Sampling weights and clustering procedures

- Data are grouped both within age/gender group and within youth center.
  - For the four age/gender groups, data were reported separately for each category.
  - Youth Center was treated as a fixed effect.
  - Because differential response rates were observed within age/gender groups across youth centers (by approximately 10% or greater), data were weighted for the final analyses (i.e., comparing baseline to endline results) to account for these differences.
  - Data were not weighted for the initial baseline only analyses for simplicity. However, as final analyses have now been conducted, sample weight variables have been added to the baseline dataset for future use.
  - To calculate the sampling weights, we calculated a sampling weight for each age/gender group at each youth center at each time point (4 groups x 9 centers x 2 time points).
    - For each calculation, we determined the following:
      - Total number of youth in age/gender group enumerated
      - Total enumerated who agreed to share contact details
      - Total number selected
      - Total number completed interviews (based on age/gender category listed at enumeration)
  - Sampling weights for each age/gender group at each youth center at each timepoint were calculated as follows: number completed interviews/number enumerated.
    - Of note, there are instances of individuals who switched age/gender categories from the time of enumeration to the time of survey administration (e.g., someone was enumerated as a 14-year-old boy but was 15 years old at the time of survey administration). To calculate the sampling weights, individuals who switched age/gender categories were accounted for in the category in which they were enumerated, not the category they were at the time of data collection. For analysis, however, individuals were analyzed accordingly to their age/gender category at the time of survey administration.
  - Additionally, two sets of response rates were calculated for each age/gender group at each youth center at each time point. Although these were calculated separately for each group, in the manuscript, aggregated responses rates will be reported. The response rates per group will be calculated as follows:
    - 1^{ST} LEVEL: Number sharing contact details/number enumerated
    - 2^{nd} LEVEL: Number with completed interviews/number selected
      - Of note, to calculate this response rate, for simplicity we included the number of interviews completed using the age/gender group assigned at enumeration (versus the time of survey administration) as this prevented instances of having a response rate >1.