

Study of the Impact of English Learner Classification and Reclassification Policies: District and School Surveys

Supporting Statement for Paperwork Reduction Act Submission

PART B: Collection of Information Employing Statistical Methods

November 2024

Contract 91990021D0004 (Task Order 2: 91990022F0057)

Submitted to:

Institute of Education Sciences
U.S. Department of Education

Submitted by:

Westat
An Employee-Owned Research Corporation®
1600 Research Boulevard
Rockville, Maryland 20850-3129
(301) 251-1500

Contents

B.1. Respondent Universe and Sample Design	B-1
B.2. Information Collection Procedures	B-2
B.2.1. Statistical Methodology for Stratification and Sample Selection	B-2
B.2.2. Estimation Procedures	B-3
B.2.3. Degree of Accuracy Needed	B-4
B.2.4. Unusual Problems Requiring Specialized Sampling Procedures	B-7
B.2.5. Use of Periodic (Less than Annual) Data Collection to Reduce Burden	B-7
B.3. Methods to Maximize Response Rates and Address Nonresponse	B-7
B.3.1. Methods to Maximize Response Rates	B-7
B.3.2. Methods to Address Nonresponse	B-9
B.4. Tests of Procedures	B-9
B.5. Individuals and Organizations Involved in the Project	B-9
References	B-11
Appendix A. District Survey	A-1
Appendix B. School Survey	B-1
Appendix C. Notification, Survey Invitation, and Reminder Materials	C-1

Exhibit

Exhibit B-1. Respondent universes and sample sizes for the district and school surveys, by state	2
Exhibit B-2. Key research questions addressed using survey data and associated estimation procedures	3
Exhibit B-3. Estimated precision for descriptive analyses	5
Exhibit B-4. Estimated precision for analyses of district factors moderating the impacts of reclassification on ELA achievement in grades 6 to 8	7
Exhibit B-5. Organizations and individuals involved in the project	10

Part B. Collection of Information Employing Statistical Methods

The U.S. Department of Education, through its Institute of Education Sciences (IES), is requesting clearance for data collection activities to support a study to evaluate and inform entrance and exit policies for English learners (ELs). Classification into and reclassification out of EL status are both high-stakes decisions with far-reaching impacts for students, educators, and education systems. To help achieve better outcomes for ELs, in 2015, the reauthorization of Title III of the Elementary and Secondary Education Act as the Every Student Succeeds Act (ESSA) required states to implement statewide standardized EL entry and exit procedures, starting in the 2017–18 school year.

This package requests clearance for district and school survey instruments and administration of these surveys. This data collection will complement an earlier data collection request (OMB Approval # [1850-0974](#)) to collect extant data on students from state longitudinal data systems (SLDSs). The surveys of districts and schools will allow an assessment of whether the standardization of procedures within states is happening as intended by ESSA and whether locally determined instructional settings, programs, and services moderate the impacts of reclassification on ELs.

B.1. Respondent Universe and Sample Design

Following plans described in this study's previously approved clearance request (OMB Approval [1850-0974](#)), the study team is collecting SLDS records from 30 of the states with the largest populations of ELs. These states collectively contained 98 percent of students who exited EL status nationally in the 2017–18 and 2018–19 school years.

The district survey will collect information about instructional settings, programs, and services applicable to the full range of grades (K to 12); hence the target population for the district survey is all ELs in the 30 study states. The school survey will collect information about a subset of these topics in middle schools, as discussed in Part A. The target population for the school survey is ELs who are in grades 6 to 8 in the study states. However, because not all schools include all three grades, we define the respondent universe of middle schools as schools offering grades 7 and 8.¹

The study team will field these surveys to the following samples selected in the 30 study states:

- The **district survey sample** will include 1,800 districts selected from a universe of 8,462 districts that enroll ELs in the 30 study states. The study team will survey one superintendent or a designee in each sampled district. The study team expects a district survey response rate of 90 percent.
- The **school survey sample** will include 1,800 middle schools nested within the sampled districts from a universe of 15,882 schools that enroll ELs in the 30 study states. The study team will survey one principal or a designee in each sampled school. The study team expects a school survey response rate of 85 percent.

Section B.2 contains additional information on how the study team will select districts and schools for the survey samples, along with a table showing respondent universes and sample sizes by state.

¹ In the 2017-18 school year, among schools offering any grade from 6 to 8, almost 63 percent offer all three grades, 15 percent start at grade 7, and 19 percent end at grade 6.

B.2. Information Collection Procedures

B.2.1. Statistical Methodology for Sample Selection

The study team’s sampling plans prioritize precision for key moderator analyses of reclassification impacts on ELs using data from surveys. As described in Section B.2.2, the descriptive analysis of survey responses will provide district and school-level prevalence estimates of certain policies, while the moderator analyses will produce *student-level estimates* by analyzing student SLDS records for all ELs in groups of districts/schools defined using survey response data and comparing estimates between survey response groups through a fixed-effects meta-analysis.

The study team will select the 1,800 largest districts in the 30 study states regardless of state. We will then select 1,800 large middle schools. However, we will not necessarily take the largest middle schools within the study states, for two reasons. First, we will constrain the selection of the 1,800 schools to be located within the 1,800 sampled districts. It would be inefficient to include schools in the sample that are not nested within the districts we select due to the research application process required by many districts. Second, we will set an upper bound on the number of schools from any one district. At the research application stage, districts can decide not to allow the study team to survey its schools. Therefore, to decrease the risk of losing many schools within the same district, we will not select more than 10 middle schools from any one district. Based on this sampling approach, Exhibit B-1 displays state-by-state counts of the number of districts and schools in the respondent universe, the number of ELs enrolled in those districts and schools (counting only ELs in grades 6 to 8 for schools), and the number of districts and schools in the sample.

Exhibit B-1. Respondent universes and sample sizes for the district and school surveys, by state

State	Number of ELs in district universe	District universe size	District sample size	Number of ELs in grades 6 to 8 in school universe	School universe size	School sample size
Arizona	76,180	186	51	15,408	563	17
Arkansas	36,465	210	21	6,646	199	8
California	988,852	877	421	200,270	1,969	483
Colorado	74,257	152	29	12,897	384	24
Connecticut	48,907	169	30	9,165	263	21
Florida	231,590	68	40	38,822	696	72
Georgia	127,643	169	47	24,282	467	39
Illinois	228,403	648	126	48,737	1,028	88
Indiana	66,566	278	42	13,782	376	24
Maryland	96,261	24	17	16,624	277	30
Massachusetts	87,659	273	47	13,793	387	28
Michigan	73,238	393	46	13,657	539	28
Minnesota	57,770	250	41	9,094	308	20
Missouri	30,090	225	26	4,779	292	2
Nevada	60,541	16	6	11,718	117	16
New Jersey	108,749	464	67	16,605	600	28
New Mexico	54,385	77	23	13,505	152	34

Exhibit B-1. Respondent universes and sample sizes for the district and school surveys, by state—continued

State	Number of ELs in district universe	District universe size	District sample size	Number of ELs in grades 6 to 8 in school universe	School universe size	School sample size
New York	219,967	561	64	39,016	1,099	36
North Carolina	114,738	118	62	27,146	558	52
Ohio	56,776	460	36	7,366	598	6
Oklahoma	57,647	353	26	12,033	365	20
Oregon	52,175	136	35	9,935	264	19
Pennsylvania	68,504	439	36	13,754	594	22
South Carolina	43,956	77	34	9,219	266	15
Tennessee	51,899	133	21	7,426	413	15
Texas	963,549	967	251	226,714	1,735	548
Utah	49,837	37	20	12,777	164	22
Virginia	116,862	131	34	18,849	341	37
Washington	118,732	226	74	15,959	456	33
Wisconsin	44,775	345	27	9,066	412	13
Total	4,406,973	8,462	1,800	879,044	15,882	1,800

EL = English learner

Note: All counts of ELs are for the 2021-22 school year. The respondent universe includes only districts and schools that enroll ELs. The school respondent universe also excludes schools that do not include grades 7 and 8; special education, charter, alternative, and fully virtual schools; and other schools not categorized as a “regular school” in the Common Core of Data. Source: U.S. Department of Education, National Center for Education Statistics, ED Data Express, 2021-22, Common Core of Data 2021-22, Civil Rights Data Collection 2020-21.

B.2.2 Estimation Procedures

The survey data collection activities described in this clearance request will allow the study team to characterize local instructional settings, programs, and services in the largest districts that moderate the effects of reclassification among ELs, as well as describe these local factors to provide context for policymakers. Exhibit B-2 shows the specific research questions that the study team will answer using the survey data. (Section A.2 of Part A lists the full set of study research questions, including some that do not require survey data.) The exhibit also indicates the types of estimation procedures the study team will use for each research question, and the following subsections contain more details about these procedures.

Exhibit B-2. Key research questions addressed using survey data and associated estimation procedures

Research question	Estimation procedures
What instructional settings, programs, and services do districts and schools offer to students?	Descriptive analyses
What is the relationship between these instructional settings, programs, and services and the impacts of reclassification on student outcomes?	Fixed-effects moderated meta-analyses

Descriptive analyses. The study team will calculate summary statistics, such as means and percentages, to describe the instructional settings, programs, and services offered by districts and schools, based on responses to the surveys. When estimating summary statistics and their variances, the study team will use normalized frequency weights to account for the size of each district’s or school’s target population to allow inference about the population of students included in the districts or schools responding to the survey. To compare groups of districts or schools, the study team will use chi-square tests to test for significant differences in proportions for categorical variables and *t*-tests to test for significant differences in means for continuous variables.

Impact moderator analyses. The study team will assess moderators of reclassification by contrasting impacts across two groups of districts, defined using survey response data: (1) districts using one type of instructional setting, program, or service; and (2) districts using an alternative instructional setting, program, or service (referred to below as a policy). Within each subgroup of surveyed districts, the study team will estimate student-level impacts of reclassification using the SLDS records already collected, as indicated in this study’s previous clearance request (OMB Approval [1850-0974](#)). These impact estimates will be based on a regression discontinuity design (RDD), which can be used to compare the outcomes of otherwise similar ELs whose English language proficiency test scores are just above and below the threshold for reclassification.

The estimate for the average RDD impact of students in districts which follow the *i*th policy is a precision weighted average of the impacts of the *i*th policy-associated RD impact across the set of *j*th states

$$\hat{\theta}_{i.} = \frac{\sum_j w_{ij} \hat{\theta}_{ij}}{\sum_j w_{ij}}, \quad i = \{1, 2\},$$

where the weights are the inverse of the RD impact estimate’s variance, $w_{ij} = (V(\hat{\theta}_{ij}))^{-1}$.

The fixed-effects meta-regression

$$\hat{\theta}_{ij} = \gamma_0 + \gamma_1 I(i=1) + \epsilon_{ij},$$

will be used to test whether $\hat{\theta}_{1.} = \hat{\theta}_{2.}$ by testing whether $\gamma_1 = 0$, where $I(i=1)$ is an indicator variable equal to 1 if $\hat{\theta}_{ij}$ is the RD impact from the state’s sample of students under Policy 1. The variance of γ_1 is the sum of the variance of the policy-specific estimates, $V(\hat{\theta}_{1.})$ and $V(\hat{\theta}_{2.})$, which in turn relate to the associated estimated variances of the average RD impacts across each *j*th state for each *i*th policy, $\hat{\theta}_{ij}$. The variance for each *i*th policy-associated mean impact across states is the inverse of the sum of the precision weights used (Hedges and Piggot 2004).

B.2.3. Degree of Accuracy Needed

The study team expects the survey sample sizes to yield sufficient precision for both descriptive analysis and key moderator analyses.

Precision for descriptive analyses. For binary measures of district practices based on the survey, the study team seeks to produce descriptive statistics with a margin of error of no more than ± 5 percentage points. The study team also seeks to reliably detect differences of at least 10 percentage points when comparing practices between policy-relevant subgroups of districts/schools that differ in key features of the state context. The descriptive analysis uses these thresholds for precision because smaller errors or differences may not meaningfully alter conclusions about the population. For example, study results might lead to the conclusion that approximately half of ELs have access to additional supports or accommodations during the monitoring period in both districts/schools where 45 percent of ELs can access these supports and districts/schools where 55 percent of ELs can access these supports.

The study team expects the district and school survey samples to yield estimates that meet or exceed these precision targets, as shown in Exhibit B-3. The study team assessed the likely precision of estimates for a practice used by 50 percent of survey respondents, which corresponds to the highest potential variance and aligns with the prevalence of ability tracking in middle schools, a potential key moderator for impacts of reclassification (Standing & Lewis, 2021). The study team also considered precision for less-/more-prevalent practice used by 25 or 75 percent of survey respondents. For each prevalence level of an outcome, all margins of error, as measured by the half-width of 95 percent confidence intervals, are less than 5 percentage points for both the district and school samples—overall and for policy-relevant subgroups. Additionally, Exhibit B-3 shows that the minimum detectible difference (MDD) between subgroup means is no more than 8 percentage points for comparisons of both districts and schools.

Exhibit B-3. Estimated precision for descriptive analyses

Survey measure/sample	Expected number of responses	Standard error of mean	Half-width of 95 percent confidence interval	MDD for contrasts subgroup means
Instructional setting, program, or service with a prevalence of 50 percent				
District respondents				
Full sample	1,620	1.24 pp	2.43 pp	n.a.
50 percent subgroup	810	1.76 pp	3.44 pp	6.94 pp
33 percent subgroup	535	2.16 pp	4.24 pp	7.35 pp
School respondents				
Full sample	1,530	1.28 pp	2.51 pp	n.a.
50 percent subgroup	765	1.81 pp	3.54 pp	7.14 pp
33 percent subgroup	505	2.22 pp	4.36 pp	7.56 pp
Instructional setting, program, or service with a prevalence of 25 percent or 75 percent				
District respondents				
Full sample	1,620	1.08 pp	2.11 pp	n.a.
50 percent subgroup	810	1.52 pp	2.98 pp	6.25 pp
33 percent subgroup	535	1.87 pp	3.67 pp	6.68 pp
School respondents				
Full sample	1,530	1.11 pp	2.17 pp	n.a.
50 percent subgroup	765	1.57 pp	3.07 pp	6.44 pp
33 percent subgroup	505	1.93 pp	3.78 pp	6.88 pp

MDD = minimum detectable difference; pp = percentage point; n.a. = not applicable.

Note: The reported standard errors, confidence intervals, and MDDs all assume that staff in 90 percent of sampled districts and 85 percent of sampled schools will respond to surveys. The reported MDDs for each subgroup are based on comparisons of means with the complementary subgroup—that is, comparisons between two subgroups that each comprise 50 percent of the full sample and comparison between subgroups comprising 33 percent and 67 percent of the full sample. For MDDs, the study team also assumed a target level of statistical power of 80 percent, and that comparisons of means will use *t*-tests and a 5 percent level of statistical significance for testing.

Precision for key moderator analyses. The variance of each standardized mean difference impact in a fixed-effects meta-regression is a function of the sample sizes associated with each impact, which allows us to estimate the precision expectations for our moderator analysis based on an effective sample size. We compute an effective sample size using the expected sample size, an expected response rate, and RDD design effects (Schochet, 2009), which we empirically estimated using available SLDS data. (See Deke and Dragoset [2012] for a similar approach to estimating RDD design effects.) We convert the expected total sample to an effective sample size with $N_{eff} = N \times \frac{RR}{DEFF}$, where RR is the expected response rate and $DEFF$ is the RDD design effect. Using our effective sample sizes, we compute the MDD of the fixed meta-regression results using typical two-group study formulas (Hedberg, 2017), assuming 50 percent of the effective RDD sample was on either side of the cutoff in the RDD estimates, and that p_1 of the districts or schools followed the first of two policies,

$$MDD = 2.8 \times \sqrt{\frac{4}{p_1 \times (1 - p_1) \times N_{eff}}}$$

where 2.8 is a factor which is derived from quantiles of the standard normal distribution associated with a Type I error rate of .05 split between 2 tails and Type II error rate of 0.2 (power is 0.8).

When comparing the impacts of reclassification between policy-relevant groups of districts, the study team seeks to reliably detect differences of at least 0.10 standard deviations. IES evaluations often specify a minimum detectable effect size of 0.10 standard deviations because smaller effect sizes may not be educationally meaningful and require cost-prohibitively large samples. This study’s moderator analysis also uses 0.10 standard deviations as a threshold for educationally meaningful gains related to districts’ use or adoption of an instructional setting, program, or service.

The study team expects the moderator analysis estimates to meet these precision targets for the school sample of ELs in key grade spans, based on the MDDs reported in Exhibit B-4. These MDDs focus on moderators of the impacts of reclassification on English language arts (ELA) achievement among ELs in typical middle school grades 6 to 8 across three school years (see B.1 and part A for discussion of the middle school focus). We expect an 85 percent response rate from schools. The exhibit presents MDDs for hypothetical contrasts of districts grouped based on moderators that differ in prevalence. Assuming a moderator with a prevalence of 50 percent, the MDD is 0.078 standard deviations—below the study’s precision target. For a moderator with a prevalence of 25 percent or 75 percent, the MDD is 0.090 standard deviations—which also meets the study’s precision target. The district sample is expected to have a higher response rate, 90 percent, and thus has lower MDDs which meet the targets as well.

Exhibit B-4. Estimated precision for analyses of district factors moderating the impacts of reclassification on ELA achievement in grades 6 to 8

Sample	Students sampled	Effective sample	MDD values by proportion of districts in first group		
			50 percent	33 percent	25 percent
School	1,257,531	20,774	0.078	0.083	0.090
District	2,455,092	47,980	0.051	0.054	0.059

MDD = minimum detectable difference in standard deviation units.

Note: The number of students sampled equals the total number of English learners in grades 6 to 8 in 1,800 sampled schools over three years (first row) and English learners in grades 6 to 8 in 1,800 sampled districts over three years (second row). Each MDD assumes that the study team will estimate impacts using student-level data from 1,530 schools (first row) and 1,620 districts (second row) responding to this study’s survey. The MDD was further adjusted to account for RDD design effects. Since reclassification may be affected by test scores and other factors, the study team will use fuzzy RDD methods to measure the causal effects of exiting EL status among students with test scores near the reclassification threshold (Lee & Lemieux, 2010; Calonico et al., 2019). The reported MDDs are based on these sample sizes, along with assumptions about how districts are divided into subgroups based on the prevalence of the moderator (per the columns of the table), and estimates of the variance of RDD impacts in each moderator subgroup. The study team estimated design effects of RDD impacts relative to RCT impacts based on preliminary results from analyses of SLDS records collected for this study. The design effects were used to adjust the students sampled in each state to compute an effective sample size to estimate the MDDs using two-group comparison formulas. All MDDs assume that (1) the target level of statistical power is 80 percent; and (2) comparisons of impacts will use two-tailed statistical tests with a 5 percent threshold level of significance.

B.2.4. Unusual Problems Requiring Specialized Sampling Procedures

There are no unusual problems requiring specialized sampling procedures.

B.2.5. Use of Periodic (Less than Annual) Data Collection to Reduce Burden

Both the district and school surveys will be fielded only once starting in January 2025.

B.3. Methods to Maximize Response Rates and Address Nonresponse

B.3.1. Methods to Maximize Response Rates

To maximize response rates, the study team will work closely with districts and schools in the survey sample using strategies that have been successful for other large-scale surveys (e.g., surveys conducted for the *Implementation of Title I/II-A Program Initiatives* study (OMB # 1850-0967) and the *Title II, Part A Use of Funds Study and Analytic Support* (OMB # 1810-0618). The study team’s general approach is to clearly communicate with potential respondents throughout the process of fielding surveys to set expectations, build relationships, and encourage follow-up.

At the start of the survey, the study team will work with school districts and schools to explain the importance of the data collection efforts and make it as easy as possible to comply, by:

- Sending notification letters to the superintendents of sampled districts and survey invitation letters to superintendents and to the principals of sampled schools about the surveys (Appendix C). These letters will include clear descriptions of the study’s purpose, design, and importance; a summary of survey content; contact information for the study team; and OMB clearance

information. The district letter will note that districts are expected to participate, per ED regulations, and the school letter will note that participation is voluntary. The invitation letters, on U.S. Department of Education letterhead and signed by the federal project officer for the study, will also include a link to the survey website and log-in credentials that districts and schools can use to access the survey. The study team will send letters both by email and postal mail to increase the likelihood that addressees receive the letters in a timely manner.

- Identifying a primary contact at the district and school. The district superintendent, EL/Title III director, or a designee will serve as the primary contact for the district survey; and the school principal or a designee will serve as the primary contact for the school survey. The study team will follow all required procedures needed to obtain approval from the district for its participation in the study, and for participation of the sampled schools.
- Answering questions from district and school staff using efficient and responsive procedures. Potential respondents may contact the study team through a toll-free hotline and a study email address included in the invitation letter. The study team will assign trained staff to answer the study hotline and reply to emails in the study mailbox. These staff will be trained to readily answer questions about the purpose and logistics of the survey. They also will be able to quickly provide consistent information to district and school staff by referring to a document with frequently asked questions, and the study leadership team will be available to answer new questions that arise.

The study team will accept completed surveys in multiple formats. Although the primary mode for completing the survey will be the web, the study team will make available an electronic version of the survey (e.g., PDF or Word document) that respondents can return by email or postal mail if they prefer.

The study team will track nonresponse and be courteous but persistent in following up with participants who do not respond in a timely manner. Specifically, the study team will:

- Follow up with nonrespondents by email and telephone. About one week after the start of data collection, the study team will begin contacting districts by telephone that have not logged into their survey to confirm they received the survey invitation letter and answer any questions. Emails will be sent to schools one week after the invitation letter is mailed and the study team will begin following up with schools by telephone one week later. The survey management system will allow interviewers to send personalized email messages to respondents, answering their questions and providing them with their survey login information if needed.
- Monitor responses, review submitted survey instruments for completeness, and continue to respond to questions received via the toll-free hotline or the study email account.

The study team is also proposing the use of survey incentives for principals. As noted in Part A, because principals face numerous data collection requests, incentives may be needed to obtain high response rates so that the school survey data can yield reliable answers to the study's research questions. If approved by OMB, the proposed incentive of \$50 will be paid upon survey completion.

B.3.2. Methods to Address Nonresponse

If unit or item-level response rates for either survey are below 85 percent, the study team will conduct analyses to address and mitigate potential challenges in generalizing from a survey sample to the universe of district or school respondents.

The study team will analyze response rates, overall and by stratum, to check for differences across subgroups of districts/schools defined by (1) size, based on the number of ELs (see Section B.2); and (2) average EL characteristics (e.g., race/ethnicity, disability status, eligibility for free or reduced-priced lunch). The study team will create these subgroups using SLDS records for all sampled districts. Large or significant differences in response rates may indicate potential nonresponse bias, in which case respondents may not be directly representative of the corresponding universe of districts/schools.

Should this analysis indicate large standardized differences, the study team will construct nonresponse weights to limit potential bias and increase alignment between the sample and the universe (Brick, 2013). The study team will also conduct sensitivity tests in the main descriptive and RDD analysis to determine if the findings change when excluding schools/districts in particular size categories or with specific characteristics associated with low response rates. The study team will not, however, impute survey data, since a major goal for collecting these data is to contrast subgroups defined by survey responses and the inherently imperfect imputation of responses would bias such contrasts.

B.4. Tests of Procedures

The study team pretested each survey with nine or fewer respondents to ensure that questions are clear and that the average survey completion time is within expectations. The team conducted these pretests via telephone calls or videoconferences with superintendents (or Title III directors or designees) and principals in districts and schools with varying sizes and average characteristics of ELs. The pretests used a cognitive interviewing format, in which respondents were asked for feedback on the format, content, and wording of the survey instrument. The study team used probing questions to identify instructions or questions that respondents may not fully understand. Based on what was learned in the pretests, the study team made improvements to the surveys to make sure that respondents can provide accurate and reliable responses and to limit undue burden. The study team also used timing information from the pretests to confirm burden estimates.

B.5. Individuals and Organizations Involved in the Project

Westat is the prime contractor for the evaluation, supported by its subcontractors WestEd and New York University. Exhibit B-5 shows the leadership team as well as at least one key contact from each participating organization.

Exhibit B-5. Organizations and individuals involved in the project

Name	Role	Organization	Phone number
Eric Isenberg	Project Director	Westat	240-314-7542
Molly Faulkner-Bond	Co-Principal Investigator	WestEd	202-816-3508
Joseph Cimpian	Co-Principal Investigator	New York University	212-998-5049
Ann Webber	Deputy Project Director	Westat	301-738-3627
Eric Hedberg	Data Analysis Task Leader	Westat	301-251-8253
Michael Steketee	Data Acquisition Task Leader	Westat	240-453-2603
Atsushi Miyaoka	Data Processing Task Leader	Westat	301-610-4948

In addition, we consult with a technical working group (TWG) of researchers and practitioners to provide input on the sampling plan, data collection instruments, and eventually the interpretation of results. The TWG consists of researchers with expertise in issues related to the acquisition of language proficiency; state and local policies for entry and exit of ELs; curricula and strategies to support ELs; and RDD. TWG members include:

- Rebecca Callahan, Professor, College of Education and Social Services, The University of Vermont
- John Deke, Senior Fellow, Mathematica
- Anjelica Infante-Green, Commissioner, Rhode Island Department of Education
- Madeline Mavrogordato, Associate Professor, College of Education, Michigan State University
- Sean Reardon, Professor, Stanford Graduate School of Education
- Nami Shin, Senior Research Associate, University of Kansas
- Emily Tanner-Smith, Thomson Professor, College of Education, University of Oregon

References

- Brick, J. M. (2013). Unit nonresponse and weighting adjustments: A critical review. *Journal of Official Statistics*, 29(3), 329.
- Calonico, S., Cattaneo, M. D., Farrell, M. H., & Titiunik, R. (2019). Regression discontinuity designs using covariates. *Review of Economics and Statistics*, 101(3), 442-451.
- Deke, J., and Dragoset, L. (2012). Statistical Power for Regression Discontinuity Designs in Education: Empirical Estimates of Design Effects Relative to Randomized Controlled Trials. Working Paper. Mathematica Policy Research, Inc.
- Hedberg, E. C. (2017). *Introduction to Power Analysis: Two-Group Studies*. SAGE Publications.
- Hedges, L. V., and Pigott, T. D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods*, 9(4), 426-445.
- Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of economic literature*, 48(2), 281-355.
- Schochet, P. Z. (2009). Statistical power for regression discontinuity designs in education evaluations. *Journal of Educational and Behavioral Statistics*, 34(2), 238-266.
- Standing, K., and Lewis, L. (2021). *Pre-COVID ability grouping in U.S. public school classrooms*. Data Point (NCES 2021-139, for U.S. Department of Education, National Center for Education Statistics). Washington, DC.
- U.S. Department of Education. (2022). *Digest of education statistics*. Washington, D.C.: ED, National Center for Education Statistics.