Fraud Detection Methods in Online Surveys

Authors: Michelle Gosney, Jamie Atkisson, Alyssa Haskins, Edward Johnson, Matthew Deihl

Abstract

Fraud detection is an important tool for gathering quality data in online market research. As "fraudsters" in online surveys become more and more savvy, it is crucial to seek out and use methods that will effectively identify them and reduce their ability to infiltrate survey data.

Growing evidence points to problems with "character misrepresentation" (Bell et al., 2022), multiple completions by the same respondent (Teitcher et al., 2015), and insincere respondents tending to select positive answer choices (Kennedy et al., 2020).

Research Questions

In-survey Quality (ISQ) tools in online survey research help to identify fraudulent activity. One such tool is an algorithm used to judge whether Open-End responses are likely produced by bots (RobotFlag) vs. human respondents. Adding to this bot detection, a newly designed "Honeypot Bot" flag has been developed to further help identify bots taking our surveys.

Another tool used, "Respondent Instruction", flags whether a respondent is paying attention. This type of question has typically been presented towards the end of the survey, but we wonder if "false-positives" may be flagged at this point in the survey simply due to respondent fatigue.

In-survey Quality Tests included in this study:

- Minimum LOI
- Respondent Instruction
- Illogical Choice Combination
- ISQ Honeypot
- Real Answer Score
- Digital Fingerprinting
  - Duplicate respondent check
  - Fraud score check

Manual methods considered in

- Age/Zip/Demo comparisons
- Numeric outliers

This research seeks to answer the following questions:

1) Does our newly designed "HoneyPot Bot" question improve bot detection compared to just using the established "RobotFlag" for Real Answer (RA)?

2) Does location of the "Respondent Instruction" influence the occurrence of failure at this question incorrectly?

Methodology

To test our questions, we programmed our survey to randomly assign (1) whether respondents receive the HoneyPot Bot question or not and (2) where the Respondent Instruction question is presented – whether earlier in the survey or later.

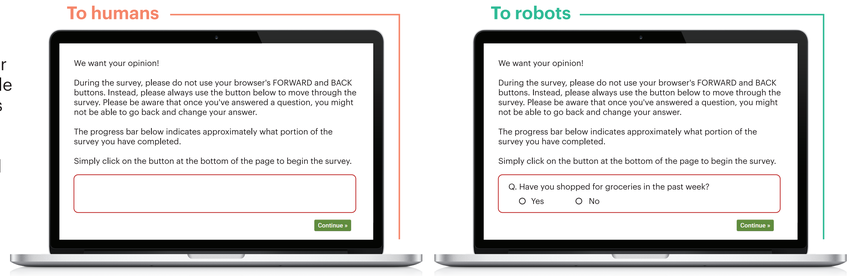
The survey data were collected online by The Harris Poll from 8,524 US adults age 18+ across ten different online opt-in sample provider blends. Sample sizes for each blend ranged from n=850 to n=858. The 15-minute survey fielded August 16-30, 2022.

To answer our research questions regarding fraud detection, we analyzed unweighted data among all completed survey interviews (including qualified respondents, non-qualified respondents, and overquotas, totaling n=11,332), we measured the effects of use or non-use of HoneyPot Bot and/or rotated placement of RI tools on fraudulent activity.

HoneyPot (HP) Bot Detector Example

Not Seen by Humans

Description: We added a new HoneyPot bot detector question which is not visible on screen to humans but is visible to online bots. The bots see this question and answer it thinking it's a real survey item, while true respondents are unable to see/answer it.



HoneyPot (HP) Bot Detector – Frequency of Bots

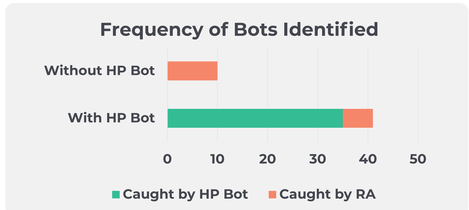
HoneyPot (HP) caught 35 "robots" and immediately termed them – saving from "manual" review.

Therefore:

- Data "with HP" does not include the 35 auto-termed (assumes then, that these are actual respondents, with fewer bots)
- Data "without HP" means we don't know if they would have been "bots" caught immediately and auto-termed

In conjunction with the established RobotFlag which identifies if an open-end question was answered as a result of a programmatic insertion, we were able to increase the number of bots detected from 10 (not using the ISQ Bot) to 41 (using the HoneyPot Bot).

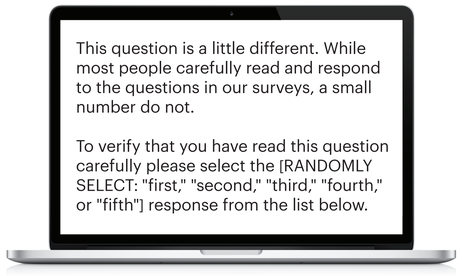
Even with HP in use, 6 bots passed the test but were identified later in the survey by the RA RobotFlag.



Without HP, directionally more respondents displayed "fraudulent" activity.

Respondent Instruction Test

Examples - Randomly Assigned



[Display]

- O Strongly Agree
- O Somewhat Agree
- O Neither Agree or Disagree
- O Somewhat Disagree
- O Strongly Disagree

[OR]

- O Extremely Important
- O Very Important
- O Somewhat Important
- O Not at all important
- O Don't Know

[OR]

- O Very Good
- O Good
- O Fair
- O Poor
- O Very Poor

[OR]

- O 0-5
- O 6-10
- O 11-25
- O 26-50
- O 50+

Alternating Placement of Respondent Instruction (RI) Test

Regardless of where RI was placed within the survey (earlier vs later), there was no significant change to the number of qualified respondents who failed the Respondent Instruction task (332 when asked earlier in the survey, and 307 when asked later).

This implies that survey fatigue does not result in more falsely flagged cases to be fraud.

Additionally, there is no significant change to the frequency of cases failing RI who were finalized as Fraudulent. This suggests that researchers should not be concerned with the placement of RI items of their survey.

Implications

HoneyPot Findings

Our research on the use of a newly designed bot detector, HoneyPot, shows a significant increase of bots being detected in our sample! However, since a small portion of cases passed the HoneyPot test and were caught by the following Open-End bot detector, we recommend adding the HoneyPot tool to the collection of fraud checks rather than replacing other bot detectors.

Respondent Instruction Findings

Concerns on whether our Respondent Instruction question could result in false-positives due to survey fatigue was found to be undeserved. There was no evidence to show an increase of false-positive cases (Qualified cases who failed the RI) being flagged as fraud later in the survey, nor any change in false-negatives (Fraudulent cases who passed the RI).

Limitations

HoneyPot

Currently the HoneyPot bot detector is treated as an automatic disqualification. Therefore, we were unable to monitor for the number of bots identified by failing both the HoneyPot and the Open-End bot detector. Additionally, we were unable to monitor for potential false positives, though our programmers have guaranteed no human

Respondent Instruction Findings

While we were able to find any evidence that RI placement effected the frequency of fraud detection and therefore unaffected by survey fatigue; we did not monitor for potential bias in qualified survey data. Such bias could result in the jarring experience of the thematic change in questions mid-survey, but also the annoyance of being 'tested' for honesty.

Future Research

This research could be extended into examining the placement and presentation of other fraud indicators, such as "Real Answer" open-end responses, and the effects of auto-terming respondents upon violating fraud indicators vs adjudicating them after the survey is completed.

While this research has largely been focused on effects of fraud indicators on frequency of identified fraud and its influence on qualified data, we could also examine fraud indicators' influence on user experience and potential bias resulting from being overly monitored by fraud detectors. It is important to use fraud detection to keep your data clean of bots and click-farm responses but is it worth creating a negative survey experience for your real respondents.

Learn More

Michelle Gosney
Director
Media Communications Research
michelle.Gosney@harrispoll.com

Matt Deihl
Research Manager
Media Communications Research
matt.Deihl@harrispoll.com

Use the following QR code to download this poster as a pdf and view other posters presented by The Harris Poll.

Bell, A.M., & Gift, T., Fraud in Online Surveys: Evidence of a Nonprobability, Subpopulation Sample, Journal of Experimental Political Science, Cambridge University Press, May, 2022.
Teitcher, J.E.F., Bockting, W.O., Bauermeister, J.A., Hofer, C.J., Miner, M.H., Klitzman, R.L., Detecting, Preventing, and Responding to "Fraudsters" in Internet Research: Ethics and Tradoffs. J Law Med Ethics, 2015 Spring 43(1), 116-133.
Kennedy, C., Hatley, N., Lau, A., Mercer, A., Keeter, S., Ferno, J.,& Asare-Marfo, D., "Assessing the Risks to Online Polls From Bogus Respondents" Pew Research Center, February, 2020.