An Evaluation of the True North AmeriCorps Program

IMPACT REPORT

August 2022

Report Authors:
Peter Nelson, Ph.D.   –   peter@serveminnesota.org
Patrick Kaiser, M.Ed. –   patrick@serveminnesota.org

Project:
True North AmeriCorps

NATIONAL
Science & Service
COLLABORATIVE
Powerful Insights. AmeriCorps Ingenuity.

AmeriCorps   ServeMN

About the National Science & Service Collaborative

We believe partnerships between researchers, AmeriCorps programs, and communities can transform research and practice, leading to sustainable, community-driven solutions. We value a broad and inclusive definition of "collaboration" because improving societal outcomes is maximized when the tools of science, expertise of communities, and resources of AmeriCorps are deployed in a truly collaborative way.

The Center's portfolio includes projects to evaluate the impact of AmeriCorps programming, projects to advance the existing knowledge base in education, and development projects to bring new and innovative programming to communities across the nation.

https://nssc.serveminnesota.org/

NATIONAL
Science & Service
COLLABORATIVE
Powerful Insights. AmeriCorps Ingenuity.

Table of Contents

Introduction

Launched in 2003, the True North AmeriCorps program (TNAC) is an affiliate of the Duluth Area Family YMCA, serving economically disadvantaged children in the city of Duluth and three surrounding rural counties. The program has undergone significant changes in recent years. Prior to 2019, TNAC members were placed in partnering schools and after-school programs to provide reading and math support as well as general support for school engagement.

In the last grant cycle, the program shifted focus in at least two important ways. First, the program almost exclusively served students in after-school programs (e.g., daily structured programming at schools and drop-in programming at community and youth centers), shifting away from school-day partnerships. Second, the program modified activities to focus more intently on supporting students' engagement in school-related activities by way of fostering positive adult relationships. As such, the program no longer directly assessed students' reading and math achievement.

Despite promising changes to the program's theory of change, a clear need to strengthen the program for evidence-building persisted. In their 2020 evaluation, Nelson & Kaiser observed a statistically significant and negative effect of treatment assignment when comparing treated and untreated student scores on the SAYO-Staff survey distributed by the National Institute of Out-of-School Time.

Yet the results observed in the previous evaluation were illustrative of important

limitations of program infrastructure and areas for ongoing improvement. For example, although members were trained more directly in building relationships with students and a variety of training approaches were adopted to align with this focus, the SAYO-S was retained as the primary outcome of interest for the program during the 2020 evaluation. This is relevant because the nature of member activities and outcome measures needed improvement.

Theory of Change

TNAC's current Theory of Change holds that the provision of a caring adult for students will create an environment that promotes increases in school engagement, noting that those increases are associated with long-term academic success outcomes (e.g., Reschly & Christenson, 2006).

Currently, all participating students are referred to TNAC members by a representative at the site, school, or home. Participating students receive targeted sessions in which the member provides a variety of formal "check-in" supports characterized by goal-oriented problem-solving and social-emotional skill building as well as informal support in

the form of after-school activity participation and homework help. TNAC members may serve a maximum of 40 students (approximately 20-40 depending on site). The program does not currently graduate or exit students based on performance so participating students typically receive support for the full program year.

Current Evaluation

The present evaluation sought to address many of the limitations associated with previous evaluations. TNAC was instrumental in positioning the program for a revised evaluation by introducing a new assessment and approach to data collection—these two programmatic changes resulted in an assessment more aligned with member activities and a data infrastructure that better positioned the program to engage in evaluative activities.

The current evaluation focused specifically on student scores on the Devereux Student Strengths Assessment (DESSA), which is a CASEL-aligned social emotional assessment with defensible psychometric data (described in more detail within the report). The DESSA is composed of eight subscales and produces a composite score. The primary evaluation questions for the current evaluation explicitly focus on the composite score and the resulting qualitative designation of post-test DESSA scores. More formally, the evaluation was guided by two research questions:

1. Relative to students of a similar age and baseline DESSA score, to what extent do end-of-year adult ratings of students served by TNAC members differ from ratings of students not served by TNAC members?

2. To what extent does the probability of receiving a composite score in the "typical" range differ between students served by a TNAC member and a matched comparison group?

Impact Assessment

As described previously, the evaluation was restricted to the DESSA survey at TNAC partner sites. Outcomes were collected at two time points: fall (baseline) and spring for nearly all students at a given site. Regardless of assignment to support from a TNAC member, all participants received some form of adult support per the nature of the intervention setting (i.e., out of school support programs). However, students support by a TNAC member received individualized support on setting personal goals and improving SEL skills addressed by the DESSA.

DESSA ratings were obtained by either (1) AmeriCorps members serving at the site, or (2) an adult familiar with the student at the site.

Matching and Participants

Data for the current evaluation were obtained from an extant data file from the 2021-2022 program year. That file included demographic and DESSA data for all students served by True North members. To be included in the treatment group for the evaluation, students must have had a fall and spring DESSA score and have begun support from a TNAC member before December 1, 2021 and continued through April 1, 2022. Regardless of treatment provision both fall and spring DESSA scores were required to be collected from the same rater (i.e., the member or a staff member at the partner site).

The initial data file included 521 students, 282 of which were students assigned to the caseload of a TNAC member. Thus, there were 239 students

in the data file who did not receive support from a TNAC member.

After applying the aforementioned inclusion criteria, there were 262 "treated" students and 150 comparison students. To create a matched comparison group, we used propensity scores. Propensity matching is largely advocated as a valid and useful quasi-experimental method for evaluating group differences, as it is designed to balance groups across a number of pre-existing factors that could otherwise account for differences in the outcome measures (Smith, 1997; Stuart & Rubin, 2008). In the current evaluation we used the MatchIt package in R to create a matched sample (Ho, Imai, King, & Stuart, 2011). The matching process included two analytic steps. In the first step, logistic regression was used to calculate each students' propensity (or likelihood) for receiving TNAC support based on (1) baseline DESSA composite scores and (2) age. The second analytic step involved matching cases according to their propensity scores, where students in the treatment group were matched with a student who did not receive the treatment but had a similar propensity for receiving the treatment. In the present analysis we used nearest neighbor matching with replacement to pair cases based on their likelihood of participating in TNAC (Rubin, 1973). Comparison cases not matched to a treatment case were excluded from further analysis. The final sample consisted of 341 students (262 in treatment and 79 comparison students). The loss of comparison students in the matched file was directly related to differences in baseline DESSA scores

and/or age. Descriptive data for the analytic sample separated by group are included in Table 1. Students in the treatment received an average of 46 sessions from a TNAC member focused on SEL skills and an additional 51 sessions in other areas. The large standard deviation for sessions indicates substantial variability in the number of sessions students received.

All data reported on the analytic sample—including inferential analyses—reflect case weighting adjustments produced by the matching process. A series of chi-square and t-tests indicated that there were no statistically significant differences between groups in regard to any of the demographic frequencies or fall group averages outlined in Table 1.

Table 1. Fall Demographic and DESSA Data across Groups

Variables	Treatment	Comparison
Demographics		
Age	7.77 (SD = 2.10)	7.62 (SD = 1.77)
Female	44%	51%
White	67%	72%
Asian	1%	3%
Black/African American	10%	5%
Native American	4%	3%
Multi-Racial	16%	18%
Service		
Avg. SEL Sessions	45.64 (44.14)	-
Avg. "Other" Sessions	50.98 (47.28)	-
DESSA Fall Composite Scores		
DESSA Composite Score	45.43 (10.00)	45.00 (10.10)
Students Scoring in "Typical" Range	61%	60%

Measures

The DESSA is a standardized, norm-referenced behavior rating scale that assesses social-emotional competencies in kindergarten through the eighth grade. The measure is organized into eight social-emotional competency scales aligned to the CASEL framework: self-awareness, social-awareness, self-management, goal-directed behavior, relationship skills, personal responsibility, decision making, and optimistic thinking.

The DESSA can be completed by parents, teachers, or staff at schools and child-serving agencies, including after-school programs. For each of the 72 items on the measure, the rater is asked to indicate on a five-point scale how often the student engaged in each behavior over the past four weeks.

Upon completing the measure, the student receives a raw score in each of the eight competencies. Raw scores are converted to standard T-scores so that the separate scales of the DESSA can be directly compared and so an individual student's behavior can be compared to that of the other children in the standardization sample. Students also receive a Social-Emotional Composite

score which is a combination of the eight scales. The Composite score is the most reliable and valid overall indicator of strengths within the DESSA.

The DESSA categorizes each of the student's eight scale scores and the composite score into range descriptions. Students with DESSA scale T-scores 40 and below are identified as "need for instruction", meaning the student was rated as showing few behaviors associated with the particular social-emotional strength and are considered at risk for exhibiting or developing social-emotional problems. Scale scores of 41 to 59 are described as "typical". Students in this range would likely benefit from universal strategies designed to promote the social and emotional competence of all children. Finally, scores of 60 and above are considered "strengths" for that student (LeBuffe, Shapiro, & Naglieri, 2014).

Analysis Procedures

As outlined previously, there were two outcomes of interest in the current evaluation: differences in end-of-year DESSA composite scores and the probability of receiving a DESSA composite score within the "typical" range. That is, the second outcome of interest represents a slightly different interpretation of the same score of interest in the first research question. The use and interpretation of DESSA scores in the "typical" range was included

because it is a substantively meaningful criterion for intervention. That is, it is meaningful and consistent with the program's theory of change to estimate whether or not a student needs additional intervention or perceived to be functioning in the typical or strong range of social-emotional skills.

To assess the degree to which DESSA composite scores differed across groups at the end of the year, we fit a linear regression model to the data that included fall DESSA scores, student age, and a dummy-coded variable for treatment assignment. Likewise, to assess differences in rate of reaching the "typical" rating, we fit a logistic regression model to the data with the same covariates adopted for the linear regression model.

Results

Descriptive results for DESSA composite scores are displayed in Table 2. In general, the average DESSA score for treated and untreated students increased across time, from a t-score of approximately 45 at baseline to a spring t-score of 51.52 among students receiving TNAC support and 49.65 among comparison students. The slightly larger increase among treatment students was also reflected in the proportion of students receiving a composite t-score within the typical range when comparing treated (86%) and comparison (77%) students.

Table 2. Descriptive data for Fall and Spring DESSA Ratings

| | Fall | | Spring | |
	Baseline Score	Baseline Typical %	End-of-Year Score	End-of-Year Typical %
TNAC Treatment	45.43 (10.00)	61%	51.52 (10.93)	86%
Comparison	45.00 (10.10)	60%	49.65 (11.53)	77%

To further examine the degree to which changes in DESSA ratings differed as a function of TNAC service, the end-of-year DESSA t-scores were regressed on the aforementioned demographic and prior performance covariates (see Table 3). There was a statistically significant and positive association between DESSA fall composite scores and end-of-year scores (B = 0.68) as well as age and end-of-year scores (B = 0.80). There was no meaningful association observed between receiving TNAC support and end-of-year DESSA scores (p = 0.10). Results were generally similar in the logistic regression model with the exception that there was a statistically significant and positive increase in the probability of receiving a DESSA composite score in the "typical" rating category among students receiving TNAC support. For example, among age seven students who did not receive a DESSA composite score within the typical range, the probability of receiving ratings within the typical category at the end of the year increased by 18% among students receiving TNAC support.

Table 3. Linear and Logistic Regression Results

| | End of Year DESSA Composite | | Probability of Typical Rating | |
	B	SE	B	SE
Age	0.80***	0.25	0.18*	0.13
DESSA Fall Composite	0.68***	0.05	-	-
"Typical" Criterion at Fall Rating	-	-	1.93***	0.35
Treatment	1.84	1.13	0.76*	0.35

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

Conclusion and Discussion

In the current evaluation, we sought to examine the degree to which students assigned to receive TNAC support demonstrated increases in social-emotional competencies as measured by the DESSA. Descriptive analysis showed treatment students – those who received targeted support from a TNAC member – increased their DESSA composite score by an average of 6.1 T-score points while comparison students who did not receive targeted support increased their score by an average of 4.7 points. Guidance from DESSA indicates a change of 2-4 points in one program year is considered a small change while a 5-7 point change is considered medium. Following this guidance, both students receiving targeted TNAC support and students who did not receive targeted support generally made medium gains, with the TNAC supported students making slightly larger gains. However, the difference in composite scores between the two groups was not statistically significant.

The difference in the percentage of students scoring in the "Typical" range showed slightly more promising results. The percent of treatment students in the "Typical" range increased by 25 percentage points (61% to 86%) while the comparison students increased by 17 percentage points (60% to 77%). This difference was statistically significant. That is, although students in the two groups did not differ in regard to their overall composite scores, students assigned to TNAC support had a slightly larger probability of receiving ratings within the typical range at the end of the year.

Implications for Practice

TNAC appears to support student SEL skill development; however, the effects observed in the study were relatively small. With the program's goal of continuous improvement, the current evaluation results offer insight on potential program changes.

Student Selection

TNAC selects students for targeted support based on teacher and staff recommendations. After selection, the member completes the baseline DESSA on these students. This process led to 61% of students in the TNAC service group starting the year already in the "Typical" range for their social-emotional competencies. Conversely, 40% of students who were not served scored in the "need for instruction" range at baseline. Thus, a large number of students at a lower risk level received support while many higher risk students did not receive support. By using baseline DESSA scores as part of the selection criteria process, TNAC could better target their support to students with the highest need. In addition, the use of a cut score for service may facilitate other means for evaluation design (e.g., regression discontinuity).

Session Delivery

Participating students had substantial variation in the number of sessions they received. The average student received a total of 97 sessions while the range was 0 to 382 sessions. The extensive distribution indicates potential inconsistency in delivery across members and sites. Since it is reasonable to

expect more sessions would produce greater improvement in student skills, students would likely benefit from increased consistency in session delivery.

Finally, student sessions were generally evenly divided between SEL and "other" skills. Encouraging members to spend a greater percentage of sessions on SEL skills may lead to more student improvement on the DESSA measure. However, if potential student improvements from content covered during sessions focused on "other" skills are important to TNAC's theory of change, the program would benefit from collecting additional data on these skills. For example, if some sessions focus on improving student attendance, school or social connectedness, or self-efficacy the program could pursue the collection of additional data. In this case, data such as self-efficacy or perceived social connection could be obtained directly from students.

Documentation

TNAC members currently select two competencies per student to focus on during their sessions. The program year that was the focus of the current evaluation was the first year TNAC attempted to track the focus area for students, and member reported data was incomplete. Clearer and more consistent documentation on the focus areas, including the focus area of each session and the strategies used by the member, would help evaluate differences observed between particpaiting students and assist the program with more targeted continuous improvement efforts.

Limitations

The primary objective of the evaluation was to assess the effect of the TNAC program on students' social-emotional skills. In order to achieve this objective, the evaluation was designed to measure program effects using a quasi-experimental design. However, unavoidable limitations inherent in the study design and in working with sites and students constrained some aspects of the evaluation's design, implementation, and analysis.

The DESSA measure used in the analysis was completed by either the TNAC member or another adult at the site. There are limitations with any rating tool as raters may have different levels of understanding about the student or there could be variability in how the measure is completed, potentially impacting the reliability of the results.

As with most programs, there are feasibility limits to data collection. In the current evaluation, TNAC was able to collect limited information on program implementation, particularly related to the specific interventions and content used by members with each student. Having additional information in future evaluations would help determine if some activities produce a greater positive impact.

Relatedly, incomplete data on individual student activities limited the ability to use the individual competency captured by the DESSA in a targeted way. Instead, the evaluation focused on the DESSA composite score. Capturing additional details on member activities with each student and their related social-emotional competency would allow evaluators to analyze program

impact using a sub-measure that is more closely aligned with the member activities with each student.

Conclusion

During the previous grant cycle, TNAC has shown a strong commitment to program improvement initiatives. Since the 2019-2020 evaluation, the program has refined its program model, changed its primary outcome measure, and updated its data collection system. The results in the current evaluation demonstrate the impact of these changes. Rather than finding null and negative effects, the current evaluation observed patterns of growth in the desired direction and statistically significant effects on the rate at which students scored in the "Typical" range on the DESSA. Additionally, descriptive analysis, though not statistically significant, showed TNAC supported students increased their DESSA composite score by an average of 6.1 points, while comparison students increased their score by an average of 4.7 points. In both cases, the rate of improvement from baseline to post-test was at or above national norms on the DESSA.

By continuing its focus on program improvement, TNAC can build on these results and position itself for greater impact going forward.

References

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software, http://gking. harvard. edu/matchit.*

LeBuffe, P. A., Shapiro, V. B., & Naglieri, J. A. (2014). Devereux Student Strengths Assessment: A measure of Social-Emotional Competencies of Children in Kindergarten Through Eighth Grade. Aperture Education, 100 Main Street, Suite 201, Fort Mill, SC 29715.

Reschly, A. L., & Christenson, S. L. (2006). Prediction of dropout among students with mild disabilities: A case for the inclusion of student engagement variables. *Remedial and Special Education*, *27*(5), 276-292.

Rubin, D. B. (1973). The use of matched sampling and regression adjustment to remove bias in observational studies. *Biometrics*, 185-203.

Smith, H. L. (1997). 6. Matching with Multiple Controls to Estimate Treatment Effects in Observational Studies. *Sociological methodology*, *27*(1), 325-353.

Stuart, E. A., & Rubin, D. B. (2008). Best practices in quasi-experimental designs. *Best practices in quantitative methods*, 155-176.

About the Authors

Peter Nelson, Ph.D.

Director of Research and Innovation, ServeMinnesota
Peter engages with ServeMinnesota pilot programs and existing programs to ensure key principles of effective implementation and evidence-building occur. Before joining ServeMinnesota, Peter was an Assistant Professor of School Psychology at Penn State University. Formerly a high school teacher, he regularly publishes and presents research related to data-based decision making and academic intervention in the school setting. Peter serves on the editorial board for School Psychology Review, School Psychology Forum, and Assessment for Effective Intervention.

Patrick Kaiser, M.Ed.

Director of Education Evaluation, ServeMinnesota
Patrick manages evaluation and reporting for ServeMinnesota's programs with a focus on educational programming. He also supports the evaluation needs of partner organizations replicating Minnesota programs across the country. Patrick previously taught fifth grade in Savannah, GA as an AmeriCorps member and worked in education policy advocacy in Atlanta, GA. Patrick attended the University of Notre Dame where he received a B.A. in Political Science and Economics and a M.Ed. in Elementary Education.

NATIONAL
Science & Service
COLLABORATIVE
Powerful Insights. AmeriCorps Ingenuity.

AmeriCorps | ServeMN