

# NATIONAL EVALUATION OF SCHOOL TURNAROUND AMERICORPS

---

## SUPPORTING STATEMENT FOR PAPERWORK REDUCTION ACT SUBMISSIONS

### B. COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS

**B1. Describe (including numerical estimate) the potential respondent universe and any sampling or other respondent selection method to be used. Data on the number of entities (e.g., establishments, State and local government units, households, or persons) in the universe covered by the collection and in the corresponding sample are to be provided in tabular form for the universe as a whole and for each of the strata in the proposed sample. Indicate expected response rates for the collection as a whole. If the collection had been conducted previously, include the actual response rate achieved during the last collection.**

#### *Study Sample*

This section describes how the sample of schools participating in the evaluation will be determined. The School Turnaround AmeriCorps National Evaluation school sample will include low-performing schools that receive the School Turnaround AmeriCorps intervention (hereafter, treatment schools) and other low-performing schools that lack a meaningful AmeriCorps presence (hereafter, comparison schools). These two sets of schools will be compared through a variety of quantitative and qualitative analyses to understand how the AmeriCorps presence affects treatment schools' turnaround efforts. To make such comparisons meaningful, each comparison school will be selected so that its profile of key school characteristics (e.g., reading and math proficiency, percent receiving free or reduced price lunch) closely matches the profile of a treatment school. We first describe the treatment school subsample, then the pool of potential comparison schools from which the comparison school subsample will be drawn, and conclude with a discussion of the matching procedure that will be followed to select the comparison schools.

#### *Treatment Schools*

The treatment school subsample will consist of a subset of the schools that started receiving their School Turnaround AmeriCorps intervention in the 2013-2014 school year. This subsample will include 5 of the 17 Teach for America (TFA) schools and all 57 non-TFA schools. The 5 TFA schools that will be included in the subsample will be chosen purposively to minimize the number of school districts to be recruited and, within this constraint, to represent a range of school levels and geographic locations. The subsample of schools excludes the new School Turnaround AmeriCorps schools from the 2014-2015 cohort, as it is likely we expect that the experiences of schools just starting the intervention will be qualitatively different from those of schools that have already been implementing the intervention for a year.

#### *Potential Comparison Schools*

We have constructed a pool of potential comparison schools, from which the comparison school subsample will be drawn. To qualify as a potential comparison school, a school must meet the following criteria:

1. It is designated a SIG or Priority school.
2. It is not receiving any form of the School Turnaround AmeriCorps intervention, and has an “at most minimal” AmeriCorps member and/or VISTA volunteer presence. The operationalization of “at most minimal” is TBD pending additional information on AmeriCorps performance measures and service activities.
3. There is a treatment school that resembles the potential comparison school, in the following ways:
  - The pair of schools is from the same state.
  - The potential comparison school offers the treatment school’s relevant grades. The School Turnaround AmeriCorps intervention is sometimes focused on a subset of the grades offered by a treatment school, and the set of grades receiving AmeriCorps assistance is referred to as the treatment school’s relevant grades. For instance, if a K-12 treatment school receives AmeriCorps assistance in grades 3-5, then its relevant grades are 3-5, and another K-5 school from the same state could qualify as a potential comparison school.
4. It does not use the closure model. Among the treatment and potential comparison schools, there are a handful of schools with the closure model, a small number with the restart model, including several treatment schools, and the vast majority of schools are transformation or turnaround schools. The closure model, which involves closing an existing school and enrolling students who attended that school in other, higher-achieving schools,<sup>1</sup> is not applicable to the School Turnaround AmeriCorps intervention, which places members in low-performing schools to support their turnaround efforts. Therefore any potential comparison schools using the closure model will be dropped from the pool of potential comparison schools.

Meeting the first, third and fourth criteria means that treatment school “apples” will not be compared to comparison school “oranges,” as pair-able schools are both low-performing, serve the relevant grades, and share the same state-level educational environment. Meeting the second criterion means that a potential comparison school is not receiving assistance comparable in nature and intensity to the School Turnaround AmeriCorps intervention.

Note that it is possible that a given potential comparison school could theoretically meet the third criterion with respect to multiple schools in the treatment school subsample.

### ***Matching Procedure***

Ideally, the matching procedure will pair each school in the treatment school subsample with a single very similar school from the pool of potential comparison schools so that no comparison school is paired with multiple treatment schools. That is, the matching procedure seeks to produce 1-1 matching without replacement. Matching a treatment school with more than one comparison school is expensive given the resources needed to recruit districts and schools and administer surveys; matching a comparison school to more than one treatment school reduces statistical power.

The matching procedure will also provide the means for selecting replacement comparison schools if, once selected, a comparison school is unwilling to participate in the evaluation.

The matching procedure includes the following steps:

---

<sup>1</sup> Handbook on Effective Implementation of School Improvement Grants. Chapter 4: Organizational Structures Only. [http://www.centerii.org/handbook/Resources/Chapter\\_4\\_Organizational\\_Structures.pdf](http://www.centerii.org/handbook/Resources/Chapter_4_Organizational_Structures.pdf). Accessed October 14, 2014.

1. Within each state, treatment schools with the same relevant grades will be grouped.
2. For each such treatment school grouping, its associated set of potential comparison schools will also be grouped.
3. For each treatment and potential comparison school within such paired groupings, the average of its mean reading proficiency percentage from 2012-2013 across its relevant grades and its mean math proficiency percentage from 2012-2013 across its relevant grades will be computed.
4. An average proficiency percentage caliper of 15% will be employed in the matching. This means that a treatment school can only be matched to a potential comparison school whose average proficiency percentage (computed in stage 3) is within 15% of its own. For instance, a treatment school with an average proficiency percentage of 35% could be matched to a potential comparison school with an average proficiency percentage within the 20%-50% range, but not to a potential comparison school outside of this range. This ensures that matches will occur between schools with reasonably similar levels of academic achievement.
5. A relevant faculty size cutoff will be employed in the matching. Relevant faculty size refers to the number of teachers teaching in one of the relevant grades. A treatment school can only be matched to a potential comparison school whose relevant faculty size is at least 80% of its relevant faculty size. For instance, a treatment school with a relevant faculty size of 25 could only be matched to potential comparison schools with relevant faculty sizes of at least 20. This promotes treatment and their matched comparison schools having comparable relevant faculty sizes. The focus on number of faculty is deliberate, and of practical importance, as the study intends to survey the same number of teachers from each treatment school and its matched comparison school, and if the relevant faculty size of the comparison school is appreciably smaller, surveying the same number of faculty would pose a greater proportional burden on the comparison school. .
6. Mahalanobis distances (Rosenbaum, 2010) will be computed between grouped treatment schools and their associated grouped potential comparison schools. These distances will be computed from schools' profiles of the key characteristics (as measured in 2012-2013): mean reading proficiency, mean math proficiency, relevant faculty size, % minority students, % Free/Reduced Price Lunch (FRPL), and % students with disabilities.
7. Order the treatment schools in a grouping as follows: convert their average proficiency percentages to z scores, standardizing within the grouping. Then order these treatment schools in descending order by the absolute value of their z score. This means that the treatment schools with the most extreme average proficiency percentages will be matched first.
8. In order, for each treatment school within the treatment grouping, select its initial matched comparison school as follows:
  - a. If there is at least one potential comparison school from the associated grouping that respects the proficiency caliper and the faculty size cutoff and is from the same school district, select the potential comparison school with the smallest Mahalanobis distance from the treatment school. Then remove this school from the potential comparison grouping.
  - b. Otherwise, if there is at least one potential comparison school from the associated grouping that respects the proficiency caliper and the faculty size cutoff but is from a different school district, select the potential comparison school with the smallest Mahalanobis distance from the treatment school. Then remove this school from the potential comparison grouping.

- c. Otherwise, there is no potential comparison school from the associated grouping that respects the caliper and the cutoff. Then select the potential comparison school whose average proficiency percentage is closest to that of the treatment school and remove this school from the potential comparison grouping.
9. The remaining schools in the potential comparison grouping will serve as potential replacements in case any of the matched comparison schools refuse to participate in the evaluation. If a treatment school experiences such a refusal, then steps 8a, 8b, and 8c will be repeated for that treatment school with respect to the remaining schools in the potential comparison group.

By utilizing the proficiency caliper, this procedure prioritizes matched pairs being similar on academic achievement. This is consistent with the What Works Clearinghouse (WWC: <http://ies.ed.gov/ncee/wwc/>) emphasis on demonstrating baseline equivalence on academic achievement. By utilizing the relevant faculty size cutoff, the procedure prioritizes matching schools of comparable faculty size, as needed in the sampling plan. By considering whether a potential comparison school is from the same district as a treatment school, it prioritizes limiting the number of school districts to recruit.

### *SIG Model*

This matching procedure ignores SIG model (after one initial screening criterion – see above) for several reasons. First, schools using the closure model will be excluded from the pool of potential comparison schools, eliminating the need to match on the closure model. Second, for the purposes of the specific intervention challenges that the School Turnaround AmeriCorps program is designed to help SIG/ Priority schools address (i.e., increased learning time, turnaround leadership, students’ nonacademic needs, and community/family engagement), the strategies used by the transformation and turnaround models are very similar. The key differences in these models are primarily related to governance and management of school personnel (e.g., hiring, evaluating performance, professional development), which fall outside the scope of AmeriCorps members’ influence.

Third, the restart model is distinct from the turnaround and transformation models, involving conversion or closure of an existing school and reopening the school as a charter school or under the control of an education management organization.<sup>2</sup> Considering SIG model as a key characteristic for matching the profiles of treatment and comparison schools is thus most relevant for schools with the restart model. However, this model is believed to be most distinct in the first few years of implementation. Within the treatment cohort, only two schools using the restart model are in their second year of implementation; all others are in the third year of the restart model or even further along. Thus, differences related to the organizational and operational challenges of undergoing a restart transition are hypothesized to be less significant in the third year or beyond, while the instructional strategies, increased learning time and other non-academic support strategies of a restart model are hypothesized to resemble those implemented in turnaround and transformation schools, justifying matching more “mature” restart schools with schools using the other two SIG models. Finally, none of the data analyses look separately at restart versus transformation/turnaround schools. There are relatively few restart schools, so such data analyses would have low statistical power.

To address a small number of exception cases, such as treatment schools in an early stage of the restart model or treatment schools that have no comparison school in the same state and the same relevant grades, we anticipate that a small number of treatment schools will need to be dropped because of lack of an adequate matched comparison school.

---

<sup>2</sup> *ibid.*

In the following section we discuss sampling teachers, principals, AmeriCorps members, and grantee staff.

### ***Sampling for Conducting Surveys***

#### *Sampling Teachers*

*Sampling teachers from treatment schools.* Each of the 62 treatment schools is conceptualized as a stratum of the teacher population of interest, which is the union of all teachers teaching in a relevant grade<sup>3</sup> in a treatment school. We refer to these teachers as *relevant teachers*. We will approximate proportional allocation stratified sampling, under which the same proportion of relevant teachers (the *sampling rate*) is sampled from each school. When implemented exactly, this results in all survey respondents being assigned the same survey weight, which leads to smaller standard errors for survey-based estimates than when there is variation in the size of sampling weights. The evaluation design calls for at least 348 completed treatment teacher surveys at each measurement occasion, and this implies that in the absence of survey nonresponse, the sampling rate would be 348 divided by the size of the relevant teacher population; given the reality of survey nonresponse, the actual sampling rate will be larger than this. We previously estimated the size of the relevant teacher population at about 2,500 teachers; this estimate was based on school faculty size data from the Common Core of Data (CCD) and other websites. In the ideal case, we will be able to determine the exact number of relevant teachers for each treatment school prior to beginning survey sampling, as this will give us the exact population size and hence will allow us to compute the exact sampling rate (assuming a perfect response rate). If it is not possible to obtain this information from all schools prior to the start of sampling, then we will estimate the relevant teacher population from whatever exact faculty size data we have obtained, plus estimated faculty sizes from the other schools, and we will then compute the sampling rate based on this estimate.

The target response rate is 80%. The sampling rate assuming perfect response divided by the anticipated response rate gives the sampling rate needed to obtain the desired 348 completed surveys. For instance, if the relevant teacher population is 2,500, then 14% of the relevant teachers per school ( $348/2500$ ) would be asked to complete a survey in the absence of any nonresponse; if the response rate is 80%, then the sampling rate should be 17.5% ( $14\%/0.80$ ). To err on the side of caution, we will use a sampling rate of 25%. This compensates for slightly missing the 80% goal and allows for erroneous teacher contact information or a possible underestimate of teacher population size. This means that about a quarter of the relevant teachers at each treatment school will be asked to complete a survey at each measurement occasion. In addition, at schools with a small relevant faculty size, we will always ask at least 5 teachers to complete surveys in an attempt to have at least 3 completed surveys per school per measurement occasion. Sampling will be implemented by first randomly ordering each school's roster of relevant teachers and then selecting the first quartile of the teachers on the roster.

Since there are multiple measurement occasions, we could sample teachers once for the entire evaluation, or at each measurement occasion. The latter approach lessens the burden on individual teachers, as given teachers will not be surveyed at every measurement occasion, but it rules out the possibility of longitudinal data analyses, in which individual teachers are followed over time. We believe lessening the burden on teachers, with a presumed bolstering of the response rate, warrants relinquishing our capacity to perform longitudinal data analyses. To amend the just-discussed sampling procedure, at the beginning of the evaluation, the first quartile of teachers on the randomly ordered roster will be asked to complete

---

<sup>3</sup> As defined in the discussion of matching, a *relevant grade* in a given treatment school is a grade receiving service from AmeriCorps members. In some schools, AmeriCorps members only provide assistance to selected grades.

the fall survey and the second quartile of teachers on the roster will be asked to complete the spring survey. This ensures that no teacher is asked to complete two surveys in a given school year.

*Sampling teachers from matched comparison schools.* At a given measurement occasion, the same number of comparison teachers will be asked to complete a survey as at the treatment school to which the comparison school is matched. All surveyed comparison teachers will be from the treatment school's relevant grades. Given that the matching procedure prioritizes similar treatment school/comparison school faculty sizes, the burden imposed on comparison school faculties should be approximately the same as that imposed on treatment school faculties. The same kind of randomly ordered roster procedure used to select treatment school teachers will be used with comparison schools.

#### *Sampling Principals*

At each measurement occasion, each principal at each participating school (treatment and comparison) will be asked to complete a survey. If a principal refuses, then his/her school's vice principal or an assistant principal will be asked to complete the survey.

#### *Sampling AmeriCorps Members*

There are about 440 AmeriCorps members working each year at the treatment schools. Every AmeriCorps member will be surveyed once.

#### *Sampling Grantee Staff*

There are 13 grantees in the first and second years of the program's operation. Grantee participation in the evaluation is mandatory under the program guidelines. A senior staff member from each grantee will be surveyed. The staff member should be familiar with the grant requirements, program theory of change and the interventions targeted to support schools' turnaround plans, school partners and the implementation of school partnership agreements, orientation and training of AmeriCorps members, and the management, tracking and monitoring of member activities. The appropriate staff member will be designated by the grantee.

### ***Sampling for Conducting Interviews***

#### *Sampling Teachers*

26 treatment teachers and 26 comparison teachers will be interviewed once. Two teachers will be selected per grantee from its associated treatment schools, and two teachers from the matched comparison schools will also be interviewed. Relevant teachers will be selected randomly from treatment schools and comparison teachers will be randomly selected from their matched treatment schools' relevant grades. The teacher survey and interview samples will be mutually exclusive to ensure that no one teacher is both surveyed and interviewed during the school year.

#### *Sampling Principals*

26 principals of treatment schools and 26 principals of comparison schools will be interviewed twice. Two principals will be randomly selected from each grantee's associated treatment schools, and the principals at the matched comparison schools will also be interviewed.

#### *Sampling AmeriCorps Members*

26 AmeriCorps members will be interviewed once. Two AmeriCorps members will be randomly selected per grantee.

*Sampling Parents*

50 parents of students served by AmeriCorps members at treatment schools will be interviewed once. 3-4 parents will be selected from each grantee. Grantees and schools will be asked to nominate parents who may be willing to participate in the interview.

*Sampling Grantee Staff*

A senior staff member from each grantee will be interviewed twice. The staff member will be designated by the grantee. The staff member should have a similar level of familiarity with the program as the member designated to respond to the grantee survey

**Surveys: Respondent Universe, Sample Size, and Target Response Rates**

The potential respondent universe for the national evaluation of School Turnaround AmeriCorps consists of all grantee staff, AmeriCorps members, principals, and teachers at schools which receive School Improvement Grant (SIG) funds to implement one of the four SIG models and in Priority schools that are implementing interventions aligned with the ESEA flexibility turnaround principles (Priority schools). In the 2013-14 school year, this universe was composed of 72 schools that received funding through one of the 13 grants issued by the School Turnaround AmeriCorps program (hereafter referred to as the Program) and the remaining schools that did not receive grants from this program. The 72 schools that received grants from the Program consist of 15 schools that are partnered with Teach for America (TFA) and 57 other schools. The evaluation will include a total of 62 schools, 5 of which are partnered with TFA.

Exhibit B-1 presents the size of the population, the size of the survey sample to be selected, the target response rate, and the target respondent sample size for surveys of school principals, teachers, program grantees, and AmeriCorps members.

**Exhibit B-1: Surveys: Population Size, Sample Size, and Target Response Rate**

<b>Sample Unit</b>	<b>Population</b>	<b>Sample to be Selected</b>	<b>Target Response Rate</b>	<b>Target Respondent Sample</b>
<b>1. Principals (one principal per school)</b>				
Program schools	72 principals	62 principals	100%	62 principals
Non-Program SIG schools	At least 72 principals from matched comparison schools	62 principals	100%	62 principals
<b>2. Teachers in grade levels served by the Program</b>				
Program schools	2,510 teachers	435 teachers	80%	348 teachers

Non-Program SIG schools	At least 2,510 teachers from matched comparison schools	435 teachers	80%	348 teachers
<b>3. Parents of students served by AmeriCorps members</b>	26,100 parents	50 parents	100%	50 parents
<b>4. Program grantees</b>	13 grantees	13 grantees	100%	13 grantees
<b>5. AmeriCorps members</b>	440 individuals	440 individuals	100%	440 individuals

**B2. Describe the procedures for the collection of information, including: Statistical methodology for stratification and sample selection; Estimation procedure; Degree of accuracy needed for the purpose described in the justification; Unusual problems requiring specialized sampling procedures; and any use of periodic (less frequent than annual) data collection cycles to reduce burden.**

The study’s multi-method data collection approach will examine: 1) how AmeriCorps members are supporting school turnaround efforts; 2) differences between school turnaround efforts at sites supported by AmeriCorps and similar sites not supported by AmeriCorps; and 3) best practices as well as implementation challenges for the program. Data collection will also gather information about local context that may affect program implementation as well as the feasibility of replication of different program elements. Together, these data will inform understanding of the value-added of AmeriCorps above and beyond the other school turnaround resources invested in School Improvement Grant (SIG) and Priority schools (see Exhibit B-2).

**Exhibit B-2: Data Sources by Research Questions, Timing of Collection and Use in Analysis<sup>4</sup>**

Research Question	Data Sources	Timing			Use in Analysis	
		Late Fall 2014	Winter / Early Spring 2014-15	Late Spring 2015	Descriptive	Pre/Post
<b>1) How do AmeriCorps members help schools implement their turnaround plans?</b>						
a) How do AmeriCorps grantees work with teachers and other school personnel to identify and target students with whom their members will engage so that the school is more	Grantee staff survey		✓		✓	
	Grantee staff interviews	✓		✓	✓	✓
	Teacher survey	✓		✓	✓	✓
	Teacher interviews		✓		✓	

<sup>4</sup> Data collection activities in Year 2 will likely mirror those in Year 1 (both type and timing).



Research Question	Data Sources	Timing			Use in Analysis	
		Late Fall 2014	Winter / Early Spring 2014-15	Late Spring 2015	Descriptive	Pre/Post
likely to achieve its turnaround goals?	Grantee progress reports		✓	✓	✓	✓
f) What are the specific direct service activities and school-level interventions that AmeriCorps members conduct at each school and how are those activities believed to support school turnaround?	Grantee staff survey		✓		✓	
	Grantee staff interviews	✓		✓	✓	✓
	AC Member interviews		✓		✓	
	Principal survey	✓		✓	✓	✓
	Principal interviews	✓	✓		✓	✓
	Teacher interviews		✓		✓	
	Parent interviews		✓		✓	
	Grantee progress reports		✓	✓	✓	✓
	Grantee activity logs		✓	✓	✓	✓
o) What are the specific capacity-building strategies that AmeriCorps members contribute to each school? How do school leaders and staff view the role and contributions of AmeriCorps members in building the school's capacity to implement their turnaround effort? What are the areas in which schools believe AmeriCorps members have the most and least influence over the school's ability to achieve its turnaround goals, and why? In what ways, if any, does the presence of AmeriCorps members allow school staff or volunteers to modify their activities in ways that might benefit students?	AC Member interviews		✓		✓	
	Principal survey	✓		✓	✓	✓
	Principal interviews	✓	✓		✓	✓
	Teacher survey	✓		✓	✓	✓
	Teacher interviews		✓		✓	
	Grantee progress reports		✓	✓	✓	✓
	Grantee activity logs		✓	✓	✓	✓

Research Question	Data Sources	Timing			Use in Analysis	
		Late Fall 2014	Winter / Early Spring 2014-15	Late Spring 2015	Descriptive	Pre/Post
v) Do the specific activities that AmeriCorps members conduct change over the course of the grant period? To what extent do grantees use data to inform continuous improvement efforts to meet changing needs and improve their interventions?	Grantee staff interviews	✓		✓	✓	✓
	AC Member interviews		✓		✓	
	Principal survey	✓		✓	✓	✓
	Principal interviews	✓	✓		✓	✓
	Grantee progress reports		✓	✓	✓	✓
	Grantee activity logs		✓	✓	✓	✓
<b>2) How and to what extent do School Turnaround AmeriCorps programs adhere to grantees' program designs across schools or exhibit flexibility to adapt to schools' needs and local contexts?</b>						
a) Which aspects of grantee-school partnerships appear to be the most promising practices in terms of satisfaction of the school leadership and the participating AmeriCorps members?	Grantee staff survey		✓		✓	
	Grantee staff interviews	✓		✓	✓	✓
	AC Member survey			✓	✓	
	Principal survey	✓		✓	✓	✓
	Principal interviews	✓	✓		✓	✓
	Teacher survey	✓		✓	✓	✓
	Grantee progress reports		✓	✓	✓	✓
h) What elements of the implementation are sensitive to local contexts and might be difficult to generalize and replicate in other contexts?	Grantee staff interviews	✓		✓	✓	✓
	Grantee staff focus groups		✓		✓	
	AC Member interviews		✓		✓	
	AC Member focus groups		✓		✓	
	Principal interviews	✓	✓		✓	✓
	Teacher interviews		✓		✓	
	Principal/ teacher focus groups		✓		✓	
	Grantee progress reports		✓	✓	✓	✓

Research Question	Data Sources	Timing			Use in Analysis	
		Late Fall 2014	Winter / Early Spring 2014-15	Late Spring 2015	Descriptive	Pre/ Post
p) Which elements of implementation are potentially replicable in other schools?	Grantee staff interviews	✓		✓	✓	✓
	Grantee staff focus groups		✓		✓	
	AC Member focus groups		✓		✓	
	Principal/ teacher focus groups		✓		✓	
	Grantee progress reports		✓	✓	✓	✓
<b>3) Are AmeriCorps members perceived by school leaders and other stakeholders to be more vital in supporting certain SIG/Priority strategies than others? Which activities pursued by AmeriCorps members are perceived as being more or less helpful in supporting schools' turnaround efforts with respect to the following outcomes, and why?</b>						
a) Overall success in school turnaround?	Grantee staff survey		✓		✓	
	Grantee staff focus groups		✓		✓	
	AC Member survey			✓	✓	
	AC Member focus groups		✓		✓	
	Principal survey	✓		✓	✓	✓
	Teacher survey	✓		✓	✓	✓
	Principal/ teacher focus groups		✓		✓	
	Parent interviews		✓		✓	
i) Academic achievement?	Grantee staff survey		✓		✓	
	Grantee staff focus groups		✓		✓	
	AC Member survey			✓	✓	
	AC Member focus groups		✓		✓	
	Principal survey	✓		✓	✓	✓
	Teacher survey	✓		✓	✓	✓
	Principal/ teacher focus groups		✓		✓	
	Parent interviews		✓		✓	
	Student outcome data			✓	†	

Research Question	Data Sources	Timing			Use in Analysis	
		Late Fall 2014	Winter / Early Spring 2014-15	Late Spring 2015	Descriptive	Pre/Post
r) Students' socio-emotional health?	Grantee staff survey		✓		✓	
	Grantee staff focus groups		✓		✓	
	AC Member survey			✓	✓	
	AC Member focus groups		✓		✓	
	Principal survey	✓		✓	✓	✓
	Teacher survey	✓		✓	✓	✓
	Principal/ teacher focus groups		✓		✓	
	Parent interviews		✓		✓	
	Student outcome data			✓	†	
aa) School climate?	Principal survey	✓		✓	✓	✓
	Teacher survey	✓		✓	✓	✓
	Parent interviews		✓		✓	
dd) School capacity to implement its turnaround effort?	Grantee staff focus groups		✓		✓	
	AC Member interviews		✓		✓	
	Principal survey	✓		✓	✓	✓
	Principal interviews	✓	✓		✓	✓
	Teacher survey	✓		✓	✓	✓

† Student outcome data will be collected from grantees and their partner schools, not districts or states. Prior to use in analysis, it will be necessary to determine the availability and completeness of student achievement and attendance/behavior data.

## Primary Data Collection

### Surveys

In order to gather information about how the program is being implemented, the school climate, and how the program is helping schools improve compared with schools without School Turnaround AmeriCorps, we will administer online surveys (taking no more than 30 minutes) to all grantee staff, AmeriCorps members, AmeriCorps and comparison principals, and a sample of AmeriCorps and comparison teachers. Surveys will provide data from the broadest group of AmeriCorps members, principal, and teacher respondents possible, relative to more targeted interviews and focus groups, and will help anchor the interpretation of qualitative data collection. Comparison school principals and teachers will be surveyed about school climate, perceptions of school improvement, and any community involvement and partnerships that support school turnaround efforts.

The initial sample and timing for administration is summarized in Exhibit B-3. All grantees will be surveyed during the spring of 2015. AmeriCorps members will be surveyed during the spring 2015. All AmeriCorps and comparison principals and a sample of teachers from each school will be surveyed twice during the 2014-15 school year (fall/winter and spring).

**Exhibit B-3: Initial Survey Sample by Respondent and Time**

	Fall/Winter 2014	Spring 2015
<b>Grantee staff</b>	N/A	13
<b>AmeriCorps members</b>	N/A	440
<b>Principals (AmeriCorps and comparison)</b>	124	124
<b>Teachers (AmeriCorps and comparison)</b>	1160*	1160*

\*The initial teacher survey size is based on calculations to obtain power=0.80 when 2-sided alpha=.05, assuming true difference in endorsement rates of 10% and allowing for nonresponse and erroneous contact information.

The contractor will work with a survey subcontractor, to program the final survey instruments in online format. The contractor’s staff will work with grantees and school districts to obtain respondent email addresses. At the beginning of each survey administration, each respondent will be emailed an individualized survey link and will have 3-4 weeks to take the survey at their convenience. Regular reminders will be emailed to respondents who have not yet completed the survey over the course of the field period.

To maximize response rates, the contractor will use district-specific liaisons to help communicate with schools about the importance of the survey effort. They will also identify a staff person at each school who can assist with encouraging and reminding educators to complete surveys. They will have access to real-time response rates to target follow-up communications, and study team staff will communicate with potential respondents and send reminders over the course of each field period. They plan to use multiple modes for reminders including email, mail (e.g., fliers for teachers’ school mailboxes), and telephone calls with school liaisons.

**Interviews**

Semi-structured telephone interviews will be conducted with grantee staff, and a sample of AmeriCorps members, principals (AmeriCorps and comparison sites), teachers (AmeriCorps and comparison sites), and parents (AmeriCorps sites only). The purpose of these interviews is to collect more detailed information from grantees, members, and AmeriCorps principals and teachers about the program’s structure, activities, implementation successes and challenges, and perceived effects of the program. Comparison school principals and teachers will be interviewed about the activities and perceived effects of any volunteers, support staff, or external partners whose role it is to support school-wide improvement.

Parent interviews will be limited to a purposeful subsample of schools where family engagement is a clearly articulated element of the SIG strategy. In schools that focus grantee activities on parent engagement, parents will be asked about their awareness of the program, general perceptions of their child’s school, and perceptions of the AmeriCorps program. The parent interviews will provide valuable information about the implementation of parent engagement activities and perceptions of the program, as well as an external stakeholder perspective on the topic of school climate. The sample and timing for the interviews are summarized in Exhibit B-4.

**Exhibit B-4: Interview Sample by Respondent and Time**

	Fall/Winter 2014	Spring 2015
<b>Grantee staff</b>	13	13

<b>AmeriCorps members</b>	N/A	26
<b>Principals (AmeriCorps and comparison)</b>	62	62
<b>Teachers (AmeriCorps and comparison)</b>	0	62
<b>Parents (AmeriCorps)</b>	0	50

At the beginning of each data collection period, trained interviewers will provide potential respondents with information about the purpose and content of the interview, and will ask for verbal or written consent, as required by the contractor’s IRB. The 30-minute telephone interviews will be scheduled in advance during times convenient for respondents. Parents will be offered a modest incentive (gift card) for their participation in the interviews.

### ***Focus Groups***

Focus groups will be conducted with grantee staff, and a sample of AmeriCorps members, principals, and teachers (AmeriCorps sites only). The focus groups will provide more in-depth information about the factors that facilitate progress for school turnaround and implementation successes and challenges. The sample and timing for the focus groups is summarized in Exhibit B-5.

### **Exhibit B-5: Focus Group Sample by Respondent and Time**

	<b>Winter 2014-2015</b>	<b>Spring 2015</b>
<b>Grantee staff</b>	13	N/A
<b>AmeriCorps members</b>	39	N/A
<b>Principals (AmeriCorps)</b>	N/A	8
<b>Teachers (AmeriCorps)</b>	N/A	24

During the winter of 2014-15, online focus groups will be conducted with grantee staff and a sample of AmeriCorps members. The web-based approach will allow an open exploration of implementation experiences and perceptions of the program. The contractor has used this approach successfully for other evaluations and will work individually with participants to ensure that they have access to the necessary technology and equipment, and provide any support needed to fully participate in the focus groups. Two grantee and up to five AmeriCorps member focus groups, lasting between 30-45 minutes will be conducted to accommodate schedules and manage the size of the groups.

During the spring of 2015, eight principal and teacher focus groups will be conducted on-site at AmeriCorps schools. Each focus group will consist of the principal and three teachers. Two trained facilitators will conduct each focus group; one will serve as the moderator to facilitate the focus group accompanied by an assistant moderator, who will assist with logistics, observe the group dynamics and, later, assist with coding or analysis of the group discussion. Focus group participants will complete consent forms in advance of the discussion. The moderator will encourage an open, free-flowing discussion by emphasizing that there are no right or wrong answers and that individual responses will be analyzed and reported at the aggregate level.

### **Secondary Data Collection**

The approach to secondary data collection will focus on collecting data related to program implementation from grantees and their partner schools. It will also involve collecting student outcome data such as achievement and attendance data from grantees and schools to determine the availability and completeness of those data, and for supplemental analysis if possible.

### ***Grantee Performance Measure Data and Mid-year and Annual Grantee Progress Reports (GPRs)***

Grantees are required to submit progress reports to CNCS twice a year, mid-way into the grant and at the end of the grant year. The annual report should include the information for the full program year through September 30th, not just for the period since the mid-year report. The contractor will collect these reports from CNCS (rather than from grantees) at two points during the Year 1 evaluation, in the fall and spring. The GPR consists of the following sections: Demographic Information, MSYs/Members, Performance Indicators, Performance Measures, and Narratives. Completed GPRs will contain performance measure data, including the outputs and outcomes the grantee selected and the targets and actual measures, and other quantitative data such as Member Service Year (MSY) and member counts. They also contain fields for grantees to provide narrative explanations of their performance data and of the program overall. Grantees are instructed to provide a narrative analysis of the impact of AmeriCorps members' service in the community that would not have been possible through existing staff and/or volunteers, how the members have enabled the program to leverage new public-private partnerships and funding, and any factors or trends that positively or negatively affected their program's performance. Optionally, grantees can include "impact snapshots" or examples of a change in beneficiary knowledge, attitude, behavior or condition that the program has been able to measure. In addition, they are required to describe any activities and accomplishments relative to member experience that were not captured in national performance measures, as well as training, technical assistance, and monitoring activities.

### ***Grantee Activity Logs***

We will collect grantee activity logs directly from grantees on a quarterly basis. Because the fourth quarter falls outside of the contract period of performance, we will collect a total of three quarters of grantee activity logs and performance data during Year 1 of the evaluation. There is no standardized form or template for grantee activity logs, therefore we expect the form and content (and possibly the instrument name) to vary across grantees. Some grantees may not use activity logs. These logs may contain a combination of quantitative and qualitative information describing the types, frequency, duration, and intensity of member activities, details about the nature, venue, and timing of activities, and information on service beneficiaries, including number of persons served and their specific needs related to the intervention.

### ***Student Achievement/Attendance Data***

Student outcome data will be collected from grantees and their partner schools, not districts, states, or comparison schools. We anticipate these will consist of student-level achievement test data and/or attendance and/or behavior data. Currently we plan to collect these data directly from grantees on a quarterly basis together with the grantee activity logs. Prior to use in analysis, it will be necessary to determine the availability and completeness of student achievement and attendance and/or behavior data. Collection of these data may occur less frequently depending on availability.

### ***Analyses, Analysis Methods, and Degree of Accuracy***

Many of the evaluation's research questions can be addressed by estimating the prevalence of responses for principals, teachers, and parents for the treatment group. For these analyses, the expected sample sizes and estimated degree of accuracy/margin of error, in terms of the half width of the 95 percent confidence interval, is provided in Exhibit B-6.

**Exhibit B-6: Sample Size and Degree of Accuracy (Margin of Error) for Analyses of the Treatment Group Only**

Respondent	Sample Size (Treatment Group Only)	Margin of Error
Principal	62 principals	None—estimate based on full population
Teacher	348 teachers	5 percentage points
Parent	50 parents	More than 10 percentage points
Program grantees	13 grantees	None—estimate based on full population
AmeriCorps members	440 individuals	None—estimate based on full population

Some of the evaluation’s research questions can be addressed by comparing the prevalence of responses for principals, teachers, and parents between the treatment group and the comparison group. For these analyses, the expected sample sizes and estimated degree of accuracy/margin of error, in terms of the half width of the 95 percent confidence interval, is provided in Exhibit B-7.

**Exhibit B-7: Sample Size and Degree of Accuracy (Margin of Error) for Analyses of Differences Between the Treatment and Comparison Groups**

Respondent	Sample Size (Treatment and Comparison)	Margin of Error
Principal	124 principals	None—estimate based on full population
Teacher	696 teachers	10 percentage points

***Analysis of Survey Results***

The survey analysis has two main goals. First, the surveys of grantee staff and AmeriCorps members will provide insights into how the AmeriCorps members are supporting school turnaround efforts. Second, the surveys of principals and teachers in low performing schools with School Turnaround AmeriCorps and principals and teachers in matched comparison schools that are also low performing but with little or no AmeriCorps presence will provide insights into the perceived effectiveness of the program.

To meet these goals, we will conduct cross-sectional and pre-post analyses, as described below. In all cases, we will conduct design-based rather than model-based analyses (Lohr, 1999), as our aim is to learn about the current populations of grantee staff, AmeriCorps members, principals, and teachers at the



specific set of 62 treatment schools (and their grantees) participating in the evaluation. Surveyed teachers and principals at comparison schools will serve as foils for the treatment school teachers and principals.<sup>5</sup>

The cross-sectional analyses examine survey responses collected at a single measurement occasion and show the state of the School Turnaround AmeriCorps world at that point in time. Because there is a single grantee staff survey and a single AmeriCorps member survey, only cross-sectional analyses will be performed for these surveys, and they will attempt to characterize the views of staff across all 13 grantees and of all AmeriCorps members serving in the 62 participating treatment schools, respectively. Because there are pre- and post-surveys for teachers and principals, there will be two cross-sectional analyses for each, one for the pre-survey and another for the post-survey. These analyses will compare the views of teachers/principals from treatment schools with those of the corresponding staff from the matched comparison schools. The aim here is to characterize how the views of the set of all teachers/principals at the 62 treatment schools may or may not systematically differ from those of the corresponding set of all teachers/principals at the matched comparison schools. For example, the proportion of treatment teachers who endorse a certain view at a given point in time may or may not differ from the proportion of comparison teachers who endorse that view.

The pre-post analyses compare responses collected at different measurement occasions and show how the state of the School Turnaround AmeriCorps world differs from one point in time to another. Only teacher and principal surveys will receive pre-post analyses. We will conduct two kinds of pre-post analyses. The first will compare the responses of teachers/principals from the post-survey to the responses from the corresponding pre-survey. These will characterize how the views of teachers/principals at the 62 participating treatment schools may or may not differ in spring, 2015 as compared to fall, 2014. The second kind of pre-post analysis is sometimes referred to as a difference-in-differences analysis, and will examine how pre-vs-post survey differences among treatment school staff may or may not differ from pre-vs-post survey differences among comparison school staff. For instance, it may be that treatment teachers tend to look increasingly optimistic about a certain aspect of their school's functioning, whereas comparison teachers' views remain constant. A difference-in-differences analysis will allow us to detect a change in response between the two groups of survey participants.

### *Examining Baseline Equivalence*

We will follow the lead of the What Works Clearinghouse (WWC) in examining baseline (i.e., pre-treatment, not pre-evaluation) equivalence of treatment and comparison schools. The WWC calls for the assessment of baseline equivalence in terms of standardized differences in treatment vs. comparison school means on key pre-treatment school characteristics, such as school proficiency percentages on reading and on math achievement. We will examine baseline equivalence on all pre-treatment school characteristics used in matching treatment and comparison schools, including proficiency percentages, ethnicity, and percent of students receiving free or reduced price lunch. The WWC operationalizes "acceptable baseline equivalence" in terms of a cutoff of .25 of a standard deviation, and any school characteristics exhibiting standardized differences above the cutoff will be highlighted.

Note that we will not be able to perform any baseline balance testing of teacher or principal characteristics. In order to do so, we would have to know the characteristics of the teachers and principals employed at the treatment and comparison schools *prior* to the start of the intervention, but we only have potential access to the characteristics of school staff who are employed at the schools one year after the

---

<sup>5</sup> Note, per the sampling plan, that the 62 participating treatment schools are conceptualized as strata, not as clusters. This means that we are focused on drawing inferences about individuals associated with this specific set of schools rather than about all the individuals associated with some larger population of schools.

start of the intervention. Also note that if a school drops out of the evaluation, then its matched school staff will be excluded from any analyses comparing treatment to comparison schools.

### *Creation of Sampling Weights*

We plan to survey all AmeriCorps members, all grantees, and all principals of the treatment and comparison schools, but only a subsample of the teachers. For teachers, sampling base weights will be created to ensure that the sampled teachers accurately represent the full populations of relevant teachers at the treatment and comparison schools. We will calculate sampling base weights as the inverses of the probability of being invited to complete a survey. As described below, we will then adjust these weights to account for survey nonresponse.

### *Computation of Standard Errors and Margins of Errors*

Computing accurate standard errors and margins of errors can be challenging in the context of non-response adjustments to sampling weights, as the usual methods for computing standard errors do not account for the extra variability introduced by estimating nonresponse adjustments.<sup>6</sup> We will compute them using a jackknife procedure employing replicate weights.<sup>7</sup> This procedure is also able to accurately account for stratification and clustering in a sampling design. Margins of error are half the width of a 95% confidence interval, and will be computed from standard errors by using normal distribution approximations.

### *Scale Development*

Creating valid and reliable scales to assess important constructs can help answer some of the important questions for CNCS. A poorly developed scale can lead to inaccurate results and inferences. We will develop relevant scales consisting of a set of survey items that define a common construct; e.g., scales for capacity to implement turnaround model, student and family engagement and school climate. To be used in data analyses, scales must exhibit an adequate level of internal consistency, which we operationalize as Cronbach's alpha of at least .65.

### *Descriptive Statistics of All Relevant Teachers*

Along with the request for teacher rosters from each treatment and comparison school, we will request demographic characteristics of the teachers. As discussed above, knowledge of these characteristics for all relevant teachers will help in developing accurate nonresponse-adjusted sampling weights. In addition, if we are able to obtain this information, it will be used to compare the characteristics of all treatment teachers with those of all comparison teachers. Note that no sampling weights are needed to conduct such comparisons, as we would have a census of all relevant teachers for these characteristics. Also note that there is no reason to expect balance on these characteristics, as they are measured one year into the intervention on the teachers currently employed.

## **Analysis of Survey Responses**

---

<sup>6</sup> Valliant, R., Dever, J. A. & Kreuter, F. (2013). Practical tools for designing and weighting survey samples, pp. 418-422.

<sup>7</sup> The contractor has successfully implemented jackknife procedure employing replicate weights in the Massachusetts Educator Evaluation Framework study as well as USDA's Evaluation of the Impact of the Summer Food.

### *Analysis of Grantee Staff and AmeriCorps Members*

For grantee staff and AmeriCorps members' surveys, we will report on the percentage of members who agree on a (dichotomized) survey question or a scaled construct, and will also report margins of error. These analyses, as well as those discussed below, can be easily implemented in Stata or SAS.

#### *Cross-sectional Analysis of Teachers and Principals (for pre-survey as well as post-survey separately)*

The cross-sectional analysis of teacher and principal survey responses will examine differences between treatment and comparison school teachers and principals on key survey questions/scaled constructs. We will report on the percentage of treatment group teachers/principals who agree to a particular survey question, or give scale means, as compared to the comparison group teachers/principals. We will also calculate the estimated difference and report on its statistical significance (p-value < 0.05). If a survey question or scaled construct is valid only for the treatment group, we will report on the percentage of respondents that agree with margins of error (similar to the grantee staff and AmeriCorps member survey).

More specifically, to compare treatment school teachers to comparison school teachers on a given scale or response, we will perform a design-based analysis of the following regression model

$$Y_{ij} = \beta_0 + \beta_1 \text{Group}_j + \beta_2 X_j + \varepsilon_{ij}$$

where  $Y_{ij}$  is the outcome of the  $i^{\text{th}}$  teacher in the  $j^{\text{th}}$  school;  $\text{Group}_j$  indicates whether school  $j$  is a treatment (= 1) or comparison (= 0) school; and  $X_j$  is a vector of pre-intervention school characteristics used to match potential comparison to treatment schools (described in the discussion of matching above). The latter helps adjust for the fact that matching is necessarily imperfect, and  $\beta_1$  gives the adjusted mean treatment-vs-comparison difference.<sup>8</sup> The use of the jackknife procedure described above provides accurate standard errors. The principal survey will be analyzed in an analogous manner.

#### *Pre vs. Post Analysis of Teachers and Principals*

As discussed above, there are two kinds of pre-post analyses for the teacher and principal surveys. The first kind of pre-post analysis compares treatment school teachers'/principals' responses from the two surveys (this analysis is for survey questions that apply specifically to the treatment school staff). A design-based analysis of the following regression will be performed

$$Y_{ijt} = \beta_0 + \beta_1 t + \varepsilon_{ijt}$$

where  $Y_{ijt}$  is the outcome of the  $i^{\text{th}}$  teacher in the  $j^{\text{th}}$  school at measurement occasion  $t$  (0 = pre, 1 = post).  $\beta_1$  gives the mean post-vs-pre difference. The use of the jackknife procedure described above provides accurate standard errors. The principal survey will be analyzed in an analogous manner.

The second kind of pre-post analysis is a difference-in-differences analysis (this analysis is for survey questions that apply to both the treatment and comparison group staff). A design-based analysis of the following regression will be performed

---

<sup>8</sup> When  $Y_{ij}$  is binary, then this regression is called a linear probability model. Analyses with this model are easier to interpret than when logistic regression is used, and the results yielded tend to be similar.

$$Y_{ijt} = \beta_0 + \beta_1 t + \beta_2 \text{Group}_j + \beta_3 t * \text{Group}_j + \beta_4 X_j + \varepsilon_{ijt}$$

where  $Y_{ijt}$  and so on are defined as before and  $\beta_3$  gives the difference in differences estimate. The use of the jackknife procedure described above provides accurate standard errors. The principal survey will be analyzed in an analogous manner.

### *Reporting of Survey Results*

For items meaningful only to the treatment group teachers or principals, we will report responses with margins of error. For grantee and AmeriCorps member surveys, we will report on the percentage of members who agree on a survey question with margins of error. See Exhibit B-8 for a sample table shell.

#### **Exhibit B-8: Sample Table Shell for Reporting Cross-Sectional Treatment Group Survey Results**

Survey Question	Percent Agree	Margin of Error
Teachers in the school are supportive of the AmeriCorps program	XX.X	YY.YY

Notes: Sample Size = ZZZ

The cross-sectional analyses of teacher and principal surveys will examine differences between treatment and comparison school teachers on key survey questions; e.g., “XX percent of treatment group teachers agree there was an improvement in grades for one or more of their students this year, compared to YY percent of comparison group teachers, an estimated difference of +/-WW percent which is statistically significant if p-value < 0.05.” See Exhibit B-9 for a sample table shell.

#### **Exhibit B-9: Sample Table Shell for Reporting Cross-sectional Treatment and Comparison Group Survey Results**

Survey Question	Percent Treatment Group Agree	Percent Comparison Group Agree	Difference	p-value
Was there improvement in grades for one or more of your students at your school this year?	XX.X	YY.YY	-/+WW.W***	0.VVV

Notes: \*\*\*-p-value<0.01, \*\*-p-value<0.05, \* p-value<0.10

The pre- vs. post- analysis of teacher and principal surveys will examine differences between the pre-post survey questions of treatment and comparison school teachers; e.g., “XX percent change in agreement of treatment group teachers that their school sets high standards for academic performance for all students, compared to YY percent change in comparison group teachers, an estimated difference-in-difference of +/-WW percent which is statistically significant if p-value < 0.05.” See Exhibit B-10 for a sample table shell.

**Exhibit B-10: Sample Table Shell for Reporting Pre- Post Treatment and Comparison Group Survey Results**

Survey Question	Change in Treatment Group	Change in Comparison Group	Difference-in-Change	p-value
Sets high standards for academic performance for all students?	XX.X	YY.Y	-/+WW.W***	0.VVV

Notes: \*\*\*-p-value<0.01, \*\*-p-value<0.05, \* p-value<0.10

**Qualitative Analysis**

This study will generate qualitative data through both primary data collection (open-ended survey responses, interviews, focus groups) and secondary data collection (primarily, narrative responses to grantee progress reports and possibly also narrative descriptions included in grantee activity logs). Producing valid and reliable results for any qualitative research effort requires databases to be carefully managed and analysis to be systematic. Qualitative analysis conducted for this study will involve setting up a rigorous framework for organizing, coding and exploring data using NVivo 10.0, a software package designed for efficient data organization, and systematic, reliable and replicable analyses. This application also permits merging of close-ended attributes with open-ended data. For example, open-ended narrative responses can be imported from surveys to perform more sophisticated mixed method analyses and more readily observe patterns in the data. The team will develop a codebook and train the analysts on the codes and their definition, and decision rules regarding codes, to identify content with a high level of inter-rater reliability, thereby reducing sources of bias introduced to the study.

Once the data have been coded, the analysis team will utilize NVivo to run a series of queries to triangulate patterns in the data. As patterns are identified, we will examine differences and similarities across the intervention types, stakeholder groups, and sites where School Turnaround AmeriCorps is being implemented. We will also compare patterns in data across treatment and comparison schools to help explain survey findings. We will examine implementation activities in relation to the six SIG strategies and the research topics, including the key turnaround outcomes (overall success in school turnaround, academic achievement, students’ socio-emotional health, school climate, school capacity to implement its turnaround effort).

To produce summary memos of study findings throughout the period of performance, we will code and analyze the qualitative data in several waves as data become available. Exhibit B-11 shows the timeline for the four planned rounds of qualitative analysis and the data sources to be included in each round. During the first round, we will train the analysis team in the codebook and perform peer coding of a sample of the first dataset to ensure that team members are interpreting the meaning of the data in the same way and consistently applying the appropriate codes to the data. We will then review and revise the codebook and analysis strategy as needed.

**Exhibit B-11: Timeline and Scope of Qualitative Analysis Rounds**

Rounds of Qualitative Analysis	Start	Finish
<b>Round 1 - coding, queries and attribute analysis</b>	<b>10/29/14</b>	<b>2/12/15</b>
Planning, develop codebook, training preparation		
Round 1: analysis of 2013-14 mid-year and annual GPRs		
Setup NVivo file, import data, peer coding, quality reviews		
Revise codebook and analysis strategy		

Round 1: analysis of teacher & principal PRE-survey narrative responses, principal PRE-interviews, and grantee PRE-interviews		
<b>Round 2 - coding, queries and attribute analysis</b>	<b>10/30/14</b>	<b>3/20/15</b>
Round 2: analysis of grantee focus groups and AmeriCorps member focus groups		
<b>Round 3 - coding, queries and attribute analysis</b>	<b>3/30/15</b>	<b>6/10/15</b>
Round 3: analysis of parent interviews		
Round 3: analysis of AmeriCorps member interviews		
Round 3: analysis of 2014-15 mid-year GPRs		
<b>Round 4 - coding, queries and attribute analysis</b>	<b>5/11/15</b>	<b>7/7/15</b>
Round 4: analysis of teacher & principal POST-interviews and focus groups, and grantee POST-interviews		
Round 4: analysis of teacher & principal POST-survey narrative responses		
Round 4: analysis of grantee survey narrative responses and AmeriCorps member survey narrative responses		

### **Descriptive Analysis of Secondary Data**

This section explains the descriptive analyses we will perform with data obtained from grantees. We first discuss analyses of grantee performance measures provided in mid-year and end-of-year grantee progress reports. Second, we discuss analyses of grantee activity logs. Finally, we consider student achievement and attendance/behavior data.

#### *Grantee Performance Measures*

Grantee performance measures are reported in the mid-year and annual grantee progress reports (GPRs). They include CNCS-defined academic performance measures such as ED2 (number of students who completed K-12 education programs) and ED5 (number of students with improved academic performance in literacy and/or math), as well as measures of the success of implementing the desired AmeriCorps presence (number of FTE AmeriCorps members working at a grantee’s schools). The GPRs compare the target for each measure (e.g., target number of students with improved academic performance in literacy and/or math) to what the grantee’s schools were actually able to accomplish, and they note whether the target was met.

For each academic and implementation performance measure, we will create a grantee-by-reporting-occasion table to document each grantee’s progress on the measure over time. Proportions of grantees meeting their target and cross-grantee averages will be reported for each reporting occasion. Based on these measure-specific tables, we will characterize grantees’ progress over time.

#### *Grantee Activity Logs*

We have not yet reviewed examples of grantee activity logs, and therefore cannot be specific about how the data they contain will be analyzed. We expect these logs will at least include qualitative information (e.g., discussions of difficulties encountered in implementing the AmeriCorps interventions) and may also include quantitative information that supplements the separately-reported performance measures. Summaries of commonly-occurring themes in the logs will be developed, and descriptive statistics will be presented, as appropriate, for quantitative data that are commonly reported in the logs.

### *Student Achievement/Attendance Data*

We expect that some grantees will provide us with student-level achievement test data and/or attendance and behavior data. In the ideal case, we will receive data from three years: 2012-2013, the year prior to the start of School Turnaround AmeriCorps; 2013-2014, the first year of the intervention; and 2014-2015, the first year of the evaluation. Because we will not be receiving the analogous data from the comparison schools matched to the grantees' treatment schools (such comparison school data are, in general, not available to grantees), the achievement and attendance data cannot be used in a quasi-experimental impact analysis.

For grantees that provide student-level achievement test data, we will tabulate treatment school grade-level means across all provided end-of-year measurements. If the student-level data also indicate which students were recipients of AmeriCorps assistance, then test score means for these students will separately be tabulated. We will also examine, using a two-level model, whether the test scores show improvements, on average, across the years of data provided.

For grantees that provide attendance and/or behavior data, we will similarly tabulate treatment school attendance (or behavior incident) percentages across the provided end-of-year measurements and examine whether there is improvement across the years.

### **Synthesis of Quantitative and Qualitative Results across Data Collection Strategies**

The sections above describe each of the study's multi-faceted data collection strategies, and how we plan to approach analyses for each distinct data collection activity. For those activities that occur once, we will describe findings as of a particular point in time, and for activities with multiple waves of data collection, we will describe observed patterns between pre- and post-interviews or surveys. Evidence obtained from any single approach to analysis has strengths and weaknesses, and by integrating multiple analytic approaches, one can have greater confidence in the results. Integrating the observations across data collection strategies, therefore, is another important feature of our analytic approach, as it will allow us to describe the findings more comprehensively, and will allow us to contextualize the findings appropriately.

In this study, we will be collecting data from different stakeholders, whose perceptions and day-to-day involvement with school turnaround efforts, and with AmeriCorps members in particular, are likely to vary. Synthesizing patterns across methods and respondents will inform our collective understanding of how school turnaround efforts take root in school communities—or not, and whether the AmeriCorps members are perceived as essential to school turnaround. For example, findings from interviews and focus groups may provide more nuanced explanations of patterns observed in close-ended survey responses. Responses to focus group questions about the factors that influence school turnaround efforts will complement a school leader survey question that asks respondents to rate the relative importance of specific school turnaround effort activities on a Likert-type scale. Should the study find educationally meaningful differences between program and comparison school leaders' survey responses, we can draw from interviews to help understand why we may be seeing such differences; conversely, should we find little or no differences between program and comparison school survey responses, interview data can help us understand why that might be so.

### **B3. Describe methods to maximize response rates and to deal with issues of non-response. The accuracy and reliability of information collected must be shown to be adequate for intended uses.**

**For collections based on sampling, a special justification must be provided for any collection that will not yield “reliable” data that can be generalized to the universe studied.**

The National Evaluation of School Turnaround AmeriCorps employs a number of strategies to maximize response rates while maintaining cost control, which are detailed in Justification Part A. To maximize response rates, the contractor will use district-specific liaisons to help communicate with schools about the importance of the survey effort. They will also identify a staff person at each school who can assist with encouraging and reminding educators to complete surveys. The contractor will have access to real-time response rates to target follow-up communications, and study team staff will communicate with potential respondents and send reminders over the course of each field period. Interviews will be conducted over the telephone and focus groups will be done in person and online. Interviews and focus groups will be scheduled at a time that is most convenient for the respondent.

We anticipate encountering a variety of types of missing data. There will be individuals who refuse to be surveyed (pre- or post-survey) and individuals who refuse some questionnaire items. We refer to these as unit nonresponse and item nonresponse, respectively. Survey nonresponse will be handled in accordance with OMB’s Standards and Guidelines for Statistical Surveys, particularly sections 1.3 and 3.2. Unit nonresponse will be handled using survey non-response adjusted sampling weights, whereas item nonresponse will be handled, as needed, using multiple imputation.

#### *Handling Unit Non-response*

Well-designed surveys may experience patterns of unit non-response that compromise the comparability of the treatment and comparison groups, potentially leading to biased estimates of the differences between the groups (IES, 2014). We plan to conduct non-response analyses by computing both overall non-response (i.e., the rate of non-response for the entire sample) and differential non-response (i.e., the difference in the rates of non-response for the treatment and comparison groups) for the principals’ and teachers’ surveys and non-response rate (only for the treatment group) for the grantee staff and AmeriCorps members’ surveys. We will implement propensity stratification non-response adjustments to the sampling base weights to create non-response adjusted weights<sup>9</sup> (separately for treatment and comparison groups for the principals’ and teachers’ surveys). The construction of base weights for teachers was described above, and the base weights for the other kinds of surveys are equal to 1 (since we are aiming to take a census of AmeriCorps members and principals). Note that applying nonresponse adjustments accounts only for differences in observed characteristics between respondents and non-respondents; it cannot rule out the possibility of non-response bias being introduced in estimates due to differences in unobserved characteristics.

In an appendix to the final report, we will include more detailed information on non-response for each survey. In the detailed tables, we will include sample sizes by treatment and comparison group, the overall response rate, and response rates by treatment and comparison group. Reporting these rates provides context for the reader, as relatively large rates of differential non-response can lead to imbalances between the treatment and comparison groups.

---

<sup>9</sup> Valliant, R., Dever, J. A. & Kreuter, F. (2013). Practical tools for designing and weighting survey samples. pp. 329-334. This involves creating weighting classes based on estimated propensity scores (i.e., on the estimated probabilities of responding to the survey given covariate values, computed using logistic regression). We will examine the adequacy of the nonresponse adjustments by checking for within-class balance on the covariates used to compute the estimated propensity scores. Also, note that the covariates used to compute the propensity score estimates must have known values for all individuals asked to complete a survey, respondents and nonrespondents. Ideally, we will have access to some known individual-level covariates (e.g., gender, age) in addition to known school-level covariates in order to generate useful propensity scores.



### *Handling Item Non-response*

We anticipate that unit nonresponse rather than item nonresponse will be the main cause of missing data. However, if, on a set of preselected key items, we discover a large amount of missing data (say, more than 20% missing), we will use multiple imputation to account for the missing item data. To impute missing data, we will use an off-the-shelf suite of multiple imputation commands, either SAS or STATA. Both sets of commands implement all three steps of the multiple imputation process: imputation, completed-data analysis, and pooling.

#### **B4. Describe any tests of procedures or methods to be undertaken.**

Survey pilot tests were completed in May 2014 with the following individuals: teachers/counselors (n=8), school leaders (n=8), AmeriCorps members (n=8), and grantees (n=6). Interview pilot tests were also completed with the following individuals: teachers/counselors (n=8), school leaders (n=5), AmeriCorps members (n=9), parents (n=3), and grantees (n=6). Two pilot test focus groups were conducted with teachers and school leaders, and one pilot test focus group was conducted with AmeriCorps members. Data collection instruments were revised in response to pilot test results.

#### **B5. Provide the name and telephone number of individuals consulted on statistical aspects of the design, and the name of the agency unit, contractor(s), grantee(s), or other person(s) who will actually collect and/or analyze the information for the agency.**

Abt Associates has been contracted to administer the survey, conduct interviews and focus groups, and analyze the data. The key staff assigned to this project are:

- Jennifer Bagnell Stuart, Project Director
- Dr. Beth Gamse, Principal Investigator
- Dr. Edward Bein, Director of Analysis

CNCS has collaborated in the design phase and will continue to collaborate on all stages of the project. The individual at CNCS assigned to this project is:

- Diana Epstein, Ph.D., Senior Research Analyst, 202-606-7564

In addition, the Project Officer for CNCS is Diana Epstein, Ph.D., Senior Research Analyst, 202-606-7564.

#### **B6. Other information – Study Limitations**

Limitations of the current study and its design include the following:

- This is an outcomes and implementation study designed to understand how AmeriCorps members contribute to schools' capacity to successfully implement their respective turnaround model and its perceived effectiveness in affecting key turnaround outcomes, including student academic achievement. However, this study is not designed to determine impacts of the program on student performance, and its results will not support conclusions about the causal relationship between the intervention and the outcomes of students served by the program. Findings will most likely be considered as meeting preliminary or moderate evidence standards, but not strong evidence standards, as defined in the AmeriCorps Notice of Funding Opportunity (NOFO).<sup>10</sup>

---

<sup>10</sup> Corporation for National and Community Service. AmeriCorps State and National *Notice of Federal Funding Opportunity*. FY 2014, pp. 27-28.

- Though this study is not an impact evaluation of academic outcomes, it uses a quasi-experimental design (QED), in that it will employ Mahalanobis matching to pair grantee (treatment schools) with other struggling schools that lack a meaningful AmeriCorps presence (comparison schools). The matching will be based on pre-intervention school characteristics such as students' reading and math achievement and demographics. These matched comparison schools are thus intended to serve as a meaningful foil for the treatment schools, and to provide insight into whether and how the intervention has influenced treatment schools' staff perceptions, relative to the perceptions of comparison schools' staff. Note that some treatment schools have a quite limited pool of potential comparison schools to which they may be matched, and therefore some of these matched pairs may not be very close in their profiles of school characteristics.<sup>11</sup> Also, because staff at treatment schools are not blind to the fact that they are at treatment schools, they may feel some pressure to express positive views of how their school is functioning. Finally, because schools were not randomly assigned to receive AmeriCorps members as part of their respective school turnaround efforts, there could be pre-existing differences between the treatment and the (yet-to-be-selected) comparison schools on unobserved characteristics. We cannot assume the two types of schools will be equivalent except for the presence of AmeriCorps members. As a result of these limitations, the study may not be able to support rigorous causal inferences that attribute any observed differences in the views of treatment school staff to the School Turnaround Initiative.
- In consideration of respondent burden, a substantial amount of data collection is cross-sectional and not longitudinal (e.g., teacher surveys and interviews). As noted earlier, the pre-surveys will be administered at the beginning of the first year of the evaluation, which is the beginning of the second full year of the School Turnaround AmeriCorps intervention. Therefore, examinations of differential pre-to-post changes in perceptions should not be assumed to reflect true pre-intervention to post-intervention differences, as they do not address differential change from prior to the intervention to the end of the second year of the intervention.
- There are a small number of School Turnaround AmeriCorps grantees (13) and schools (62) relative to the universe of over 1,600 SIG and Priority schools across four school year cohorts (2010-2014). As noted earlier, this evaluation is focused on drawing inferences about individuals associated with this specific set of School Turnaround AmeriCorps schools rather than about all the individuals associated with some larger population of low-performing schools.
- Secondary data will be collected from grantees and their partner schools, not districts or states. Prior to use of any student outcome data, it will be necessary to determine the availability and completeness of student achievement and attendance/behavior data. As part of their grant requirement, grantees established written partnership agreements with the schools in which AmeriCorps members are serving. The partnership agreements explicitly reference schools' written commitment to "share outcome data" with the grantees. However, the agreements do not specify the type of data that are available or which data they will provide to the grantees. Thus, the type of data grantees can access through their written partnerships is currently unknown and will likely vary across grantees and schools. Identifying the data made available through these agreements will help inform Year 2 of the evaluation. Therefore, we will consider the findings from the first year of the evaluation to re-visit the question of whether collecting school-level student achievement and attendance data might be valuable, and if so, how to reformulate the research questions to incorporate

---

<sup>11</sup> We anticipate that up to a handful of treatment schools will need to be dropped because of lack of an adequate matched comparison school.

collection of these data in Year 2. These insights will also inform the possible inclusion of collecting secondary student outcome data from districts and states in Year 2 of the evaluation, including the possible need for assessing the feasibility of negotiating data sharing agreements with states/districts where there is a time lag before data are publicly released.

- This is a study of a new program that has been in operation for one year prior to the start of the evaluation. As such, it may take some time (sometimes more than a year) to refine new programs within schools to the point they are operating efficiently and effectively. Thus, study findings are expected to be most useful for formative purposes of providing early learning about what works and does not work, and informing best practice development and program improvement and refinement efforts. Year 1 findings may be less useful for making summative judgments about the potential effectiveness of the program at a more mature stage or when operating at greater scale. However, Year 2 findings are intended to provide additional evidence about the potential effectiveness of the program.

## References

Gabler, Hader, and Lahiri (1999). A model based justification of Kish's formula for design effects for weighting and clustering. *Survey Methodology*, 25(1):105–106.

Hsueh, JoAnn, Desiree Principe Alderson, Erika Lundquist, Charles Michalopoulos, Daniel Gubits, and David Fein (2012). *The Supporting Healthy Marriage Evaluation: Early Impacts on Low-Income Families*, Technical Supplement. OPRE Report 2012-27. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Institute of Education Sciences (IES). 2014. What Works Clearinghouse. What Works Clearinghouse Procedures and Standards Handbook Version 3.0. [http://ies.ed.gov/ncee/wwc/pdf/reference\\_resources/wwc\\_procedures\\_v3\\_0\\_draft\\_standards\\_handbook.pdf](http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_procedures_v3_0_draft_standards_handbook.pdf).

Izrael, D., Battaglia, M. P., & Frankel, M. R. (2009). Extreme Survey Weight Adjustment as a Component of Sample Balancing (aka Raking). In SAS Global Forum paper.

Lohr, S. L. (1999). *Sampling: Design and Analysis*. New York: Cengage Learning.

Rosenbaum, P. R. (2010). *Design of Observational Studies*. New York: Springer.

Valliant, R., Dever, J. A. & Kreuter, F. (2013). Practical tools for designing and weighting survey samples, pp. 418-422.