

**NSNI Cleaning Criteria for NCHA Data  
February 2009**

**Rationale:** NSNI cleans the NCHA data files specifically to assist the schools in **EVALUATION** of the alcohol use and perception variables each year and across years. Our cleaning thus has three purposes:

1. **Assure that we know the age and sex of each respondent, since age and sex are important determinants of alcohol use;**
2. **Assure that each respondent has given valid answers to the key questions on alcohol use and perception**
3. **Assure that each respondent has reason to identify with a normative message (so, for example, we eliminate international students from our data)**

There are many other uses for the NCHA data, including design of messages for a social norms campaign (whether that campaign is focused on alcohol use or other health issues).

At NSNI, we maintain the original data files sent by each school, as well as the “cleaned” version which we use for our work with each school.

Each school should also maintain the original NCHA data file, for uses other than evaluation of the alcohol use and perception variables.

**The following CASES are deleted (i.e., the entire record is deleted, reducing the sample size):**

**1. those missing values on any of the following variables:**

- age (q45)
- sex (q46)
- days using alcohol (q9d)
- perception of typical student's days using alcohol (q10d)
- hours you drank last time partied/socialized (q12)
- number of drinks last time partied/socialized (q13)
- Perception of number of drinks typical student had last time partied/socialized (q15)

**2. those with unacceptable values (defined below) on any of the following variables:**

- q12, hours alcohol, greater than 13 eliminated
- q13, hour many drinks, greater than 20 eliminated
- q14, occasions last 2 weeks, greater than 14 eliminated
- eBAC greater than or equal to .3 eliminated
- q15, typical student drinks last time partied socialized, greater than 20 eliminated
- international students (q52)
- values 7 (adult special) & 8 (other) for year in school (q49)

Although this reduces the sample size, it greatly improves the validity of the data.