

# Comparison Group Identification for Difficult-to-Evaluate Populations: Lessons from Evaluating 1n10 LGBTQ Youth Support Services

Donald E. Friesner and Laura R. Peck  
*Arizona State University, School of Public Affairs*

The evaluation literature on Lesbian, Gay, Bisexual, Transgender, and Queer (LGBTQ) youth programs is largely absent of quantitative studies, likely in part due to the challenge of using experimental or quasi-experimental evaluation designs. This paper proposes the creative use of a national data set to overcome the problem of estimating a counterfactual for this population. In addition to discussing lessons from this approach, we describe the program and its impacts. Evidence suggests that the program under study—1n10, a local support group for LGBTQ youth in the Phoenix, AZ, metropolitan area—draws a relatively more disadvantaged group than the national average but that their levels of suicide ideation are lower. By addressing key threats to the internal validity of evaluation designs, we rule out plausible rival explanations for program impacts.

While many stressors are evident in youth in general, Lesbian, Gay, Bisexual, Transgender, and Queer (LGBTQ) youth often experience special mental health and social problems apart from their peers. Over the past several decades, the gay community in many metropolitan areas has formed non-profit groups that reach out and provide support and education programs for LGBTQ youth. Given the importance of the health and well-being of this population and the commitment of resources to serving them, it is likewise important that program evaluations yield useful results. An important challenge, however, is identifying an appropriate comparison group.

This paper discusses the evaluation of one such program in the Phoenix, Arizona metropolitan area, known as 1n10, and attempts to make headway on the challenge of identifying an appropriate comparison group. First, we describe the social and mental health problems many LGBTQ youth face. We then discuss

other existing impact evaluations that study this population. After describing the outcomes measures related to health, social and emotional well-being, we estimate the impacts on the youth participants. To do so, we use three strategies that we hope are useful in gauging program effectiveness in the absence of any preprogram data or randomization of treated individuals.

## Background

LGBTQ youth are more susceptible than their heterosexual peers to family violence and homelessness (Garofalo, Wolf, Kessel, Palfrey & DuRant, 1998; Kruks, 1991; Lock & Steiner, 1998), have higher rates of drug and alcohol abuse (DuRant, Krowchuk & Sinal, 1998; Noell & Ochs, 2001; Shifrin & Solis, 1992), and are more likely to consider and attempt suicide (D'Augelli & Hershberger, 1993; Faulkner & Cranston, 1998; Hammelman, 1993; Kourany,

1987; Ramfedi, Farrow & Deisher, 1991; Ramfedi, French, Story, Resnick & Blum, 1998; Russell & Joyner, 2001).

Suicidal tendencies appear to be a major problem in the adolescent sexual-minority population. Bell and Weinberg (1978) found that 35 percent of gay men and 38 percent of lesbians studied had attempted suicide or had serious suicide ideation before the age of 20. Saghir and Robins (1973) found that five out of six gay men who had attempted suicide had done so before age 20. In addition, studies have shown that the stigma surrounding sexual minorities affects the formation and expression of sexual identity. Some research in recent years, however, has provided conflicting information. Savin-Williams and Ream (2003) find high suicide rates among gay youth, but sexual orientation alone appears not to be a risk factor. Diamond (2003) also saw suicide and other emotional issues to be related more to adolescence than specifically correlated with sexual identity and orientation.

Because humans are social animals, we tend to be influenced by what Rosenberg (1979) labels “reflected appraisals.” Rosenberg argues that as social animals, people are deeply influenced by the attitudes of others, and over the course of time people tend to view themselves as they are viewed by others. Reflected appraisal can be especially true for adolescents who value the attitudes and judgments of parents, siblings and peers. Furthermore, Rosenberg defines self-esteem as an attitude. As such, self-esteem may be associated with other attitudes and interests of the individual (Savin-Williams, 1990).

One of the challenges of evaluating the outcomes of programs catering to LGBTQ youth is the “puzzling normality of the well-adjusted homosexual” (Aiken, West, Schwalm, Carroll & Hsuing, 1998). Many believe that the self-esteem and self-evaluation of gay men and lesbians are so fluid that they are easily altered by the changes in one’s environment or self-identification (Savin-Williams, 1990). Savin-

Williams asserts that much of the research done on gay and lesbian self-esteem assumes that in contrast to their heterosexual peers, gays and lesbians must have deficient levels of self-esteem. Furthermore, from a psychoanalytic perspective, gay males and lesbians are often portrayed as possessing self-hatred and a negative self-image. However, Weinberg and Williams (1974) found that gay males who anticipated discrimination and a negative view from heterosexuals had the lowest self-assessment. In order to raise self-esteem and self-assessment, some organizations have formed support groups for LGBTQ youth.

As long ago as the late 1970’s, the gay community has been concerned about the support and well-being of LGBTQ youth. One of the earliest support groups for LGBTQ youth was Horizons youth group in Chicago. The group first met in January 1978 with participation from about 25 youth across a diverse ethnic and social background (Herdt & Boxer, 1993). Another of the oldest gay and lesbian youth groups, the Boston Alliance of Gay, Lesbian, Bisexual and Transgender Youth (BAGLY) was founded in 1981 and provides social events and to people ages 22 and under. However, it has only been in the past 15 years that there has been a wellspring of organizations that cater to LGBTQ youth. Before that time, about the only place for LGBTQ youth to meet others like themselves was in more adult-type locations such as bars and cruising areas (DeCrescenzo, 1992). In recent years groups have been formed by and affiliated with such organizations as Parents, Families and Friends of Lesbians and Gays (PFLAG). However, not all youth support groups have long histories nor are they supported by national organizations. 1n10 is one such group. In addition, two cities, New York and Los Angeles operate LGBTQ-friendly schools for youth who dropout of or are at risk of failure in mainstream school systems (LeVay & Nonas, 1995).

Researchers have used qualitative and quantitative approaches to evaluating the

impacts of various youth support groups. In a study of the impact of belonging to a high school gay/straight alliance (GSA), Lee (2002) found that while the gay and lesbian youth experienced some level of hopelessness and despair, they also became empowered individuals through their association with the GSA. Another qualitative study on social support and LGBTQ youth emphasized the importance of an introduction to the LGBTQ community (Nesmith, Burton & Cosgrove, 1999). Quantitative studies focusing on self-esteem include Grossman and Kerner (1998) who found that self-esteem is a moderately strong predictor of emotional distress, but that satisfaction with supportiveness had no relationship to emotional distress. However, there is strong agreement on the importance of the role of self-esteem in the lives of LGBTQ youth (Savin-Williams, 1990).

## Evaluation of 1n10

The 1n10 organization, founded in 1993, is the only organization explicitly serving LGBTQ youth in the Phoenix metropolitan area. The 1n10 mission is to “inspire and enable lesbian, gay, bisexual, transgender, and questioning youth to realize their full potential and foster a sense of community.” In order to meet this mission, 1n10 operates several programs, including a youth support group, and a youth theater group in collaboration with the Phoenix chapter of the Gay, Lesbian, and Straight Education Network (GLSEN). All of 1n10’s programs are geared towards helping youth accept themselves, understand their communities, build trust, and create positive self-identities.

1n10 operates with one paid staffer and a group of volunteers, many of whom have been in the program and are now older than 22. Although the number of youth participating week to week varies, the evaluation team observed between 15 and 25 youth participating in any one week. The evaluation focused on the

implementation of the program as compared to its goals as well as its outcomes and impacts on participating youth.

The program evaluation was initially a group service-learning project for a graduate program evaluation class at Arizona State University. Because this was a class project, there were constraints in time, budget and data. In order to deal with these constraints, and still produce a useful evaluation for the client, the evaluation used a modified “shoestring evaluation” model. The shoestring evaluation (Bamberger, Rugh, Church & Fort, 2003) is typically used when a program is already underway, there is no baseline data on either the participants or a control group, and the timeline and budget are severely limited. All of these were factors in the evaluation of 1n10. The timeline was reduced due to the learning and evaluation process combined within a 15-week semester. There had been no baseline data collected on the participants, and there was no identified comparison group.

Bamberger et al. (2003) suggest the use of both quantitative and qualitative data as a way of triangulating data in order to validate information from a small sample. They then suggest forming a non-equivalent comparison group through recall and using secondary data. Identifying a useful comparison group is perhaps the greatest challenge in estimating program impacts on a difficult-to-research population such as LGBTQ youth; but a comparison group is essential in order to know what would occur in the absence of the intervention. The formation of a comparison group is much easier, even when reconstructing one from secondary data, if one is interested in general constructs such as age or race. However, forming a comparison group is much more difficult when dealing with the evaluation of LGBTQ youth, referred to academically as sexual-minority youth (Russell & Consolacion, 2003).

Ideally, a comparison group would consist of randomly selected individuals from the target

group who did not receive treatment. But this is a small population, and the idea of turning away people who express need in the interest of forming a control group may not appeal to most service providers even if it were ethical. In the case of sexual-minority youth, the defining characteristics are not easily recognizable without self-identification; and self-identification can be a problem as well, as some youth are still in the questioning stage. Due to the pressures to conform to a heterosexual concept of society and the lack of support of family and peers, many LGBTQ youth remain in the closet. It is these at-risk youth who are of interest to research.

Another possibility might be to survey a broad population of youth, hoping to identify a subset who might serve as a comparison group. The logical setting for such a survey would be in schools, but a general LGBTQ climate survey is not likely to be of interest, nor approved for distribution in the average school setting. Even if it were, gaining parental consent to administer questionnaires to a broad population of youth would be an obstacle. Because youth may not be “out” to their parents, they may be afraid to ask parents for consent.

The evaluation challenges, then, included the following: no pretest, no opportunity (or desire) for random assignment, no opportunity for broad survey to yield a comparison group. But, the evaluation goal remained to estimate program impacts. Indeed, the evaluation’s primary research question was as follows: To what extent does 1n10 affect participating youth’s outcomes related to psychological well-being, risk factors for HIV and other sexually transmitted infections, self-esteem, and suicide ideation? Qualitative research yielded useful information about program operations and participants’ and volunteers’ *perceptions* of program impact. But a quantitative estimate of the program’s impact was needed.

The evaluation team chose to use a national survey, based on the work of Stephen Russell and others, as the source for comparison group

members (Joyner & Udry, 2000; Russell & Consolacion, 2003; Russell, Franz & Driscoll, 2001; Russell & Joyner, 2001; Russell & Seif, 2002; Russell, Seif & Truong, 2001). Data come from the main in-home sample of the second wave of the National Longitudinal Study of Adolescent Health (Add Health). The study involves an in-home interview where adolescents participated only if both the parent and the youth provided written consent (Russell & Consolacion, 2003). In Wave 1 of the study, 20,745 youth from grades 7 through 12 were interviewed at home in 1994 and 1995. In Wave 2 (1996), 14,738 of the same adolescents were re-interviewed. The 1n10 evaluation used the Wave 2 public use dataset. This dataset consists of 4,834 adolescents and is nationally representative. To our knowledge, Add Health is the only public data set that contains indicators of a youth’s sexual orientation, and as such provides an opportunity for evaluators of programs targeting LGBTQ youth to identify a comparison group.

The comparison group constructed here includes those who reported at least one same-sex relationship or reported being attracted to a member of the same sex at one time. The Wave 2 survey asked two specific questions about attraction, regardless of the respondent’s sex: “Have you ever had a romantic attraction to a female?” and “Have you ever had a romantic attraction to a male?” In addition, the survey identifies the sex of participants’ partners in sexual relationships. Using these variables we identified those with same-sex attractions and those with same-sex relationships. About five percent of the public use dataset reported having a same-sex attraction or relationship. Just about one percent of the respondents reported having exclusively same-sex attractions or relationships. This compares with similar studies which have been done on the full dataset (Russell & Consolacion, 2003; Russell, Franz et al., 2001; Russell & Joyner, 2001). This selection identified 230 youth that comprise our comparison group. For the additional work

reported here, we refine this comparison group by adding age and sex cutoffs that make the comparison group more similar to the 1n10 participants.

Ours is not the first evaluation to propose use of a national sample as a comparison group. Prior work by Beals, Spicer et al (2003) used national data to compare alcohol use of two American Indian reservation populations to those from a national sample. In addition, the method has also been used in evaluating adolescent substance abuse programs (Shaw, Rosati et al, 1997), to name a few. In contrast, Bloom, et al. (2002) assert that national comparison samples are much less useful than local ones. They go on to argue that propensity score matching can be a useful tool when using national samples. But, as stated above, in our circumstance we do not have the option to build a specific, local sample; therefore, a national comparison is our sole option. Additionally, we feel this method is consistent with Patton's (1982) call for providing useful information using scarce resources.

## Outcome Measurement

To make the most out of the Add Health comparison sample, we designed an instrument using 64 specific questions from the Add Health survey. These questions from Add Health measure adolescent psychological functioning and development, self-esteem, protective measures, personality traits, community involvement, risk factors for HIV and other sexually transmitted infections, drug and alcohol use, and suicide ideation. The evaluation team administered the survey at the group's regular meeting space, the Community Church of Hope. In total, 19 surveys were completed by 1n10 participants. Because the 1n10 organization does not have a complete count of its program participants (it is a drop-in program after all), we can not compute the response rate that this represents. We estimate that these 19 survey respondents represent most of the *regular*

program participant but are more likely about half of those who ever attend the support group.<sup>1,2</sup>

The specific survey questions were transformed into six indices relating to protective factors, psychological functioning, personality measures, self-esteem, risk of sexually transmitted infections, and suicide ideation. The survey instrument and specific outcome variable construction details appear in appendices A and B, respectively. These outcomes are computed for both the 1n10 survey respondents and the Add Health comparison group. The remainder of this section discusses each index in more detail.

### *Protective Factors*

Protective factors refer to respondents' perceptions of how much adults, teachers, parents, and friends care about them as well as general family supports. The protective factor index includes responses to eight items that asked about how the respondent felt about whether others cared about them and their family resources. Individual item responses range from 0 to 4. The scale orders all elements in the same direction and simply sums them to compute a scale ranging from 0 to 32, in theory, with an observed range of 4 to 31. A higher value suggests that an individual has higher levels of protective factors, or resources, operating in his or her life.

---

<sup>1</sup> The potential bias that survey nonresponse might introduce is mixed in direction. For example, if those who show up more regularly are needier, then we would expect to underestimate program impacts if they exist. In contrast, if the more regular participants are better off in some ways than their counterparts who do not attend regularly, then we would expect to overestimate program impacts. We have no evidence to inform which source of bias might exist.

<sup>2</sup> We recognize that this is a small number of people to study, but we assert that there is value in examining this evaluation as a model for others of similar programs or on similar populations.

### *Psychological Functioning*

Psychological functioning was measured by summing nine items that asked about the love and relationship satisfaction of both the mother and father, the results of working hard, and dealing with problems. Likert responses were on a recoded five-point scale (-2 to +2), with negative numbers reflecting unfavorable responses and positive reflecting favorable responses. The resulting scale can range from -18 to +18, and we observed a range of -12 to +14 in our data. Those with a positive value on the index have favorable levels of psychological functioning, with higher values being more favorable still. The opposite is true of negative values in the scale.

### *Personality Measures*

Personality measures have to do with one's self-perception about acceptance and traits such as being shy, assertive, independent, sensitive, and so forth. In total, the measure includes 14 elements. The Likert responses were coded on a five-point scale as the psychological functioning index, reordered as needed, and then summed to create a scale ranging from -28 to +28. The observed range is -3 to +28. Those with a positive value on the index have a favorable perception of their personality traits and disposition, while those with negative values have an unfavorable perception. This measure does not necessarily capture that a respondent has a "good" personality but instead that he or she reflects favorably on his or her own personality. As such, it is a measure somewhat akin to self-esteem, but as measured via general personality traits.

### *Self-Esteem*

Self-esteem was measured by summing the responses to 13 items having to do with shaking off the blues, feeling depressed, feeling as good as others, feeling happy and hopeful about the

future, and enjoying life. Each item response ranges from 0 to 3 scale, with the overall scale range being 0 to 39. The observed range is between 9 and 39, where higher values represent higher levels of self-esteem as measured by this composite.

### *Chance of Infection*

Particularly for men in this population, the risks of infection—of HIV or other sexually transmitted infections (STIs)—may be higher than for the general population. As such, an important outcome of interest is the rate of infection. Ideally, a successful program will reduce those rates by increasing knowledge and safe sexual practices. In the absence of having direct information on the actual rates of infection, the outcome measure used here is a simple composite of two measures regarding people's expectations about being infected. Each question counts for 0 to 4 points of the overall 8-point index, where a higher value suggests greater risk.

### *Suicide Ideation*

As noted earlier, rates of suicide among LGBTQ youth are relatively higher than among straight youth. As a result, we create a composite measure of having (1) considered and (2) attempted suicide within that past 12 months. The resulting score can range from 0 to 2, and we observed scores ranging from 0 to 2. Having a score of two means that a person both considered and attempted suicide within the last year, while a score of zero means they did neither.

## Rationale for Comparison

The comparison group is a critical part of impact evaluation because it approximates the counterfactual. The counterfactual is what would occur in the absence of the treatment or intervention. In this particular instance, the

evaluation team would have no other measure of what would happen in the absence of 1n10 were it not for the comparison group drawn from the Add Health data. Having information on the outcomes of a national sample of LGB youth offers a benchmark for judging the outcomes of youth who participate in 1n10 program support services. Having a strong comparison group (or groups) can help to minimize internal validity threats, thereby increasing causal inference (Shadish, Cook & Campbell, 2002).

Internal validity refers to inferences that are made regarding the causal relationship between two measurements. In order to make such inferences, researchers must show that A precedes B in time, that A covaries with B, and that there are no other plausible explanations for the relationship of the two (Shadish et al., 2002). As do other evaluation texts, Shadish et al. (2002) describe many internal threats to validity: ambiguous temporal precedence, selection, history, maturation, regression, attrition, testing, and instrumentation. These factors are all sources of potential bias that serve to limit an evaluator's ability to detect a causal relationship between intervention and change in outcomes.

Despite having a comparison group, some of these threats remain and may limit our ability to make causal claims about the effects of 1n10. Specifically, because individuals often self-select into treatment, evaluation of that treatment must attempt to account for selection bias. In this case, the youth who volunteer to participate in 1n10 may be different from the national average gay youth in ways that might affect their outcomes. For this population, however, it is not clear whether this type of selection might suggest more favorable or less favorable outcomes. It might be that those who select to participate in 1n10 are worse off and really need the support services. In contrast, they might be the youth who are better off and able to organize themselves to attend a weekly support

group. We have no evidence to suggest which direction of bias, if any, affects our analysis.

Next, generally the internal validity threat history refers to events that occur during the evaluation time period and might affect outcomes. We face historical threats here not because of an event that took place during intervention but instead the substantially different time points at which we measure the outcomes for program participants and for the comparison group. The Add Health Wave 2 survey took place in 1996, and our evaluation took place in 2006. There might be meaningfully different environments for LGBTQ youth a decade apart such that their outcomes are simply not comparable. While enormous changes occurred between 1976 and 1996 in tolerance, acceptance and understanding regarding this population, relatively smaller changes occurred between 1996 and 2006. Nonetheless, if improvements in acceptance of LGBTQ populations occurred during this period, and if that is related to our outcomes of interest, then we may face an historical threat to internal validity.<sup>3</sup>

Maturation refers to the natural changes that occur even in the absence of treatment. Individuals age, mature, gain knowledge, and tend to become more self-confident over time. Youth can be especially susceptible to maturation, as they grow physically and mentally over short periods of time regardless of their interaction with specific treatment or services. Maturation might be a problem for us if we were tracking individuals over time, but instead we have a single outcome measure and can not know the extent to which respondents are at a developmentally appropriate stage given their own maturation.

---

<sup>3</sup> One important historical threat is that of the improvement of HIV/AIDS medication between these time points. This change may affect outcomes such as one's expectation of falling ill from the disease but it not likely to affect the self-esteem and related outcomes we examine here.

Regression artifacts arise when assessing change in extreme populations, where individuals are selected for treatment based on special needs or scores. LGBTQ youth might qualify as youth with special needs. As such, their outcomes might improve—by simple regression-to-the-mean—regardless of program efforts. If this is the case, then to have data from a comparison group of the same population is an improvement over having either only pretest measures or no comparison group at all.

One challenge in assessing impacts for the 1n10 participants is the lack of knowledge about the developmental trajectory of the youth. Many influences outside of 1n10 might affect mental health and self-esteem. It would be impossible to account for them all. Wolhwill (1973) notes that much of the behavioral science methodology was created to study static phenomena and cannot easily be applied to developmental phenomena, which is by his definition, change over time even without treatment. A further challenge is dealing with the evaluation of sexual-minority individuals. In the case of sexual-minority youth, the defining characteristics are not easily recognizable without self-identification.

The estimation of program impacts can be done in several ways. When random assignment is not possible, a quasi-experimental evaluation design can be used to infer causality. Shadish, Cook and Campbell (2002) make the distinction between designs that do a poor job of reducing the plausibility of alternative explanations for treatment effect, and those that do much better. The less difficult procedures use a pretest-posttest design where data is collected prior to the treatment, and then gathered again post-treatment. However, this is one of the weaker designs due to multiple threats to the design's internal validity, especially maturation, regression artifact and history. Due to a lack of pretest information on those who join the 1n10 youth group, and a limited amount of time to perform the study, we chose a quasi-

experimental design involving the use of a posttest-only with a non-equivalent comparison group that did not receive the services offered through 1n10.

Adding a comparison group that receives no treatment, but which is similar to the treatment group is a way of approximating the counterfactual (D'Agostino & Kwan, 1995). The posttest-only design with nonequivalent groups is often used when treatment has begun prior to evaluation and no baseline data is present (Shadish et al., 2002). There are several ways to improve the posttest-only design. One way is through the use of an independent pretest sample. The randomly formed independent sample is drawn from the same population, but may have overlapping membership. A second way of improving the design is through the use of internal comparison groups drawn from similar populations as the treatment group. A third way of improving the design is through the use of multiple nonequivalent control groups. A final improvement to the design is the use of matching or stratifying on likely correlates of the posttest. We have the opportunity to improve our design by relying on the final strategy stated here.

Care must be taken to ensure there is as much equivalence as possible between the treatment and comparison groups. Well-matched groups will help reduce bias in the estimates of the program effects. One way to gain as much equivalence as possible is through techniques of matching. Matching employs the construction of a comparison group by selecting individuals who are identical in key characteristics to those in the treatment group. To avoid bias, the variables used to match must be strongly related to the outcome on which the groups differ. In the case of 1n10 youth, these characteristics include traits that have reliability, such as sexual orientation, age and sex. Several methods can be employed to construct matching comparison groups (Marsh, 1998; Rossi, Lipsey & Freeman, 2004; Shadish et al.,



2002). These include cluster group matching, benchmark group matching, and propensity score matching. Matching methods are most useful when there are many possible outcome variables, but none have been observed, and when the size of the treatment group is smaller than the size of the comparison group (Rubin & Thomas, 1996). Our hope here is that having sex- and age-match groups of youth with similar propensities to have same-sex attractions results in a useful comparison.

Many evaluation designs for youth treatment programs in the past have been suspect due to the lack of comparison groups (Pfeiffer & Strzelecki, 1990). The lack of easily identifiable comparison groups makes evaluation of sexual-minority youth especially difficult. It is also difficult to find data that can be matched with sexual-minority youth. With adequate funding, research could be done through experimental means that would reduce much of the bias that evaluators fear. However, some studies have shown that experimental and non-experimental studies can show quite similar impacts under good conditions (Aiken et al., 1998; Cook, Shadish & Wong, 2005). Cook, Shadish and Wong (2005) point out that experiments are not always feasible in the real world, and that good non-experiments can produce comparable results when matching-based studies are carefully done and groups are closely matched. In order to reduce bias through matching, propensity score matching may be an appropriate approach to constructing comparison groups for sexual-minority youth, and other difficult-to-research populations and should be considered in future applications.

## Findings

This paper reports on two sets of findings. First, we report on our findings regarding the appropriateness of the comparison group. Then, we report on program impacts, estimated via our three comparisons. Two of the impact estimates stem from the raw comparisons of

means between the overall Add Health LGB sample and the 1n10 group and between a sex- and age-matched Add Health LBG sample and the 1n10 group. The third approach is a conventional regression-adjustment to the outcomes where we control for those characteristics that we have (sex, age, race/ethnicity) and might be associated with variation in outcomes. Discussion includes an assessment of the strengths and remaining weaknesses of this approach.

To begin, Table 1 shows the general characteristics of the 1n10 group and the two comparison groups. The “Overall” group refers to the 230 Add Health youth who have had same-sex attractions or relationships, and the “Matched” group excludes those younger than 18 and a random half of the females from the Add Health LGB sample. The Matched group includes 36 individuals. The Overall comparison group is statistically similar to the 1n10 group in race and ethnicity measures but, on average, is younger and comprised of a greater percentage of females. It is possible that both age and sex would affect program outcomes, and therefore we chose to refine the comparison sample to be more closely matched on age and sex variables. In doing so, we changed the racial and ethnic composition somewhat; but ultimately, the Matched comparison group is statistically similar in all these measurable ways to the 1n10 group. Of course we have no way of knowing the ways in which the comparison groups might differ from the 1n10 group on unmeasured traits.

Our comparison is not a no-treatment comparison because comparison sample members might have participated in some kind of support group, though we do not what kind, if any, or the extent to which any services might resemble 1n10. Instead, our comparison is between 1n10 participants and a national sample of individuals similar on key characteristics, most importantly their sexual-minority status.

Table 1  
Comparison of Treatment and Comparison Group Characteristics

Characteristics	1n10 Group	Comparison Overall	Difference	Comparison Matched	Difference
Age (in years)	18.5	16.2	2.3*	18.3	0.2
Sex (percent male)	68.4	46.2	21.7***	69.4	-0.9
Race (percent):					
White	68.4	65.5	-2.9	59.9	-8.5
African American	15.8	27.2	-11.4	31.7	-15.9
American Indian	5.3	5.5	-0.2	12.9	-7.7
Asian	0.0	0.8	-0.9	2.1	-2.1
Ethnicity (percent):					
Hispanic	10.5	13.9	-3.3	16.3	-5.8

Notes: Race and ethnicity are not mutually exclusive; totals may not add to 100 percent because of the exclusion of “other” races from the analysis.

\*\*\* statistically significant at  $p < 0.01$

\*\* statistically significant at  $p < 0.05$

\* statistically significant at  $p < 0.10$

With these two comparison groups in hand, we turn to examining program impacts by comparing the mean outcomes of the 1n10 group to the mean outcomes of the comparison cases. Table 2 reports these outcomes and impacts for the Overall comparison, and Table 3 reports them for the Matched comparison. We use a *t*-test of the differences between group means to determine whether the differences are statistically significantly distinguishable from zero. Given our small sample sizes, we would not expect to detect small impacts. Analysis reveals that three of the six indices of outcome are statistically significantly different between the treatment and comparison groups, regardless of the comparison. Because the differences between the two comparisons are not large, we will discuss only the results from

the Matched comparison since it includes individuals who, on measured characteristics (as well as same-sex attraction/relationships), are quite similar to the 1n10 group. Of the statistically significant impacts, two are unfavorable and the third is favorable. Specifically, the 1n10 group shows lower protective factors and self-esteem scores than the Add Health national average scores. Regarding protective factors, the 1n10 group scores 20.3 points (on the 32-point scale), and the Matched comparison group scores 24.0 points. This 3.8 point difference represents a 15.6 percent relative impact. Similarly, the 1n10 group scores 28.2 points (of 39 possible) on the self-esteem index, while the comparison group scored 32.7 points. This 4.5 point difference represents a 13.7 percent impact.

Table 2  
Outcomes and Impacts, Based on Overall Add Health LGB Comparison Group

Outcome Measures	1n10 Group	Comparison Overall	Difference (Impact)	Percent Impact
Protective Factors (range = 0 to 32)	20.3	23.1	-2.8	-12.3**
Psychological Functioning (range = -18 to +18)	4.1	4.7	-0.6	-12.5
Personality Factors (range = -28 to +28)	14.2	12.5	1.7	13.3
Self-Esteem (range = 0 to 39)	28.2	31.4	-3.2	-10.2**
Chance of Infection (range = 0 to 8)	5.4	6.1	-0.7	-11.3
Suicide Ideation (range = 0 to 2)	0.5	1.4	-0.9	-62.5***

\*\*\* statistically significant at  $p < 0.01$

\*\* statistically significant at  $p < 0.05$

\* statistically significant at  $p < 0.10$

Table 3  
Outcomes and Impacts, Based on Matched Add Health LGB Comparison Group

Outcome Measures	1n10 Group	Comparison Matched	Difference (Impact)	Percent Impact
Protective Factors (range = 0 to 32)	20.3	24.0	-3.8	-15.6**
Psychological Functioning (range = -18 to +18)	4.1	5.4	-1.2	-23.3
Personality Factors (range = -28 to +28)	14.2	12.2	2.0	16.5
Self-Esteem (range = 0 to 39)	28.2	32.7	-4.5	-13.7**
Chance of Infection (range = 0 to 8)	5.4	5.9	-0.5	-8.0
Suicide Ideation (range = 0 to 2)	0.5	1.4	-0.8	-61.2***

\*\*\* statistically significant at  $p < 0.01$

\*\* statistically significant at  $p < 0.05$

\* statistically significant at  $p < 0.10$

Despite these less favorable outcomes, the 1n10 group shows a much lower rate of suicide ideation than the comparison group. On the two-point scale, the 1n10 participants average 0.5 points and the Matched comparison group averages 1.4 points. This is a substantial relative difference, representing 61.2 percent lower likelihood of reporting having considered or attempted suicide in the past year.

The third test of impacts we present here is a regression-adjusted one. Table 4 shows the results of that analysis where the outcome serves as the dependent measure, and a binary

treatment indicator along with sex, age and race/ethnicity are the explanatory variables. The resulting equation intercept is interpreted as the regression-adjusted outcome for the comparison group, and the coefficient on the treatment indicator is interpreted as the impact of the treatment. Although the values change somewhat from the raw comparisons of means (Tables 2 and 3), the relative size of the effects among the three statistically significant outcomes is comparable. That is, protective factors and self-esteem show negative impacts while suicide ideation shows a favorable impact.

Table 4  
Outcomes and Impacts, based on Regression Adjustments

Outcome Measures	1n10 Group	Comparison group	Difference (Impact)	Percent Impact
Protective Factors (range = 0 to 32)	18.6	22.0	-3.4	-15.3**
Psychological Functioning (range = -18 to +18)	1.9	3.0	-1.1	-37.8
Personality Factors (range = -28 to +28)	12.7	11.7	2.0	8.6
Self-Esteem (range = 0 to 39)	28.2	31.5	-3.3	-10.5***
Chance of Infection (range = 0 to 8)	6.2	6.8	-0.6	-9.2
Suicide Ideation (range = 0 to 2)	2.2	2.8	-0.6	-21.6***

\*\*\* statistically significant at  $p < 0.01$

\*\* statistically significant at  $p < 0.05$

\* statistically significant at  $p < 0.10$

## Discussion and Conclusion

Together, what do these findings suggest? Because we have identified both favorable and unfavorable impacts, it appears that we do not face a unidirectional historical bias. If any historical bias exists, it would have to be quite sophisticated to influence protective factors and self-esteem in one direction while influencing suicide ideation in the other. If selection is biasing our results, it might actually increase confidence in the observation that 1n10 participants may be experiencing benefits from the program. Because protective factors and self-esteem are lower among the 1n10 group, it may be that those individuals with more severe problems select to seek support. In turn, while their self-esteem may not be at higher levels than a national comparison, indeed they are much better off in terms of their likelihood to contemplate or attempt suicide. This combination of findings is suggestive of an overall favorable effect of support services for LGBTQ youth, at least of the kind that 1n10 provides. That said, because of our small sample size, we urge caution in placing too much weight on these results; further, generalizability is certainly limited because of our specific, local sample.

Nevertheless, and regardless of the utility of the substantive findings from this research, we hope that the methodological findings are of

use. We have proposed a way to use national data to extract a comparison group for an otherwise difficult-to-match population. Lessons from our approach may be useful for future evaluators who have a difficult-to-match population but who seek a comparison, in that our use of a national data set yielded interesting information, and occurred in the context of a “shoestring evaluation.”

Difficult-to-evaluate populations present evaluators with interesting choices in evaluation techniques. There is no one way to evaluate such populations as sexual-minority youth given that they often remain hidden within the general population. However, the challenges for the evaluator should not restrict the evaluation of social programs for these populations. The use of both quantitative and qualitative data will help to triangulate findings in different quasi-experimental designs. The goal is to conduct evaluations that are credible and reliable, and which meet the needs of stakeholders given the conditions and constraints present.

One of the approaches to matching that we did not use is that of propensity scores. Propensity score matching is a useful tool for combining a number of covariates into a single balanced score. Optimal matching is likely to reduce bias found in many matching procedures, and it might be appropriate to implement here in the next phase of research. Matching can be very useful in evaluations of

difficult-to-research populations where a non-equivalent comparison group must be used in order to approximate the counterfactual and therefore estimate program impacts.

While vast research shows that sexual-minority youth are at greater risk for serious emotional health problems as compared to their heterosexual peers, comparatively little evaluation literature examines these difficult-to-research populations. The lack of evaluation literature may be due to the difficulties such evaluations present. Lack of evaluation research requires LGBTQ youth programs to design useful programs with little information about outcomes of such programs. Lacking information also makes funding for such programs difficult to obtain. Funders may require some type of impact evaluation in order to continue funding programs. With limited funds available, funding agencies are apt to put their money where positive impact can be proven. As a result, the evaluation of LGBTQ youth programs is important. Furthermore, if impacts can be generalized to other LGBTQ youth populations, then useful programs may be created in other parts of the country.

We recognize that the use of a national dataset for a comparison group has limitations, particularly when the data are not gathered simultaneously with the treatment group. There has been much change in the visibility of LGBTQ youth in the past several years, including Gay-Straight Alliances in many high schools, gay proms, and an overall increase in tolerance. However, the complexity of our findings suggests that the Add Health sample at least provides a good baseline comparison on the outcomes we have identified. Results were comparable for the comparison to the overall LGB sample drawn from Add Health and the more specific age- and sex-matched subsample. These two raw comparisons were validated by regression-adjusted impact estimates. In sum, evidence is suggestive that these support services work in reducing suicide ideation

among youth, even those who may be more at risk than their national average counterparts.

## Acknowledgements

We would like to thank Tonya Adamski, Lisa McGregor-Mirghani, Jenny Schultz, and Christine Getzler Vaughan, as well as Amy Beck and Toby Urvater of 1n10, Inc., for partnering in the evaluation. We also appreciate the input of panelists and participants at the 2006 American Evaluation Association National Conference, Portland, OR, and of two anonymous reviewers.

## References

- Aiken, L. S., West, S. G., Schwalm, D. E., Carroll, J., & Hsuing, S. (1998). Comparison of a randomized and two quasi-experimental designs in a single outcome evaluation: Efficacy of a university-level remedial writing program. *Evaluation Review*, 22(4), 207-244.
- Bamberger, M., Rugh, J., Church, M., & Fort, L. (2003). Shoestring evaluation: Designing impact evaluations under budget, time and data constraints. *American Journal of Evaluation*, 25(1), 5-37.
- Beals, J., Spicer, P., Mitchell, C., Novins, D., Manson, S., & the AI-SUPERPFP Team. (2003). Racial disparities in alcohol use: Comparison of two American Indian reservation populations with national Data. *American Journal of Public Health*, 93 (10), 1683-1685.
- Bell, A., & Weinberg, M. (1978). *Homosexualities: A study of diversity among men and women*. New York: Simon and Shuster.
- Bloom, H., Michalopoulos, C., Hill, C., & Ying, L. (2002). Can nonexperimental comparison group methods match the findings from a random assignment evaluation of mandatory welfare-to-work programs?, Working Paper, Manpower Demonstration Research Corporation.

- Cochran, W. G., & Chambers, S. P. (1965). The planning of observational studies of human populations. *Journal of the Royal Statistical Society. Series A*, 128 (2), 234-266.
- Cook, T. D., Shadish, W. R., & Wong, V. C. (2005). Within study comparisons of experiments and non-experiments: Can they help decide on evaluation policy? (pp. 1-46): French Economic Society Meeting on Program Evaluation.
- D'Agostino, R. B., & Kwan, H. (1995). Measuring effectiveness: What to expect without a randomized control group. *Medical Care*, 33(Supplement), AS95-AS105.
- D'Agostino, R. B., & Rubin, D. B. (2000). Estimating and using propensity scores with partially missing data. *Journal of the American Statistical Association*, 95(451), 749-759.
- D'Augelli, A., & Hershberger, S. (1993). Lesbian, gay, and bisexual youth in community settings: Personal challenges and mental health problems. *American Journal of Community Psychology*, 21, 421-448.
- DeCrescenzo, T. (1992). The brave new world of gay and lesbian youth. In B. Berzon (Ed.), *Positively Gay*. Berkeley, CA: Celestial Arts.
- DuRant, R. H., Krowchuk, D., & Sinal, S. (1998). Victimization, use of violence, and drug use at school among male adolescents who engage in same-sex sexual behavior. *Journal of Pediatrics*, 133(1), 113-118.
- Faulkner, A. H., & Cranston, K. (1998). Correlates of same-sex sexual behavior in a random sample of Massachusetts high school students. *American Journal of Public Health*, 88(2), 262-266.
- Garofalo, R., Wolf, R. C., Kessel, S., Palfrey, S. J., & DuRant, R. H. (1998). The association between health-risk behaviors and sexual orientation among a school-based sample of adolescents. *Pediatrics*, 101, 895-902.
- Grossman, A. H., & Kerner, M. S. (1998). Self Esteem and supportiveness as predictors of emotional distress in gay male and lesbian youth. *Journal of Homosexuality*, 35(2), 25-39
- Hammelman, T. (1993). Gay and lesbian youth: Contributing factors to serious attempts or considerations of suicide. *Journal of Gay and Lesbian Psychotherapy*, 2(1), 77-89.
- Herd, G., & Boxer, A. (1993). *Children of Horizons*. Boston: Beacon Press.
- Joyner, K., & Udry, J. R. (2000). You don't bring me anything but down: Adolescent romantic relationships and depression. *Journal of Health and Social Behavior*, 41(December), 369-391.
- Kourany, R. (1987). Suicide among homosexual adolescents. *Journal of Homosexuality*, 13, 111-117.
- Kruks, G. (1991). Gay and lesbian homeless street youth: Special issues and concerns. *Journal of Adolescent Health*, 12, 515-518.
- Lee, C. (2002). The impact of belonging to a high school gay/straight alliance. *The High School Journal*, Feb/Mar 2002.
- LeVay, S., & Nonas, E. (1995). *City of Friends: A portrait of the gay and lesbian community in America*. Cambridge, MA: The MIT Press.
- Lock, J., & Steiner, H. (1998). Gay, Lesbian, and Bisexual Youth Risks for Emotional, Physical, and Social Problems: Results from a community-based survey. *Journal of the American Academy of Child and Adolescent Psychiatry*, 38(3), 297-304.
- Marsh, H. W. (1998). Simulation study of non-equivalent group-matching and regression-discontinuity designs: Evaluation of gifted and talented programs. *Journal of Experimental Education*, 66, 163-192.
- Nesmith, A., Burton, D., & Cosgrove, T. (1999). Gay, lesbian, and bisexual youth and young adults: Social support in their own words. *Journal of Homosexuality*, 37(1), 95-108.
- Noell, J., & Ochs, L. (2001). Relationship of Sexual Orientation to Substance Abuse, Suicidal Ideation, Suicide Attempts, and Other Factors in a Population of Homeless Adolescents. *Journal of Adolescent Health*, 29(1), 31-36.
- Patton, M. C. (1982). *Practical Evaluation*. Newbury Park, CA: Sage.

- Peck, L. R. (2003). Subgroup Analysis in Social Experiments: Measuring Program Impacts Based on Post-Treatment Choice. *American Journal of Evaluation*, 24(2), 157-187.
- Pfeiffer, S., & Strzelecki, B. (1990). Inpatient Psychiatric Treatment of Children and Adolescents: A Review of Outcome Studies. *Journal of the American Academy of Child and Adolescent Psychiatry*, 29, 847-853.
- Ramfedi, G., Farrow, J., & Deisher, R. (1991). Risk factors for attempted suicide in gay and bisexual youth. *Pediatrics*, 87, 869-875.
- Ramfedi, G., French, S., Story, M., Resnick, M., & Blum, R. (1998). The relationship between suicide risk and sexual orientation: Results of a population-based study. *American Journal of Public Health*, 88(1), 57-60.
- Rosenbaum, P. R., & Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, 70(1), 41-55.
- Rosenbaum, P. R., & Rubin, D. B. (1984). Reducing Bias in Observational Studies Using Subclassification on the Propensity Score. *Journal of the American Statistical Association*, 79(387), 516-524.
- Rosenberg, M. (1979). *Conceiving the self*. New York: Basic Books.
- Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation: A Systematic Approach*. Thousand Oaks, CA: Sage.
- Rubin, D. B., & Thomas, N. (1996). Matching Using Estimated Propensity Scores: Relating Theory to Practice. *Biometrics*, 52, 249-264.
- Russell, S. T., & Consolacion, T. B. (2003). Adolescent Romance and Emotional Health in the United States: Beyond Binaries. *Journal of Clinical and Adolescent Psychology* 32(4), 499-508.
- Russell, S. T., Franz, B. T., & Driscoll, A. K. (2001). Same-sex Romantic Attraction and Experiences of Violence in Adolescence. *American Journal of Public Health*, 91(6), 903-906.
- Russell, S. T., & Joyner, K. (2001). Adolescent Sexual Orientation and Suicide Risk: Evidence from a national study. *American Journal of Public Health*, 91(8), 1276-1281.
- Russell, S. T., & Seif, H. (2002). Bisexual female adolescents: A critical review of past research, and results from a national study. *Journal of Bisexuality*, 2, 73-94.
- Russell, S. T., Seif, H., & Truong, N. L. (2001). School Outcomes of Sexual Minority Youth in the United States: Evidence from a national study. *Journal of Adolescence*, 24(1), 111-127.
- Saghir, M., & Robins, E. (1973). *Male and female homosexuality: A comprehensive investigation*. Baltimore: Williams and Wilkins.
- Savin-Williams, R. (1990). *Gay and lesbian youth: Expressions of identity*. New York: Hemisphere Publishing Corp.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.
- Shaw, R., Rosati, M., Salzman, P., Coles, C., & McGeary, C. (1997). Effects of adolescent ATOD behaviors and attitudes of a 5-year community partnership. *Evaluation and Program Planning*, 20 (3), 307-313.
- Shifrin, F., & Solis, M. (1992). Chemical dependency in gay and lesbian youth. *Journal of Chemical Dependency Treatment*, 5(1), 67-76.
- Smith, H. L. (1997). Matching with Multiple Controls to Estimate Treatment Effects in Observational Studies. *Sociological Methodology*, 27, 325-353.
- Smith, J. A., & Todd, P. E. (2001). Reconciling Conflicting Evidence on the Performance of Propensity-Score Matching Methods. *The American Economic Review*, 91(2), 112-118.
- Ury, H. K. (1975). Efficiency of Case-Control Studies with Multiple Controls per Case: Continuous or Dichotomous Data. *Biometrics*, 31, 643-649.
- Weinberg, M., & Williams, C. (1974). *Male homosexuals: Their problems and adaptations*. New York: Penguin.
- Wohlwill, J. (1973). *The Study of Behavioral Development*. New York: Academic Press.

## Appendix A: 1n10 Survey Instrument

Thank you for agreeing to complete this important survey for 1n10 youth.

This survey is completely voluntary, confidential, and anonymous. Your responses will not be associated with your name, and your participation in 1n10 will not be affected by whether you complete this survey or your specific answers. Please answer as honestly as possible. Your responses will be compiled with others and will help evaluate the 1n10 program.

Answer each question with the option that fits best you and your situation.

### Section I. General Questions

1. What is your age? \_\_\_\_\_ years
2. What is your race? (mark the box that most closely applies)
  - White
  - Black or African American
  - American Indian or Native American
  - Asian or Pacific Islander
  - Hispanic or Latino origin
  - Other
3. What is your biological sex?
  - Male    Female
4. Are you Transgendered?
  - Yes    No    I don't know/I'm not sure
5. Are you currently attending school?
  - Yes    No

### Section II.

Below is a list of statements. Choose the answer that best describes *how often* each of the following was true *during the past week*. Mark only one box per statement.

---

During the past week...	Most all the time	A lot of the time	Some- times	Never or rarely
-------------------------	----------------------	----------------------	----------------	--------------------

---



6.	You were bothered by things that usually don't bother you.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7.	You felt that you could not shake off the blues, even with help from your friends.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8.	You felt that you were just as good as other people.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9.	You felt depressed.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10.	You felt hopeful about the future.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11.	You thought you had been a failure.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12.	You felt fearful.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13.	You were happy.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14.	You felt lonely.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15.	People were unfriendly to you.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16.	You enjoyed life.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17.	You felt sad.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18.	You felt that people disliked you.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19.	You felt life was not worth living.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**Section III.**

Below is a list of statements. Choose the answer that best describes how strongly you agree or disagree with each statement. Mark only one box per statement.

		Strongly agree	Agree	Neither agree or disagree	Disagree	Strongly disagree
20.	Most of the time my mother/guardian is warm and loving towards me.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
21.	Overall you are satisfied with your relationship with your mother/ guardian.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

		Strongly agree	Agree	Neither agree or disagree	Disagree	Strongly disagree
22.	Your mother/guardian usually knows what's going on in your life	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
23.	Most of the time my father/ guardian is warm and loving towards me.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
24.	Overall you are satisfied with your relationship with your father/ guardian.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
25.	When you get what you want, it's usually because you worked hard for it.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
26.	You usually go out of your way to avoid having to deal with problems in your life.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
27.	Difficult problems make you very upset.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
28.	When making decisions, you usually go with your "gut feelings" without thinking too much about the consequences of each alternative.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
29.	You have a lot of good qualities.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
30.	You have a lot to be proud of.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
31.	You like yourself just the way you are.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
32.	You feel like you are doing everything just about right.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
33.	You feel socially accepted.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
34.	You feel loved and wanted.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
35.	You like to take risks.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
36.	You are independent.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
37.	You are shy.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
38.	You are assertive.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

	Strongly agree	Agree	Neither agree or disagree	Disagree	Strongly disagree
39. You are sensitive to other people's feelings.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
40. You are emotional.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
41. You can pretty much determine what will happen in your life.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
42. You live your life without much thought for the future.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
43. You are quite knowledgeable about how to use a condom correctly.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**Section IV.**

The following questions relate to sexually transmitted infections (STI's) including HIV and AIDS. Please rate the following questions regarding the risk you feel they pose in your life.

	Almost certain	A good chance	A 50-50 chance	Some Chance	Almost no chance
44. What do you think your chances are of getting AIDS?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
45. What do you think your chances are of getting another sexually transmitted disease, such as gonorrhea or genital herpes?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**Section V.**

Please answer yes or no to the following questions about sexually transmitted diseases and suicide.

	Yes	No
46. Have you ever been told by a doctor or a nurse that you had any of the following: chlamydia, syphilis, gonorrhea, HIV or AIDS, genital herpes, genital warts, trichomoniasis, syphilis, hepatitis B, bacterial vaginosis, or non-gonococcal vaginitis?	<input type="checkbox"/>	<input type="checkbox"/>

47. In the past 12 months have you been tested for a sexually transmitted disease?
48. During the past 12 months, did you ever seriously think about committing suicide?
49. During the past 12 months how many times did you actually attempt suicide?
50. Have any of your friends tried to kill themselves during the past 12 months?
51. → If yes to #50: Have any of them succeeded?

**Section VI.**

Please select the best answer to the following question about voting and community involvement.

- |   | Yes                      | No                       | NA                       |
|---|--------------------------|--------------------------|--------------------------|
| 52. Are you registered to vote? (If you are <18, please mark NA)                              | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 53. Did you vote in the most recent presidential election? (If you were <18, please mark NA)  | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 54. During the last 12 months did you perform any unpaid volunteer or community service work? | <input type="checkbox"/> | <input type="checkbox"/> |                          |

**NA refers to not applicable**

55. → If yes to #54: Which of the following types of organization(s) have you been involved with in your volunteer or community service work in the last 12 months? (mark all that apply)

- Youth organizations, such as little league or scouts
- Service organizations, such as Big Brothers or Big Sisters
- Political clubs or organizations
- Solidarity or ethnic support groups, such as the NAACP
- Church or church related groups (not including worship services)
- Community centers, neighborhood improvement associations, or social action groups.

**Section VII.**

The following questions relate to alcohol and drug use. Please write in the answer that best applies to your situation.

56. During the past 12 months, on how many days did you drink alcohol?

→ \_\_\_\_\_ (enter number)

57. Think of all the times you have had a drink during the past 12 months. How many drinks did you usually have each time?

→ \_\_\_\_\_ (enter number)

58. During the past 30 days, how many times have you used marijuana?

→ \_\_\_\_\_ (enter number)

59. During the past 30 days, how many times have you used cocaine?

→ \_\_\_\_\_ (enter number)

60. During the past 30 days, how many times have you used inhalants? (such as glue or solvents)

→ \_\_\_\_\_ (enter number)

61. During the past 30 day days, how many times did you use any of these types of illegal drugs – LSD, PCP, ecstasy, mushrooms, speed, ice, heroin, or pills without a prescription.

→ \_\_\_\_\_ (enter number)

**Section VIII.**

Below is a list of questions. Choose the answer that best describes how your answer. Mark only one box per question.

---

	Very much	Quite a bit	Very little	Not at all
62. How much do you feel that adults care about you?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
63. How much do you feel that your teachers care about you?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
64. How much do you feel that your parents care about you?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
65. How much do you feel that your friends care about you?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
66. How much do you feel that people in your family understand you?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

67. How much do you feel that you want to leave home?
68. How much do you feel that you and your family have fun together?
69. How much do you feel that your family pays attention to you?
-

## Appendix B: Index Construction Details

### **Protective Factors Index** (0-4 points per element)

- How much do you feel that adults care about you?
- How much do you feel that your teachers care about you?
- How much do you feel that your parents care about you?
- How much do you feel that your friends care about you?
- How much do you feel that people in your family understand you?
- How much do you feel that you want to leave home? (reversed)
- How much do you feel that you and your family have fun together?
- How much do you feel that your family pays attention to you?

*Response options and coding:*

not at all=0, very little=1, somewhat=2, quite a bit=3, very much=4

### **Psychological Functioning** (-2 to +2 point per element)

- Most of the time my mother is warm and loving towards me
- Overall you are satisfied with your relationship with your mother
- Your mother usually knows what's going on in your life
- Most of the time my father is warm and loving towards me
- Overall you are satisfied with your relationship with your father
- When you get what you want, it's usually because you worked hard for it
- You usually go out of your way to avoid having to deal with problems in your life
- Difficult problems make you very upset (reversed)
- When making decisions, you usually go with your "gut" feelings without thinking too much about the consequences of each alternative

*Response options and coding:*

strongly disagree=-2, disagree=-1, neither agree or disagree=0, agree=1, strongly agree=2

### **Personality Measures** (-2 to +2 point per element)

- You have a lot of good qualities
- You have a lot to be proud of
- You like yourself just the way you are
- You feel like you are doing everything just about right
- You feel socially accepted
- You feel loved and wanted
- You like to take risks
- You are independent
- You are shy (reversed)
- You are assertive
- You are sensitive to other people's feelings
- You are emotional (reversed)
- You can pretty much determine what will happen in your life
- You live your life without much thought for the future (reversed)

*Response options and coding:*

strongly disagree=-2, disagree=-1, neither agree or disagree=0, agree=1, strongly agree=2

**Self-Esteem Index** (0-3 points each element)

You felt that you could not shake off the blues, even with help from your friends (reversed)

You felt that you were just as good as other people

You felt depressed (reversed)

You felt hopeful about the future

You thought you had been a failure (reversed)

You felt fearful (reversed)

You were happy

You felt lonely (reversed)

People were unfriendly to you (reversed)

You enjoyed life

You felt sad (reversed)

You felt that people disliked you (reversed)

You felt life was not worth living (reversed)

*Response options and coding:*

never=0, not often=1, some of the time=2, most of the time=3

**Chance of Infection** (0-4 points per element)

What do you think your chances are of getting AIDS?

What do you think your chances are of getting another STI?

*Response options and coding:*

none at all=1, very low=1, low=2, high=3, very high=4

**Suicide Ideation** (0 or 1 point per element)

Have you considered suicide in the past 12 months?

Have you attempted suicide in the past 12 months?

*Response options and coding:*

no=0, yes=1