English Learner reclassification: Impacts on student outcomes and potential mechanisms

Mark J. Chin
Harvard University

Background:

More than five million students in US public schools are classified as English Learners (ELs) (Musu-Gillette et al., 2017). Given ELs' unique instructional needs, educational agencies maintain policies outlining how to identify students as ELs, how to support them towards English proficiency, and when to reclassify them as fully English proficient. With the projected growth of ELs in classrooms, evidence on the impact of these policies is crucial. Results from reclassification studies in particular provide vital information to policymakers on what happens to students who lose instructional supports associated with being classified as an EL when they might still need them.

Extant studies leveraging regression discontinuity (RD) designs provide rigorous evidence on the impacts of reclassification policies. These studies, typically focused on cognitive outcomes, find mixed results (e.g., Carlson & Knowles, 2016; Pope, 2016; Robinson, 2011), and have generally not assessed impacts on students' noncognitive abilities or the potential mechanisms that could impact future outcomes. Yet noncognitive skills contribute significantly to students' long-term success (e.g., Heckman et al., 2006) and other studies argue that EL classification may have unintended consequences (e.g., influencing interactions with teachers and/or other students; Umansky, 2016, 2018).

Research Question:

To address these gaps in the literature, in this study I ask the following research questions:

1) What is the impact of being reclassified as fully English proficient on student cognitive outcomes (e.g., standardized tests)?
2) What is the impact of reclassification on student noncognitive outcomes?
3) What are the potential mechanisms for the impact of reclassification?

Setting, Data, and Population:

To answer these questions, I use administrative data from a large, southeastern school district (SSD). In SSD, EL students are reclassified if performance on the annual reclassification assessment indicates English proficiency. Because this policy stipulates that *only* performance on the assessment dictates reclassification status, I arrive at credibly unbiased impact estimates on student end-of-year (EOY) standardized test scores using an RD design. Furthermore, I am uniquely able to assess the impact on other outcomes using data from the SSD student survey. This includes noncognitive outcomes and measures of potential mechanisms for effects. In Table 1, I provide reliabilities and item text for these measures. In analyses, I rescale all outcomes to be z-scores to ease interpretation.

[Insert Table 1 here.]

I focus on data from SSD students ever classified as an EL in third grade, as the majority of reclassified EL students in the district are reclassified starting after third grade. SSD also surveys students in fifth and eighth grade to measure their perceptions of their learning experiences. By focusing on third grade ELs, I can thus identify the shorter- (fifth grade) and longer-term (eighth grade) impacts of reclassification on surveyed outcomes (in addition to EOY test scores).

Research Design:

By combining an RD approach with an instrumental variables approach (RD-IV), I can identify the causal impact of reclassification on outcomes, while also accounting for the fact that some ELs who are not reclassified immediately after third grade eventually lose their EL status over time. To implement this RD-IV, I first predict the number of years each student is reclassified as fully English proficient between third and fifth/eighth grade using: a linear term for reclassification exam performance (i.e., the "running variable"); an indicator for whether a student scores above reclassification thresholds in third grade (i.e., the "treatment" variable); the interaction of these two measures; and a vector of control variables (e.g., student gender, race, disability status) to improve the precision of estimates. I then use these predictions of reclassification status, which depends solely on (as good as random) variation caused by whether a student *just* passes reclassification thresholds in third grade, to predict fifth- and eighth-grade outcomes. Finally, to test for sensitivity to outliers, I estimate the RD-IV models across a range of bandwidths of reclassification assessment performance.

Results:

From Tables 2 and 3, where I present results from estimation of the RD-IV models, I observe several patterns.

[Insert Tables 2 and 3 here.]

First, reclassified ELs demonstrate substantial (.05- to .1-SD gains per year), significantly higher cognitive outcomes (Math and ELA EOY scores) than continuing ELs; this is true in both fifth and eighth grade. Second, the impact of reclassification on students' noncognitive outcomes is less clear, with inconsistent impacts across grades and bandwidths on students' goals and civic engagement, and null effects on students' grit. Finally, evidence for the impacts of reclassification on the three mechanism measures is slightly stronger. Across models, reclassified students generally report less challenging work from their teachers and worse relationships with their teachers and peers. This suggests that, unlike what prior work has hypothesized (Umansky, 2016, 2018), the classification label has *positive* unintended consequences for ELs with regards to how others in the school treat them.

Conclusions:

In my study, I find causal evidence that being reclassified as fully English proficient leads to significant gains in standardized test scores. Prior research using RDs argue that non-null

impacts of reclassification suggest that a better policy regime exists (Robinson, 2011); for example, SSD policymakers might reduce the stringency of thresholds to allow more ELs to be reclassified—at least, up until the point where reclassification no longer had a positive effect. Null effects, however, do not necessarily indicate that the current policy is the *best* policy (Robinson, 2011). Reclassifying students as fully English proficient may lead to unintended consequences (Umansky, 2018) overlooked by impact evaluations focused on cognitive outcomes. For example, I find evidence that reclassification negatively impacts students' reports on how others in school treat them. Given these negative impacts, loosening criteria for reclassification—as results from the impact evaluation on cognitive outcomes might suggest—may in fact be undesirable. SSD policymakers instead might identify how to better support students in their in-school relationships when transitioning out of EL services. More broadly speaking, these results encourage RD evaluations of policies to always consider a variety of outcomes.

References

Carlson, D., & Knowles, J. E. (2016). The effect of English Language Learner reclassification on student ACT scores, high school graduation, and postsecondary enrollment: Regression discontinuity evidence from Wisconsin. *Journal of Policy Analysis and Management, 35*(3), 559–586.

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics, 24*(3), 411–482.

Musu-Gillette, L., de Brey, C., McFarland, J., Hussar, W., Sonnenberg, W., & Wilkinson-Flicker, Sidney. (2017). *Status and trends in the education of racial and ethnic groups 2017.* NCES 2017-051. Washington, DC: National Center for Education Statistics.

Pope, N. G. (2016). The marginal effect of K-12 English language development programs: Evidence from Los Angeles Schools. Economics of Education Review, 53, 311–328.

Robinson, J. P. (2011). Evaluating criteria for English Learner reclassification: A causal-effects approach using a binding-score regression discontinuity design with instrumental variables. *Educational Evaluation and Policy Analysis, 33*(3), 267–292.

Umansky, I. M. (2016). To be or not to be EL: An examination of the impact of classifying students as English Learners. *Educational Evaluation and Policy Analysis, 38*(4), 714–737.

Umansky, I. M. (2018). According to plan? Examining the intended and unintended treatment effects of EL classification in early elementary and the transition to middle school. *Journal of Research on Educational Effectiveness, 11*(4), 588-621.

Table 1. Survey text and reliabilities of composite outcome variables.

Challenging Work measure
My teachers give me challenging work.

Teacher relationships composite, Cronbach's Alpha = .86
Overall, adults at my school treat students fairly.
Adults at my school listen to the students.
At my school, teachers care about students.
My teachers are there for me when I need them.
The school rules are fair.
Overall, my teachers are honest and open with me.
I enjoy talking to the teachers here.
I feel safe at school.
Most teachers at my school are interested in me as a person, not just as a student.

Peer relationships composite, Cronbach's Alpha = .84
Other students at school care about me.
Students at my school are there for me when I need them.
Other students here like me the way I am.
I enjoy talking to students here.
Students here respect what I have to say.
I have some friends at school.

Future Goals composite, Cronbach's Alpha = .83
I plan to continue my education following high school.
Going to school after high school is important
School is important for achieving my future goals.
My education will create many future opportunities for me.
I am hopeful about my future.

Civic Engagement composite, Cronbach's Alpha = .64
I believe I can make a difference in my community.
When I am old enough, I plan to vote in most elections.
I care a great deal about who is elected to be our next president.
I pay attention to what is going on in the news.
I think politics and government are boring (reversed).
I participate in projects in my community.
I often volunteer my time to help others.

Grit composite, Cronbach's Alpha = .58
New ideas and projects sometimes distract me from previous ones (reversed).
Setbacks (delays and obstacles) do not discourage me. I bounce back from disappointments faster than most people (reversed).
I have been obsessed with a certain idea or project for a short time but later lost interested (reversed).
I am a hard worker.
I often set a goal but later choose to pursue (follow) a different one.
I have difficulty maintaining (keeping) my focus on projects that take more than a few months to complete (reversed).
I finish whatever I begin.
I am diligent (hard working and careful).

Table 2. Effect of years of reclassification on fifth grade students' cognitive, noncognitive, and mechanism outcomes

Outcome variable	BW -6 to 5	BW -8 to 7	BW -7 to 6	BW -5 to 4	BW -4 to 3
Stacked Math/ELA EOY	.071*	.086**	.082**	.079~	.057
	(.034)	(.027)	(.030)	(.042)	(.037)
Noncog					
Future Goals	.084*	.073	.090*	.014	.026
	(.042)	(.045)	(.039)	(.045)	(.037)
Civic Engagement	.035	.021	.042	-.023	-.055
	(.042)	(.039)	(.036)	(.053)	(.065)
Grit	.024	.054	.037	-.054	.047
	(.054)	(.043)	(.041)	(.070)	(.070)
Mechanisms					
Challenging Work	-.044	-.062~	-.069~	-.053~	-.049
	(.030)	(.035)	(.037)	(.031)	(.037)
Teacher Relationships	-.041	-.038	-.039	-.093**	-.095***
	(.028)	(.025)	(.025)	(.028)	(.025)
Peer Relationships	-.100***	-.101**	-.090*	-.101***	-.096***
	(.026)	(.035)	(.038)	(.024)	(.022)
Stacked Mechanisms	-.062***	-.067**	-.066**	-.082***	-.080***
	(.015)	(.019)	(.021)	(.012)	(.021)

Note: Each column represents results using different bandwidths of reassignment assessment scores to test for sensitivity of results. Relationships between the running variable and the outcome are modeled linearly. For "stacked" models, each student appears in the data once for each outcome. In these models, I include a fixed effect for outcome type. Standard errors clustered at the reassignment assessment score level are displayed in parentheses. *~p<.1; *p<.05; **p<.01; **p<.001.*

Table 7. Effect of years of reclassification predicting eighth grade students' cognitive, noncognitive, and mechanism outcomes

Outcome variable	BW -6 to 5	BW -8 to 7	BW -7 to 6	BW -5 to 4	BW -4 to 3
Stacked Math/ELA EOY	.052***	.090***	.071***	.030*	.024**
	(.0149)	(.024)	(.019)	(.012)	(.008)
Noncog					
Future Goals	-.076*	-.091**	-.078**	-.062~	-.014
	(.032)	(.031)	(.030)	(.033)	(.040)
Civic Engagement	-.065~	-.069~	-.076~	-.089*	-.050
	(.039)	(.040)	(.042)	(.044)	(.037)
Grit	.028	.015	-.004	-.001	.026
	(.056)	(.054)	(.062)	(.072)	(.063)
Mechanisms					
Challenging Work	-.081*	-.079*	-.105**	-.040	-.007
	(.035)	(.035)	(.036)	(.029)	(.041)
Teacher Relationships	-.068*	-.072*	-.063~	-.076*	-.010
	(.034)	(.029)	(.033)	(.039)	(.040)
Peer Relationships	-.011	-.005	.008	-.042	-.052~
	(.045)	(.039)	(.040)	(.041)	(.028)
Stacked Mechanisms	-.053**	-.052**	-.053**	-.053**	-.023
	(.019)	(.017)	(.017)	(.019)	(.015)

Note: Each column represents results using different bandwidths of reassignment assessment scores to test for sensitivity of results. Relationships between the running variable and the outcome are modeled linearly. For "stacked" models, each student appears in the data once for each outcome. In these models, I include a fixed effect for outcome type. Standard errors clustered at the reassignment assessment score level are displayed in parentheses. *~p<.1; *p<.05; **p<.01; **p<.001.*