

TEEN HEALTH AND TECHNOLOGY PHASE FOUR NATIONAL ONLINE SURVEY: METHODOLOGY REPORT

INTRODUCTION

Survey Description and Objectives

Harris Interactive Inc. conducted the *Online Benefits and Risks* study on behalf of Internet Solutions for Kids and funded by the National Institutes of Health. This study examines the benefits and risks of the online world to teens. Being connected to the Internet provides users with seemingly endless possibilities, from developing relationships and connecting to others to getting support on a relevant problem or topic. By the same token, there are also potential risks of being engaged in the online world, including solicitation and online predators. Of particular interest is to understand the impact of the benefits and risks associated with being online for gay, lesbian, bisexual and transgendered (LGBT) youth.

The overall research objective is to better understand the perceptions and experiences of youth, particularly LGBT youth, as it relates to the benefits and risks of being online.

Specifically, research objectives are to:

- Identify benefits of Internet use for LGBT youth that may be similar or unique as compared to non-LGBT youth,
- Identify threats to positive youth development posed by the Internet for LGBT and non-LGBT youth, and
- Identify risk and protective factors at the individual and contextual level that influence observed differences and therefore need to be taken into account when designing prevention and Internet safety programs for LGBT and non-LGBT youth.

The research was conducted in two phases: qualitative research was conducted as phase 1; quantitative research was conducted as phase 2.

This report documents the methodology for the quantitative phase of this study, which was conducted from August 4, 2010 to January 17, 2011.

Survey Method

The study was conducted from August 4, 2010 to January 17, 2011. A sample of 5,907 U.S. 13-18 year olds were surveyed online. Sample was obtained from two sources: 1) the Harris Poll Online (HPOL) opt-in panel (n=3,989 respondents); and 2) through referrals from GLSEN (Gay, Lesbian

and Straight Education Network) (n=1,918 respondents). The median survey length was 23 minutes for HPOL respondents and 34 minutes for GLSEN respondents.

Project Responsibility and Acknowledgments

The Harris team responsible for the survey included Dana Markow, Ph.D., Vice-President, Robyn Bell Dickson, Research Director, and Helen Lee, Project Researcher. Internet Solutions for Kids had the primary responsibility of the questionnaire design. GLSEN (Gay, Lesbian and Straight Education Network) also contributed to the questionnaire development. Dr. Michele Ybarra (Internet Solutions for Kids), Josephine Korchmaros (Internet Solutions for Kids), Kim Mitchell (Internet Solutions for Kids), Joseph Kosciw (GLSEN), and Emily Greytak (GLSEN) worked with Harris Interactive Inc. to provide support and guidance in crafting the final questionnaire.

Public Release of Survey Findings

All Harris Interactive Inc. surveys are designed to comply with the code and standards of the Council of American Survey Research Organizations (CASRO) and the code of the National Council of Public Polls (NCPP). Because data from the survey may be released to the public, release must stipulate that the complete report is also available.

METHODOLOGY

This methodology report describes the sampling and interviewing procedures used by Harris Interactive Inc., for the quantitative phase of the *Online Benefits and Risks* study.

The survey questionnaire was self-administered online by means of the Internet. Qualified respondents were defined as:

- U.S. residents;
- Ages 13 to 18;
- In 5th grade or above; and
- Consented to participate in the survey

A total of 5,907 U.S. residents ages 13 to 18 were surveyed online. Sample was obtained from two sources: 1) the Harris Poll Online (HPOL) opt-in panel; and 2) referrals from the Gay, Lesbian and Straight Education Network (GLSEN). Of the 5,907 respondents, 3,989 were recruited through HPOL and 1,918 were recruited through GLSEN.

Sample Selection

Sample from Harris Poll Online (HPOL)

Respondents were invited through password protected email invitations to participate in a survey about their online experiences. Invitations for the HPOL panel were emailed to a stratified random sample of U.S. residents among four groups:

- 13 to 18 year olds
- Adults with a 13 to 17 year old in their household
- Adults with a child under 18 in their household
- A general population of adults

In the cases where parents or other adults received the email invitation, the invitation noted that the survey was intended for a 13 to 18 year old in the household and asked the adult to forward the survey link to the teen.

The HPOL panel has been recruited through hundreds of sources using diverse recruitment methods in order to minimize selection bias, including:

- Co-registration offers on partner websites
- Targeted emails sent by online partners to their audience
- Graphical and text banner placements on partner websites
- Refer-a-friend program
- Client supplied sample opt-ins
- Trade show presentations

- Targeted postal mail invitations
- TV advertisements
- Telephone recruitment of targeted populations

Sample from GLSEN (Gay, Lesbian and Straight Education Network) Referrals

Due to the interest in examining the online experiences of LGBT youth, an oversample of LGBT teenagers was surveyed through a public (non-password protected) link. This oversample was recruited through GLSEN’s referral efforts. GLSEN recruited most respondents through the following two methods:

- Emails sent with the survey link to their distribution list
- Publicizing the survey through an ad on Facebook

Sample Disposition

The following tables represent the disposition of the sample for the survey.

Harris Poll Online (HPOL) Sample Disposition

Invitations

Panelists sent invitations	514,744
Panelists whose invitations and/or reminders bounced back	18,433
Panelists sent invitations excluding bouncebacks	496,311

Interviews

Total number of respondents <i>qualified, non-qualified, quota met, suspended interviews</i>	35,627
Qualified respondents in final sample	3,989
Over-quota qualified respondents	878
Qualified respondents removed from final sample after field ¹	76
Suspended interviews	4,759
Non-qualified respondents	25,925

Response rate for this study among HPOL sample was 7.2% based on the following calculation:

¹ Upon review of in-survey quality measures (such as straightligning, length of interview and illogical responses), ISK and Harris jointly identified respondents who appeared to be fraudulent. All of those identified, were removed from the sample.

(# qualified respondents + # non-qualified respondents + # quota met + # suspended respondents)

Total number of invitations sent - # invitations that bounced back as undeliverable
GLSEN Sample Disposition

Interviews

Total number of respondents <i>qualified, non-qualified, quota met, suspended interviews</i>	4,035
Qualified respondents in final sample	1,918
Over-quota qualified respondents	-
Qualified respondents removed from final sample after field	5
Suspended interviews	1,818
Non-qualified respondents	294

Control of the HPOL Sample

To maintain the reliability and integrity in the HPOL sample, the following procedures were used:

- Password protection. Each invitation contained a password-protected link to the survey that was uniquely assigned to that email address. Password protection ensures that a respondent completes the survey only one time.
- Reminder invitations. To improve overall response rates, up to three rounds of reminder invitations were mailed after the initial invitation to those respondents who had not yet participated in the survey.
- “Instant Results” of selected survey findings. To improve overall response rates, respondents were invited to access results to pre-determined, selected questions after completing the survey.
- HIPointsSM. HPOL panel members (age 13 and older, and not recruited through parent or other adult in their household) are enrolled in the HIPoints rewards program in which respondents earn points for completing surveys. These points can be redeemed for a variety of merchandise and gift certificates. Only respondents who were 13 to 18 years old and recruited directly from Harris (not through a parent or other adult) were able to receive HIPoints.

Online Interviewing Procedures

Interviews were conducted using a self-administered online survey via Harris' proprietary, web-assisted interviewing software. The Harris Online interviewing system permits online data entry by the respondents. Online questionnaires are programmed into the system with the following checks:

1. Question and response series
2. Skip patterns
3. Question rotation
4. Range checks
5. Mathematical checks
6. Consistency checks
7. Special edit procedures

For mandatory questions with pre-coded responses, the system only permits answers within a specified range; for example, if a question has three possible answer choices ("Agree," "Disagree," "Not Sure," "Do not want to answer"), the system will accept only one response from these choices.

Survey Security and Programming

For this study, Harris Interactive carried out the following measures to enhance survey security:

- Utilized a daughter window (pop-up) without browser buttons for both HPOL and the public survey. This removed the browser buttons (back, forward, print, etc.) and only allowed respondents to navigate through the survey using the embedded survey buttons or links. The right-click mouse functionality was also disabled in order to discourage respondents from copying and pasting the survey contents from the screen.
- Respondents recruited through HPOL and GLSEN were able to pause and then resume taking the survey at a later time.

Editing and Cleaning the Data

The data processing staff performs machine edits and additional cleaning for the entire data set. Harris edit programs act as a verification of the skip instructions and other data checks that are written into the program. The edit programs list any errors by case and type. These are then resolved by personnel who inspect the original file and make appropriate corrections. Complete records are kept of all such procedures.

Since personally identifiable information (PII) was not collected in this study, it was not possible to check for repeat respondents through the public link based on PII during the backend data cleaning process. However, in order to reduce this risk and others, Harris implemented a variety of measures to detect fraudulent respondents. The measures that were executed include: examining length of time for respondent to take the survey, cookie detection, straightlining, incomplete responses at an open-end question, and illogical responses.

Weighting the Data

Two groups were weighted in this study: a general population of teens and LGBT teens. The following steps were taken to achieve these weights:

1. The HPOL general population sample was first weighted to known demographics of 13 to 18 year olds based on the 2009 Current Population Survey (CPS). These demographics include: biological sex, age, race/ethnicity, parents' highest level of education, school location, and U.S. region.
2. From the weighted general population HPOL sample, a demographic profile was created for LGBT teens (those who identified as Lesbian, Gay, Bisexual, Transgender and/or Queer). The profile was applied to the GLSEN LGBT teens and included the following demographics: gender/biological sex combined classification, age, race/ethnicity, parents' highest level of education, school location, U.S. region, and sexual orientation² (by biological sex).³
3. After it was determined that the demographic weighting alone did not bring GLSEN and HPOL LGBT teens into alignment, a propensity score was created to adjust for behavioral and attitudinal differences between the two groups. This propensity model is based on the following questions:
 - Born-again or evangelical Christian (Q1345)
 - Participation in after-school programs or activities run or organized by school (Q805/1)
 - Attends Gay/Straight Alliance (GSA) meetings (Q825)
 - Has parents that know what websites respondent goes to (Q1020/4)
 - Has been bullied or harassed because of being or perceived as being gay, lesbian or bisexual (Q1505/1)
 - Attends programs or groups for LGBTQ people outside of school (Q2028)
 - Uses Internet to connect with other LGBTQ people (Q2030)
 - Are "out" to their parents (their parents know respondent is LGBTQ) (Q2035/4)
 - Amount of time spent online using a computer at home (Q2405)

Similar to the demographic weight, the propensity score weights GLSEN data to HPOL data.

² GLSEN transgender teens were not weighted to HPOL on orientation.

³ There are 204 LGBT teens in the HPOL sample (out of a total of 3,989 HPOL respondents) and 1,775 LGBT teens in the GLSEN sample (out of a total of 1,918 GLSEN respondents). Non-LGBT teens from the GLSEN sample are not included in the gen pop or LGBT weight.

4. After completing the demographic and propensity weighting, a postweight was applied so that GLSEN and HPOL LGBT each account for 50% of the combined total LGBT population.
5. Because the GLSEN LGBT sample was more than 8 times the size of the HPOL LGBT sample, the resulting weight factors after the postweight were larger than desired. In order to reduce the size of the weights resulting from the postweight, as a final step, the GLSEN LGBT sample was randomly split in two, and within each group, respondents with the lowest weights (i.e., those overrepresented in the data) were deselected, leaving approximately 600 GLSEN respondents in each group (no 13 year olds or transgender respondents were deselected). Steps 2 through 4 were repeated with each of the reduced GLSEN samples, once again, weighting GLSEN to HPOL.

The data file includes four separate weight variables: 1) the HPOL general population weight, 2) a LGBT weight including all GLSEN and HPOL LGBT respondents, 3) an LGBT weight including a selection of GLSEN LGBT respondents plus all HPOL LGBT respondents and 4) an LGBT weight including a second selection of GLSEN LGBT respondents plus all HPOL LGBT respondents.

Reliability of Survey Percentages

All sample surveys and polls, whether or not they use probability sampling, are subject to multiple sources of error which are most often not possible to quantify or estimate, including sampling error, coverage error, error associated with nonresponse, error associated with question wording and response options, and post-survey weighting and adjustments. Therefore, Harris Interactive avoids the words "margin of error" as they are misleading. All that can be calculated are different possible sampling errors with different probabilities for pure, unweighted, random samples with 100% response rates. These are only theoretical because no published polls come close to this ideal.

Respondents for this survey were selected from among those who have agreed to participate in Harris Interactive surveys or referred by someone who has agreed to participate in Harris Interactive surveys and among those who responded to the GLSEN invitation. Because of how the sample was obtained, no estimates of theoretical sampling error can be calculated.

Non-Sampling Error

Sampling error is only one way in which survey findings may vary from the findings that would result from interviewing every member of the relevant population. Survey research is susceptible to human and mechanical errors as well, such as data handling errors. However, the procedures used by Harris Interactive, including the data processing quality assurance process described earlier, keep these types of errors to a minimum.