



## **Estimating the Association between Generation Citizen's Curriculum and Student Civic Skills**

July 2013

Report prepared by:

Rebecca Casciano, Ph.D.  
Glass Frog Solutions  
rebecca@glassfrog.us  
1.888.609.3372, ext. 1  
www.glassfrog.us

The author thanks the teachers and administrators at the host schools for accommodating the study in their classrooms, as well as the students who participated in the study. Jonathan Davis, Pavithra Nagarajan, and Saba Abebe provided excellent research assistance. Direct correspondence to Rebecca Casciano at Glass Frog Solutions, P.O. Box 9, Skillman, NJ 08558 or by e-mail at [rebecca@glassfrog.us](mailto:rebecca@glassfrog.us).

**TABLE OF CONTENTS**

EXECUTIVE SUMMARY ..... ii

RESEARCH & SAMPLING DESIGN ..... 0

MEASURING CIVIC SKILLS..... 1

RESULTS ..... 4

THREATS TO INTERNAL VALIDITY ..... 6

CONCLUSION..... 7

APPENDIX A: Civic Skills Assessments ..... 8

APPENDIX B: Generation Citizen Civic Skills Assessment Rubric ..... 10

APPENDIX C: Grading the Assessments ..... 11

APPENDIX D: Descriptive statistics by sub-group. .... 12

## EXECUTIVE SUMMARY

This report presents results from an evaluation conducted on behalf of Generation Citizen (GC), a nonprofit organization that places college student volunteers in disadvantaged urban schools to implement an action civics curriculum alongside classroom teachers. The program aims to cultivate civic engagement and close the civic engagement gap in historically marginalized communities. A primary mechanism by which GC seeks to do this is by increasing students' civic skills. This study examines GC's impact on civic skills. It draws on data from a pre/post written assessment of civic skills that was administered to students in a sample of GC classrooms and students in a sample of comparison classrooms in Spring 2013.

## RESEARCH & SAMPLING DESIGN

To estimate GC's impact on student civic skills, we employed a pretest-posttest nonequivalent groups design, administering a civic skills assessment at two points in time to students participating in two GC classrooms and students in two classrooms not participating in the program. The original sampling frame for the study included all New York City classrooms using the GC program during Spring 2013. From this list, we selected two classrooms in schools large enough to have other social studies classrooms with similarly-aged students not participating in the GC program. We then identified one social studies classroom in each school to act as comparison classrooms and then administered the assessment to students in all four classrooms. The baseline civic skills assessment was administered in February 2013, prior to the start of the GC program, and the follow-up was administered in May and June 2013, after the program was completed. Student retention rates from baseline to follow-up were very low, which has implications for our confidence in the results. We discuss these implications below.

## MEASURING CIVIC SKILLS

The GC curriculum is designed to impact three civic skills: persuasive communication, critical analysis, and group collaboration. This evaluation relied exclusively on written assessments of student civic skills and therefore focused on two of these three skills—critical analysis and persuasive communication.

We created original written assessments to gauge students' critical analysis and persuasive communication skills at baseline and follow-up. Both the pre- and post-assessment presented students with a social problem of local concern—the pre-assessment described issues related to Hurricane Sandy and the post-assessment described issues related to New York City's stop and frisk policy. The assessments were graded using an original rubric that graded students along four criteria: organization, analysis, feasibility, and persuasiveness. Students were given a score of 1 ("Developing"), 2 "Proficient", or 3 ("Exemplary") on all criteria. At both waves, students tended to score somewhere between "Developing" (1 point) and "Proficient" (2 points) on each criterion, though scores rose slightly from baseline to follow-up. We also summed the scores on each criterion to create a total civic skills score. The average total civic skills score at baseline was 5.8, compared to 6.5 at follow-up.

To test whether participating in GC is associated with an increase in student civic skills over the course of their time in the program, we perform three sets of analyses:

1. We compare mean scores of GC students and non-GC students at both baseline and follow-up;
2. For each group, we compare students' scores at baseline to their own scores at follow-up; and
3. We compare *growth of scores* from baseline to follow-up between GC and non-GC students.

## RESULTS

At baseline, GC students scored significantly lower on all criteria of the civic skills assessment with the exception of Persuasiveness. On the overall civic skills measure, GC students scored significantly lower (1.5 points or roughly .7 standard deviation units) than the non-GC students.

By the time of the follow-up, however, GC students were performing slightly better than non-GC students on all of the criteria, though the differences were not statistically significant. On average, GC students experienced significant growth from baseline to follow-up on all four of the components and on the overall civic skills measure. On the overall measure of civic skills, scores rose 1.9 points (this is equal to a gain of roughly .8 standard deviation units). In contrast, scores among students in the comparison group stayed more or less constant from baseline to follow-up. In other words, though they started with high scores overall, students in non-GC classrooms experienced no growth in civic skills over the course of the semester.

The difference between baseline and follow-up scores for GC students was significantly greater than it was for non-GC students. This pattern was consistent across all of the criteria measured and for the overall civic skills score.

## FACTORS UNDERMINING CONFIDENCE IN OUR FINDINGS

The results suggest that GC students in this sample experienced significant gains in civic skills over the course of their time with the program, while students in the comparison classrooms did not, lending support to the hypothesis that participating in GC is associated with gains in civic skills. There are several factors that undermine our ability to say with certainty that GC is *causing* gains in civic skills. The GC classrooms may differ systematically from the non-GC classrooms in ways that have nothing to do with the GC curriculum. Moreover, the attrition rates were very high from baseline to follow-up, which means that the results presented in this report may not represent gains among all students who participated in the program. That said, because the results presented here are promising, we recommend replicating this study in Fall 2013, administering the baseline study in September 2013 and the follow-up study in December 2013 (when attendance is likely to be much higher). We would also recommend devoting more time in the pre-study phase to finding classrooms that are more comparable to the GC classrooms (for example, it would be better if we could use Participation in Government classrooms as the comparison classrooms).

## Estimating the Association between Generation Citizen's Curriculum and Student Civic Skills

This report presents results from an evaluation conducted on behalf of Generation Citizen (GC), a nonprofit 501(c)(3) charitable organization that places college study volunteers ("Democracy Coaches") in disadvantaged urban schools to implement an action civics curriculum alongside classroom teachers. The program aims to cultivate civic engagement and ultimately close the civic engagement gap in historically marginalized communities. A primary mechanism by which GC seeks to do this is by increasing students' civic skills, which the organization hypothesizes will increase students' ability to effect change as well as their confidence in their ability to do so. This study examines GC's impact on student civic skills. It draws on data from a pre/post written assessment of civic skills that was administered to students in a sample of GC classrooms and students in a sample of comparison classrooms in Spring 2013.

### RESEARCH & SAMPLING DESIGN

To estimate GC's impact on student civic skills, we employed a pretest-posttest nonequivalent groups design, administering a civic skills assessment at two points in time to students participating in two GC classrooms and students in two classrooms not participating in the program. This quasi-experimental design is preferable in settings where experimental methods, namely random assignment, are undesirable or unfeasible for some reason. (In this case, random assignment was not feasible because classrooms had already partnered with the program at the time the study was being designed.) A pretest-posttest nonequivalent groups design involves administering a pre-test ("baseline") and post-test ("follow-up") to two groups, one of which receives a "treatment" (in this case, the GC program) and one of which does not. Some of the limitations of this method as well as the limitations of this study in particular are described in more detail in the section titled "Threats to Internal Validity."

The original sampling frame for the study included all New York City classrooms using the GC program during Spring 2013. From this list, we selected two classrooms in schools large enough to have other social studies classrooms with similarly-aged students (i.e., students in grades 11 and 12) not participating in the GC program. We identified one social studies classroom in each school to act as comparison classrooms and then administered the assessment to students in all four classrooms. Both GC classes were Participation in Government classes, while the final non-GC classes included a U.S. History course and an Economics course. Both the comparison classrooms had some emphasis on analysis, critical thinking, and writing, though neither used an action civics curriculum like the one used in GC classrooms.

The baseline civic skills assessment was administered in February 2013, prior to the start of the GC program, and the follow-up was administered in May and June 2013, after the program was completed. All students who were present in class on the day the baseline assessment was administered took the assessment. As it was the beginning of the semester, some of the classrooms had not yet finalized their

rosters. Teachers indicated that they would be more likely to lose students than to gain students in the coming weeks, so we opted to administer the assessment to everyone in the class and then later discard the assessments of the students who transferred out. We re-administered the assessments to everyone in the selected classrooms at the end of the semester.

Student retention rates from baseline to follow-up were very low in some cases, ranging from 33% in one classroom to 79% in the classroom with the highest retention (Table 1 shows retention rates). The overall retention rate was 56%. In classrooms where participation at follow-up was very low, we returned to the classrooms again to re-administer the assessment. Teachers and school administrators reported that absenteeism is chronically high at the end of the school year. We discuss the implications of this in the “Threats to Internal Validity” section.

Table 1. Retention rates, by classroom, between baseline and follow-up administrations of the civic skills assessment.

	Baseline #	Follow-up #	Retention rate
GC Classroom 1	18	8	44%
GC Classroom 2	24	19	79%
Non-GC Classroom 3	16	10	63%
Non-GC Classroom 4	21	7	33%
Total	79	44	56%

## MEASURING CIVIC SKILLS

### *Generation Citizen's Targeted Civic Skills*

The Generation Citizen curriculum is designed to impact three civic skills: persuasive communication, critical analysis, and group collaboration.<sup>1</sup> This evaluation relied exclusively on written assessments of student civic skills and therefore focused on two of these three skills—critical analysis and persuasive communication. The third skill, group collaboration, is not measurable through a written assessment. Moreover, the program aims to impact both written and oral communication skills; our assessment only evaluates written skills.

### *Assessment*

We created original written assessments to gauge students' critical analysis and persuasive communication skills before they participated in the Generation Citizen program and after they had completed the program. Students in the comparison group completed the assessments at the same points in time. Both assessments presented students with a social problem of local concern—the

<sup>1</sup> Source: Generation Citizen Civic Skills Assessment Plan (Draft January 2011). Obtained via correspondence with Molly Mills, December 27, 2012.

baseline assessment described issues related to Hurricane Sandy and the follow-up assessment described issues related to New York City’s stop and frisk policy. The assessments asked students to create a plan to address the problem, describe the goal of their plan, and then explain the steps they would take to enact this plan. Students were also asked to explain why the problem is important to address and how their plan would successfully address the problem. Students were given 25 minutes to complete the assessments at both baseline and follow-up. Copies of the assessments are in Appendix A.

*Rubric*

To grade the assessments, we created a rubric that measures the targeted civic skills, namely critical analysis and persuasive communication. Table 2 shows how the criteria we used for grading correspond to these skills. For each criteria, we developed two or more components for assessing students’ mastery of a given skill. Students were awarded 1 (Developing), 2 (Proficient), or 3 (Exemplary) points for each criteria. In some instances, students failed to complete one or more of the sections, in which case they were awarded 0 points for the corresponding component. The full rubric can be found in Appendix B.

Table 2. Mapping Generation Citizen’s Targeted Civic Skills to Skills Measured in Assessment

<b>Generation Citizen skill</b>	<b>Assessment skill component</b>
Critical analysis	Organization
	Analysis
	Feasibility of Implementation
Persuasive written communication	Persuasiveness
Persuasive oral communication	Not measured
Group collaboration	Not measured

The Organization, Analysis, and Feasibility of Implementation criteria of the assessment rubric measure students’ critical analysis skills. The Organization criterion evaluates how well students set goals and sequence their plans for civic change. Students were awarded 1 to 3 points on each of three components: their ability to (a) create a goal that addresses the root cause of the problem and articulate a plan of action with steps that are (b) logical (i.e., steps that build upon previous steps) and (c) related directly to the final goal.

The Analysis criterion evaluates how well students critically analyze the civic issue(s) they are trying to address. Specifically, students were awarded 1 to 3 points on each of four components: their ability to (a) identify the parts of the problem they were trying to address, (b) explain how their plan would address the problem, (c) discuss how they would integrate key stakeholders and decision-makers into their plan for addressing the problem, and (d) analyze the individual, community, and systemic issues related to the problem.

The Feasibility of Implementation criterion evaluates how well students critically analyze the feasibility of implementing their plan. Students were awarded 1 to 3 points on two feasibility components: their ability to develop a plan that is (a) financially feasible and has multiple avenues to achieve success and (b) could feasibly be implemented.

The Persuasiveness criterion addresses Generation Citizen’s persuasive communication skill, evaluating how well students are able to persuade readers that the problem they are addressing is significant and their plan is best-suited for addressing that problem. Students were awarded 1 to 3 points on two components: their ability to justify in their writing (a) why the social problem is important to address and (b) why their plan would successfully address the problem.

*Grading*

To create scores for each criterion of the rubric, we averaged scores across the components. For instance, to create a score for a student’s organizational skills at baseline, we averaged his/her scores on the three components used to evaluate organizational skills. If a student left a section blank, s/he received a score of 0 on that component. Thus, each criterion has a theoretical minimum score of 0 and a theoretical maximum score of 3. We then summed the scores on each component to create a total civic skills score. Table 3 shows descriptive statistics on each component for the full sample of students at both baseline and follow-up. At both waves, students tended to score somewhere between “Developing” (1 point) and “Proficient” (2 points) on each component, though scores rose slightly from baseline to follow-up. No students received the maximum number of points (3) on any of the four components. The total civic skills score at baseline was 5.8, compared to 6.5 at follow-up. Means and standard deviations by group (GC vs. non-GC) are presented in Table D1 of Appendix D.

Table 3. Descriptive statistics for student civic skills at baseline and follow-up (n=44).

	Baseline assessment			Follow-up assessment		
	Mean	S.D.	Min, Max	Mean	S.D.	Min, Max
Organization	1.7	.6	.0, 2.7	1.8	.5	.3, 2.8
Analysis	1.4	.5	.0, 2.1	1.6	.5	.1, 2.5
Feasibility	1.5	.6	.0, 2.5	1.6	.6	.0, 2.5
Persuasiveness	1.2	.7	.0, 2.3	1.5	.5	.0, 2.5
Total	5.8	2.1	.5, 9.4	6.5	1.9	1.8, 9.7

*Analytic strategy*

This study asks whether participating in GC is associated with an increase in student civic skills over the course of their time in the program, relative to students who did not participate in the program. To answer this question, we perform three sets of analyses. First, we compare mean scores of GC students and non-GC students using two-group mean comparison t-tests on differences between baseline and follow-up scores. Second, we use two-sample mean comparison t-tests to determine whether students



in each group experienced significant growth from baseline to follow-up. Finally, we use two-group mean comparison t-tests to compare differences in *growth of scores* from baseline to follow-up between GC students and non-GC students. This final test is used to determine whether GC students experienced significantly greater (or lesser) growth than students in the comparison group.<sup>2</sup> Because of the low sample size, we use a probability cutoff of .10 as the criterion for determining statistical significance on all tests.

**RESULTS**

Table 4 shows mean scores at baseline and follow-up for both groups, as well as results from mean comparison t-tests on differences between the two groups. At baseline, GC students scored significantly lower on all criteria of the civic skills assessment with the exception of Persuasiveness (where students scored lower but we cannot rule out that this difference is due to chance). On the overall civic skills measure, GC students scored 1.5 points (or roughly .7 standard deviation units) lower than the non-GC students. By the time of the follow-up, however, GC students were performing slightly better than non-GC students on all of the criteria, though the differences were not statistically significant. We show these differences graphically in Figure 1.

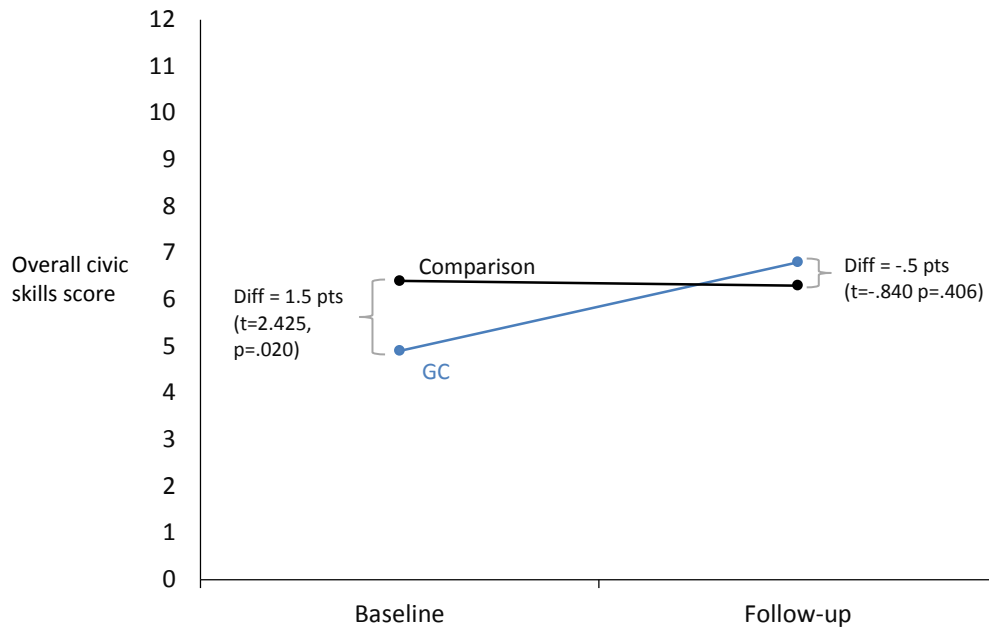
Table 4. Results from comparison of means between GC (n=18) and non-GC (n=26) students at baseline and follow-up.

	Baseline			Follow-up		
	GC	Non-GC	Sig. of t value (p value)	GC	Non-GC	Sig. of t value (p value)
Organization	1.4	1.8	.02	1.9	1.7	.38
Analysis	1.2	1.6	.01	1.7	1.6	.46
Feasibility	1.3	1.6	.09	1.6	1.5	.60
Persuasiveness	1.0	1.4	.14	1.6	1.5	.40
Overall (total)	4.9	6.4	.02	6.8	6.3	.41

Table 5 shows the difference between baseline and follow-up scores for GC and non-GC students, as well as results from the comparison of means tests (t-tests). On average, GC students experienced significant growth from baseline to follow-up on all four of the components and on the overall civic skills measure. On the overall measure of civic skills, scores rose 1.9 points (this is equal to a gain of roughly .8 standard deviation units). In contrast, scores among students in the comparison group stayed more or less constant from baseline to follow-up. In other words, though they started with high scores overall, students in non-GC classrooms experienced no growth in civic skills over the course of the semester.

<sup>2</sup> We also ran difference-in-difference regression models to identify between-group differences in scores from baseline to follow-up. The results confirm the results presented in this report; thus, we only report results from the simpler t-tests.

Figure 1. Change in student scores from baseline to follow-up, GC vs. non-GC students.



Our final test involves performing a t-test to compare differences between the two groups in *growth of scores* from baseline to follow-up. The results from the significance tests are reported in the final column of Table 5 (we only report p-values from the significance tests here; more detailed results are available upon request). The difference between baseline and follow-up scores for GC students was significantly greater than it was for non-GC students. This pattern was consistent across all of the criteria measured and for the overall civic skills score.

Table 5. Results from within- and between-groups comparison of means tests.

	GC (n=18)		Non-GC (n=26)		Sig. of between-groups difference (p-value)
	Difference in scores (Post-Pre)	Sig. of t value (p value)	Difference in scores (Post-Pre)	Sig. of t value (p value)	
Organization	+ .5	.02	- .1	.39	.01
Analysis	+ .5	.00	.0	.86	.00
Feasibility	+ .3	.06	- .1	.54	.09
Persuasiveness	+ .6	.00	+ .1	.41	.06
Overall (total)	+1.9	.00	- .1	.80	.01

## THREATS TO INTERNAL VALIDITY

A study is said to have high “internal validity” if it can show with certainty that the independent variables under study are causing changes in the dependent (or outcome) variables under study. For the present study, under conditions of high internal validity, we would be very confident that the GC curriculum—and not some other variable or combination of variables—is causing student gains in civic skills. Many studies have design features, sampling issues, and other factors that reduce internal validity. We discuss here the issues and factors that may undermine the internal validity of this study.

The results presented above suggest that GC students in this sample experienced significant gains in civic skills over the course of their time with the program, while students in the comparison classrooms did not, lending support to the hypothesis that participating in the GC curriculum is associated with gains in student civic skills. With that in mind, there are several factors that undermine our ability to say with certainty that GC is *causing* gains in civic skills.

First, the design of the study is not itself impervious to threats to internal validity. The only way to assure that students in treatment and comparison groups are comparable at the outset of a study is to randomly assign students to these groups. We could not do this here (e.g., we were working with classrooms that had already decided to use the program), so we chose an alternative method, a nonequivalent groups design, and administered the pre/post assessment to classrooms that were not equivalent, but were located within the same school, taught within the same discipline (e.g., social studies), and taught similarly-aged students (i.e., students in grades 11 and 12). Administering the pre-test also enabled us to identify any pre-existing differences between the two groups. In this study, we were able to account for the fact that the GC students actually started out with lower civic skills than students in the comparison classrooms.

The comparison classrooms also differed from the classrooms in which GC was working. Most notably, the GC classrooms were both teaching a Participation in Government curriculum, while the comparison classrooms were focused on the more traditional topics U.S. History and Economics. Thus, we cannot rule out the possibility that the Participation in Government curriculum (or perhaps the superior teaching skills of teachers in these classrooms) boosted students' civic skills and not GC.

Another serious threat to internal validity arose from high attrition rates from baseline to follow-up. As mentioned previously, students in both schools had high truancy at the end of the academic year. According to the teachers, this is typical and unlikely to be related to anything about the presence or absence of GC in the classrooms (in other words, absenteeism appears to be random and not systematic). That said, the students who remained in the classroom at the end of the year (and were therefore present to take the assessment) are likely a select subset of the overall sample of students enrolled in these classes. Specifically, they are more likely to be conscientious and therefore it is possible that they would be more likely to be both engaged in the GC program and engaged in the assessment, which could lead to an overestimation of the program's impact. That said, the same is true in the comparison classrooms, where attendance was also low at the end of the year. It is plausible that

scores among students who remained in those classrooms at the end of the year would also be higher than scores among the full class of students. In reality, it is impossible to know precisely how selective attrition impacted the students' mean scores.

Because the results presented here are promising, we recommend replicating this study in Fall 2013, administering the baseline study in September 2013 and the follow-up study in December 2013 (when attendance is likely to be much higher). We would also recommend devoting more time in the pre-study phase to finding classrooms that are more comparable to the GC classrooms (for example, it would be better if we could use Participation in Government classrooms as the comparison classrooms).

## CONCLUSION

To summarize, the results presented herein suggest the following:

1. At baseline, GC students scored significantly lower on the civic skills assessment than students in the comparison group. By the time of the follow-up study, the GC students were scoring significantly higher than they had on the baseline assessment. Scores among students in the comparison group did not change.
2. GC students experienced significantly greater growth in measured civic skills than students in the comparison group.
3. The study design, as well as selective attrition out of the study (due to chronic absenteeism), undermine our ability to say with confidence that GC is *causing* gains in student civic skills. That said, we feel that a study with better student retention rates would give us more confidence in our ability to assess the association of the GC curriculum with student civic skills. We therefore recommend replicating this study in Fall 2013.

**APPENDIX A: Civic Skills Assessments****BASELINE ASSESSMENT:*****Issue***

You live in a community that was hit very hard by Hurricane Sandy. As a result of the storm, half of the homes in your community were destroyed and most of the community's residents were displaced and had to leave the community. A few months after the storm, only a small portion of those displaced have been able to return. Few of the wealthier residents have returned, while many of the poorer residents have stayed in their homes despite unlivable conditions.

Over these past few months, your community has also encountered a dramatic increase in crime after the devastating effects of Hurricane Sandy. People have been looting stores and other businesses that have struggled to re-open for business. There has also been a rise in armed robberies and muggings of residents.

You and other frustrated residents have decided to create a plan to address one of two major problems that has impacted your community after Sandy: 1) "cleaning up" the crime problem in the community; or 2) addressing the displacement of so many residents. The plan you create will be presented at an open city council meeting.

***Task***

To address this problem, you must include the following:

- 1) Create a plan, in **2-3 paragraphs**, to address the crime problem OR displacement of residents in your community. You will present this plan to the city council and community members and must:
  - a. Describe the overall goal of your plan. (1 paragraph)
  - b. Explain the steps you would take to address the problem. Be as specific as possible. (1-2 paragraphs)
- 2) Write **2-3 paragraphs** persuading the city council and community members:
  - a. Why the storm-related crime OR displacement of residents is important to address. (1 paragraph)
  - b. Why your plan will successfully address the storm-related crime OR displacement of residents. (1-2 paragraphs)

**FOLLOW-UP ASSESSMENT****Issue**

You live in a community where the police are always stopping, questioning, and frisking teenagers. Every teenager you know has been stopped, questioned, or frisked at least twice every year since they began middle school. There have been some crime problems in your neighborhood, but you feel like teenagers are being unfairly targeted and harassed, as very few people stopped have actually committed any crimes.

Many parents have tried to confront the police to protect their children from this "harassment," but nothing has worked.

You and other frustrated residents have decided to create a plan to address the unfair amount of stopping, questioning, and frisking that has taken place in your neighborhood by the NYPD. The plan you create will be presented at an open city council meeting.

**Task**

To address this problem, you must include the following:

- 1) Create a plan, in **2-3 paragraphs**, to address the problem of stopping, questioning, and frisking by the NYPD in your community. You will present this plan to the city council and community members and must:
  - a. Describe the overall goal of your plan. (1 paragraph)
  - b. Explain the steps you would take to address the problem. Be as specific as possible. (1-2 paragraphs)
  
- 2) Write **2-3 paragraphs** persuading the city council and community members:
  - a. Why the NYPD harassment is important to address. (1 paragraph)
  - b. Why your plan will successfully address the problem. (1-2 paragraphs)

**APPENDIX B: Generation Citizen Civic Skills Assessment Rubric**

	Developing (1)	Proficient (2)	Exemplary (3)
Organization	<ul style="list-style-type: none"> <li>Student creates a <i>vague goal</i> that does not address the root cause of the problem. (1.a.)</li> </ul>	<ul style="list-style-type: none"> <li>Student creates a <i>clear, but not well-defined goal</i> that somewhat addresses the root cause of the problem.</li> </ul>	<ul style="list-style-type: none"> <li>Student creates a <i>well-defined goal</i> that addresses the root cause of the problem.</li> </ul>
	<ul style="list-style-type: none"> <li>Student creates a plan of action that <i>does not have logical steps</i> to achieve the end goal. (Logic is based on each step of the plan building on the next.) (1.b)</li> </ul>	<ul style="list-style-type: none"> <li>Student creates a plan of action that has <i>somewhat logical steps</i> to achieve their end goal. (Logic is based on each step of the plan building on the next.)</li> </ul>	<ul style="list-style-type: none"> <li>Student creates a plan of action that has <i>logical steps</i> to achieve their end goal. (Logic is based on each step of the plan building on the next.)</li> </ul>
	<ul style="list-style-type: none"> <li><i>No individual step</i> is related directly to the final goal. (1.b)</li> </ul>	<ul style="list-style-type: none"> <li><i>Some individual steps</i> are related directly to the final goal.</li> </ul>	<ul style="list-style-type: none"> <li><i>Each individual step</i> is related directly to the final goal.</li> </ul>
Analysis	<ul style="list-style-type: none"> <li>Student’s plan <i>identifies one basic component of the problem</i>. (1.b)</li> </ul>	<ul style="list-style-type: none"> <li>Student’s plan <i>identifies only one component of the problem, but in a sophisticated way</i> OR <i>identifies multiple components of the problem in a basic way</i>.</li> </ul>	<ul style="list-style-type: none"> <li>Student’s plan <i>identifies multiple components of the problem that are well developed</i>.</li> </ul>
	<ul style="list-style-type: none"> <li>Student’s plan <i>only superficially acknowledges what the problem is</i>. (1.b &amp; 2.a)</li> </ul>	<ul style="list-style-type: none"> <li>Student’s plan <i>adequately</i> addresses each component of the problem <i>beyond a superficial acknowledgment of what the problem is</i>.</li> </ul>	<ul style="list-style-type: none"> <li>Student’s plan <i>sophisticatedly</i> addresses each component of the problem beyond a superficial acknowledgment of what the problem is.</li> </ul>
	<ul style="list-style-type: none"> <li>Student’s plan <i>does not acknowledge or integrate</i> utilizing decision-makers and people in power into their plan to address the problem. (1.b)</li> </ul>	<ul style="list-style-type: none"> <li>Student’s plan <i>acknowledges, but does not integrate</i> utilizing decision-makers and people in power into their plan to address the problem.</li> </ul>	<ul style="list-style-type: none"> <li>Student’s plan <i>successfully integrates</i> utilizing decision-makers and people in power into their plan to address the problem.</li> </ul>
	<ul style="list-style-type: none"> <li><i>Student does not address or analyze</i> the individual, community, and systemic issues related to the problem and may acknowledge up to two of the three with no analysis (individual, community, and systemic issues). (1.b &amp; 2.a)</li> </ul>	<ul style="list-style-type: none"> <li>Student <i>acknowledges</i> the individual, community, and systemic issues related to the problem and may analyze up to two of the three (individual, community, and systemic issues).</li> </ul>	<ul style="list-style-type: none"> <li>Student <i>analyzes</i> the individual, community, and systemic issues related to the problem.</li> </ul>
Feasibility of Implementation	<ul style="list-style-type: none"> <li>Student comes up with a plan that is <i>prohibitively expensive</i> and <i>lacks multiple avenues</i> to achieve success. (1.b &amp; 2.b)</li> </ul>	<ul style="list-style-type: none"> <li>Student comes up with a plan that is <i>financially feasible</i> but lacks <i>multiple avenues</i> to achieve success.</li> </ul>	<ul style="list-style-type: none"> <li>Student comes up with a <i>financially feasible plan</i> that has <i>multiple avenues</i> to achieve success.</li> </ul>
	<ul style="list-style-type: none"> <li>Student’s plan <i>is not logical and could not be feasibly implemented</i>. (1.b&amp; 2.b)</li> </ul>	<ul style="list-style-type: none"> <li>Students plan <i>is logical but could not be feasibly implemented</i>.</li> </ul>	<ul style="list-style-type: none"> <li>Students plan <i>is logical and could be feasibly implemented</i>.</li> </ul>
Persuasiveness	<ul style="list-style-type: none"> <li>Student <i>does not justify</i> why the social problem is important to address. (2.b)</li> </ul>	<ul style="list-style-type: none"> <li>Student <i>adequately justifies</i> why the social problem is important to address.</li> </ul>	<ul style="list-style-type: none"> <li>Student <i>convincingly justifies</i> why the social problem is important to address.</li> </ul>
	<ul style="list-style-type: none"> <li>Student <i>does not justify</i> why their plan will successfully address the problem. (2.b)</li> </ul>	<ul style="list-style-type: none"> <li>Student <i>adequately justifies</i> why their plan will successfully address the problem.</li> </ul>	<ul style="list-style-type: none"> <li>Student <i>convincingly justifies</i> why their plan will successfully address the problem.</li> </ul>

## APPENDIX C: Grading the Assessments

Both the baseline and follow-up assessments were graded blindly, in random order, by two trained graders. The two sets of assessments were graded the same day. To maximize inter-rater reliability, the graders reviewed the rubric together and each graded three sample assessments to norm their scores. (The sample assessments were drawn from the pool of assessments that were discarded because they lacked a pre- or post-companion.) When the graders were confident that their grading methodology would yield reliable scores, Grader 1 began grading the even-numbered pre-assessments and Grader 2 began grading the odd-numbered pre-assessments. After every seven assessments, the graders would once again score a sample assessment to ensure they were still grading consistently. They would then continue. When each grader completed his/her first pile, they repeated the norming process, and then switched piles, with Grader 1 grading the odd-numbered assessments and Grader 2 grading the even-numbered assessments. When the graders finished grading the baseline assessments, they graded the follow-up assessments using the same protocol. Overall, the correlation between Grader 1's scores and Grader 2's scores was high at both baseline and follow-up across the four components. Table C1 shows a matrix with correlations between Grader 1's scores and Grader 2's scores for each of the components of the rubric.

We averaged the graders' scores together to create a single score for each skill component for each student.

Table C1. Between-grader correlation coefficients.

Component	Between-grader correlation (Pearson's r)	P-value
<i>Pre-assessment</i>		
Organization	.73	.00
Analysis	.80	.00
Feasibility	.73	.00
Persuasiveness	.85	.00
<i>Post-assessment</i>		
Organization	.79	.00
Analysis	.73	.00
Feasibility	.86	.00
Persuasiveness	.70	.00



**APPENDIX D: Descriptive statistics by sub-group.**

Table D1. Means and standard deviations at baseline and follow-up, GC vs. non-GC.

		<b>Organization</b>		<b>Analysis</b>		<b>Feasibility</b>		<b>Persuasiveness</b>		<b>Overall Civic Skills</b>	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Baseline	Non-GC	1.8	0.4	1.6	0.3	1.6	0.5	1.4	0.7	6.4	1.6
	GC	1.4	0.7	1.2	0.6	1.3	0.7	1.0	0.7	4.9	2.4
Follow-up	Non-GC	1.7	0.5	1.6	0.5	1.5	0.7	1.5	0.5	6.3	2.0
	GC	1.9	0.5	1.7	0.5	1.6	0.5	1.6	0.5	6.8	1.9