Model B2. Indiana CCSS pause and withdrawal effect on teachers' Pinterest use

Teachers Pinterest use_t=

$\beta_0+\beta_1 CCSS\ pause_t+\beta_2 Month_t+\beta_3 Month_t \times CCSS\ pause_t+\beta_4 CCSS\ withdraw_t+\beta_5 Month_t \times CCSS\ withdraw_t+\beta_6 April_t+\beta_7 June_t+\beta_8 July_t$

In model B1, Teachers Pinterest use_t is the total volume of use at month t, $CCSS_t$ is a dummy variable which takes on a value of 0 before CCSS implementation, and 1 after CCSS implementation, $Month_t$ is a continuous variable over 59 months, $Month_t \times CCSS_t$ is the interaction between the CCSS implementation and month where month is centered at the time for CCSS implementation. β_0 is the baseline for teachers' Pinterest use in Jan 2012, β_1 is the change in the volume of teachers' Pinterest use immediately after CCSS implementation, β_2 is the monthly increase or decrease in total volume of teachers' Pinterest use before CCSS implementation, β_3 is the change on the monthly increase or decrease in the volume teachers' Pinterest use after CCSS implementation relative to before, β_4 through β_6 represent seasonal effects. In contrast, in model B2, Indiana experienced the CCSS pause after the initial enactment. We test the immediate CCSS pause effect (β_1 in model B2) and its gradual effect of the monthly change on the increase or decrease of Pinterest use after CCSS pause relative to before (β_3 in model B2). We test the immediate effect of official CCSS withdrawal (β_4 in model B2) along with its gradual effect of the monthly change on the increase or decrease of Pinterest use after the official CCSS withdraw relative to before (β_5 in model B2). Table B1 presents the CCSS enactment effects on teachers' Pinterest use volume in Illinois and the CCSS pause and withdraw effect on teachers' Pinterest use in Indiana.

Model 3. Indiana CCSS pause and withdrawal effect comparative to Texas on teachers' Pinterest use

Teachers Pinterest use_t=

$\beta_0+\beta_1 CCSS\ pause_t+\beta_2 Month_t+\beta_3 Month_t \times CCSS\ pause_t+\beta_4 CCSS\ withdraw_t+\beta_5 Month_t \times CCSS\ withdraw_t+\beta_6 Treated\ State_t+\beta_7 Treated_t \times Month_t+\beta_8 Treated_t \times CCSS\ pause_t+\beta_9 Treated_t \times CCSS\ pause_t \times Month_t +\beta_{10} Treated_t \times withdraw_t+ \beta_{11} Treated_t \times withdraw_t \times Month_t +\beta_{12} CountTeacher_t+\beta_{13} Feburary_t+\beta_{14} June_t+\beta_{15} July_t+\beta_{16} August_t$

In model B3, we use a CITS model and focus on testing the interaction between the treatment status, i.e. state's CCSS participation, and the immediate and gradual policy pause (β_8 and β_9) and withdrawal effect (β_{10} and β_{11}), comparing Indiana to Texas while holding constant other resource access differences beyond the CCSS participating status.

Table 1, 2, Figure 1, 2, here

Findings: Results indicated a significant negative effect of CCSS enactment on teachers' Pinterest use in Illinois sample with an immediate drop of 811 pins on teachers' total volume of Pinterest use. Over time, Illinois teachers' monthly Pinterest use decreased by 21 pins. Thus, after CCSS, teachers continued to use Pinterest to access educational resources, however, they accessed less overall. In contrast, we observed a sharp downward trend of teachers' monthly Pinterest use in Indiana, with a gradual decrease of 176 pins each month, after they announced a pause on CCSS but before the official policy withdrawal. After official withdrawal, the gradual decrease on teachers' Pinterest use became less steep relative to the pause period (β_5=70, *p*=.10). In the CITS, we find significant difference on teachers' resource access trend before the CCSS policy pause, a significance difference immediately after the policy pause, and a borderline significant difference on the gradual change after the CCSS policy withdraw between Indiana and Texas.

Conclusions: Teachers' response to educational policy uncertainty may be observed in significant drops in overall educational resource access. We conclude state educational policy context shapes how teachers experience education policy. Their experience may relate to responses in anticipatory and subsequent ways. Furthermore, as teachers engage as street level bureaucrats (Lipsky, 1980) in reform enactment and sense-making, their behavior may be observed through social media and relate back to their local and state context.

[Word count 998]

References

Author (2011).
Author (2018).
Lipsky, M. (1980). Street Level Bureaucracy: The Dilemmas of the Individual in Public Service. New York: Russell Sage Foundation.
Supovitz, J. A., Daly, A. J., & del Fresno, M. (2015). # CommonCore: how social media is changing the politics of education.
Wellman, B. "Computer Networks as Social Networks". *Science* 293.5537 (2001): 2031–2034. Web.

Table 1. Summary table of regression analysis for total volume of teachers' Pinterest use (Months=59)

Independent variables	Illinois		Indiana	
	B	SE B	B	SE B
Baseline	1098.86***	140.25	1400.51***	176.90
Month	93.29***	12.55	83.30***	19.88
CCSS	-811.14***	182.37		
Month*CCSS	-114.53***	13.33		
CCSS pause			-510.48^{+}	288.98
Month*CCSS pause			-175.73***	40.65
CCSS withdraw			451.35	275.28
Month*CCSS withdraw			69.43^{+}	36.38

Note. $^{+}p<.10$, $^{*}p<.05$, $^{**}p<.01$, $^{***}p<.001$

Table 2. Summary table of CITS regression analysis for total volume of IN teachers' Pinterest use (N=118)

Independent variables	Indiana	
	B	SE B
Baseline	2083.18***	464.50
Month	-150.71***	35.96
CCSS pause	1117.55*	469.62
Month*CCSS pause	-26.11	67.38
CCSS withdraw	840.40+	447.82
Month*CCSS withdraw	161.36**	59.64
Treated State	-2725.64***	585.61
Treated*Month	122.13*	50.31
Treated*CCSS pause	-1581.40*	665.53
Treated*Month*CCSS pause	19.69	96.63
Treated*CCSS withdraw	-915.38	621.70
Treated*Month*CCSS withdraw	-160.56+	82.48

Note. $^{+}p<.10$, $^{*}p<.05$, $^{**}p<.01$, $^{***}p<.001$. We hold constant the number of teachers in each State.

Figure 1. Plot of the trend on teachers' Pinterest use before and after CCSS implementation

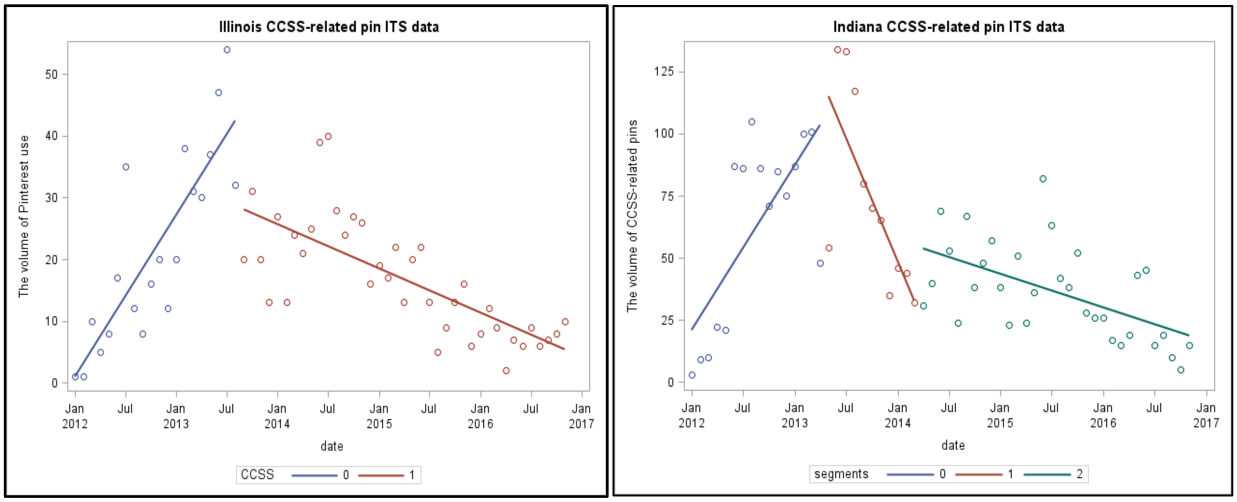
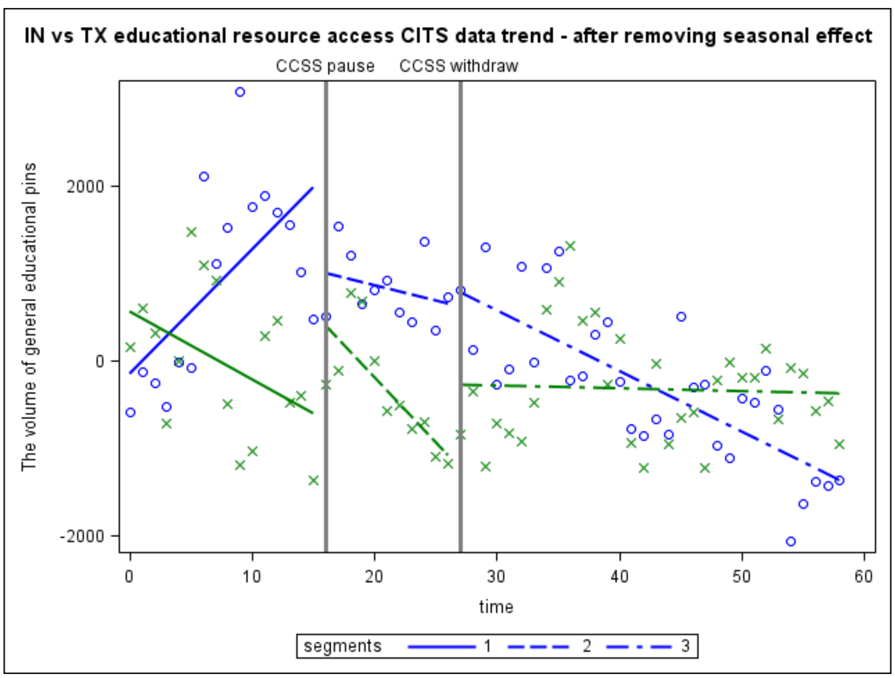


Figure 2. Plot of the trend on Indiana (blue) and Texas (green) teachers' Pinterest use before and after CCSS policy enactment



Metagenomic effects in human behavior: The case of adolescent smoking

Background/Context: A growing literature has suggested that the genomes of those around us affect our own phenotypes. Much of the empirical evidence for such “metagenomic” effects comes from animal studies, where the socio-genetic environment can be easily manipulated. Among humans, it is more difficult to identify the impact of others’ genomes on phenotypes given the non-random distribution of genes and environments.

Purpose/Objective/Research Question: In order to test for metagenomic effects in humans, we leverage the as-if-random distribution of grade-mates conditional on school-level variation in a nationally representative sample. Specifically, we evaluate whether the genetic propensity to smoke of one’s peers affects one’s own smoking behavior net of one’s own genotype.

Setting: The data for this study come from the National Longitudinal Study of Adolescent to Adult Health (Add Health), a nationally representative cohort study on the health and behavior of adolescent school children first interviewed in 1994-95 across a sample of 132 middle and high schools across the United States.

Population/Participants/Subjects: After the quality control measures, genotype data were available for 9,975 individuals, which were linked to their in-school and in-home survey results. Identifying classmates required supplementary information, which came from the *The Adolescent Health and Academic Achievement (AHAA)* study. It provides the high school transcripts for Add Health Wave III sample members (original Wave I sample members who were re-interviewed at Wave III). After collecting the transcripts, the AHAA constructed academic network using overlap in course-taking and ran clustering analyses on those networks. Individuals were assigned to clusters with other students with whom they took courses. We used these clusters to identify classmates, and again, for each student, averaged over the variables of their classmates to provide estimates of their peers’ mean smoking behavior, mean smoking genetic risk, and demographic composition.

Research Design: Individual smoking behavior measures how many cigarettes a day an individual smoked, on the days that they smoked. We include school and grade fixed effects to control for school and grade level. We further control for the sex, race, maternal education, familial smoking behavior, number of siblings and family income of every individual in the sample as well as the average level of those variables at the grade level. Further, following previous research on grade-level peers, for all analyses, we limit the sample to students in schools with a 12th grade (which results in excluding middle schools, but retaining high schools with 7th and 8th grades) and who were assigned sample weights. Then, for all grade-level analyses, we further exclude students who attended a grade with fewer than 20 total students in our sample. In all, 74 schools and 148 grades are represented in the data. This leaves us with 3,895 respondents who have genotype data and are in grades with a reasonable number of peers. For all friend analyses, we only include students who had at least one other friend in their grade, giving us 3,708 respondents. A number of controls were included in the analyses in order to ward against various forms of population stratification. In terms of race, we include individuals of all races in the analyses, but control for the individual's own race, PCs, the racial composition and average PCs of the grade, classroom, or friend group. Because it is possible that genes are a proxy for parent or sibling smoking behavior which might directly influence other individuals in a grade, we block this pathway by including an indicator for the presence of household smokers and older sibling both on the individual and on the grade level. In addition, we control for an individual's own smoking polygenic score in the models such that the results provide estimates of others' polygenic scores on individual smoking net of individual's own polygenic score.

Data Collection and Analysis: Our measure of genetic propensity to smoke is a polygenic score (PGS). A polygenic score is a genome-wide score that summarizes the presence of presence and associated weights of risk alleles discovered by genome-wide association studies (GWASs). We use the GWAS summary statistics from the Tobacco and Genetics Consortium on cigarettes per day to construct our PGS.

Findings/Results: Results show that peer genetic propensity to smoke has a large effect on an individual's smoking outcome, likely even larger than the individual's own genetic smoking risk. This is true for contexts where the individual plays an active role in shaping the metagenomic environment, such as the friend group, and also for contexts where they do not, as in the case of grade-mates. In secondary analyses, we explore these effects further and show that a few people with high genetic risk to smoke can greatly affect the smoking behavior of an entire grade. We find the causal effect of gradewide average genotypes on individual smoking behavior to be at least 60% as large as that of an individual's own genetic contributions. By using peer genotypes as an instrumental variable, we also confirm that the association between peer genotypes and individual outcomes operates almost entirely through peer behavior.

Conclusions: Our design affords new insight into the nature of peer influence in educational settings and its relationship to metagenomic effects. It also has substantial implications for future research beyond the case of smoking. That an individual's smoking behavior, and ultimately health outcomes, are affected by his or her peers' genes is important to consider for understanding social multiplier effects. Further, our framework relies only on the assumption of random variation in the local metagenomic environment and, as a result, can be used to estimate both metagenomic effects and peer effects for any behavior that has a genetic basis. This is especially valuable given the difficulty of estimating causal effects of peers, whether they result from their behaviors or genes. Future work may seek to apply this method to other socially driven outcomes, including health outcomes, that display an element of contagion and are in part influenced by genetic disposition — be that actual communicable disease as driven by the immunological profiles of peers around us, or educational performance and choices. [word count, 997]