infosurv

YOU WILL LEARN

• 3 Proven Methodologies to Avert Disengaged or Fraudulent Respondents

• How to Identify & Eliminate Common Data Manipulation Tactics

• The Cost of Not Following Data Integrity Best Practices

Protecting Survey Data Integrity:
Best Practices for Online Research

Author
Chris Gwinner, Sr. Project Manager, Infosurv

Executive Summary

Online research is quickly becoming the preferred methodology to companies worldwide, and for good reason. An online survey environment provides faster turn-around time, lower cost, and in most cases, higher data quality when compared to traditional survey methods. To achieve this however, actions must be taken to protect the integrity of your survey data before the first response is collected. In this white paper, you will learn 3 proven methodologies for protecting survey data integrity and how to identify and defend against common data manipulation tactics.

Table of Contents

PROTECTING SURVEY DATA INTEGRITY

The increased popularity of online research is forcing the market research industry to address various concerns around data collection methodologies. If an online market research firm cannot confidently provide statistically accurate data to their clients, then their service provides little value. One of the first opportunities to increase data integrity in your research is in the survey design phase of an online survey. In this white paper, we will explore three specific practices that researchers can use in order to increase the integrity of their data.

What types of invalid respondents should we be on the lookout for? There are generally three types of invalid respondents; (1) inattentive, (2) fraudulent, and (3) speeding. Each of these types of respondents can be identified via "traps" within your market research survey.

Specific Methodologies:

1. Verification ratings are becoming more popular in online surveys and can be very effective in certain formats. This type of verification method is most effective in catching the inattentive survey respondents who might otherwise straight-line through your survey questions. The table below in an example of this preventative measure:

	Strongly Disagree	Disagree	Neither disagree nor agree	Agree	Strongly agree
Company X provides a quality product	○	○	○	○	○
Company X delivers its products on-time	○	○	○	○	○
Please verify your place in the survey by checking the 2nd box from the left labeled "Disagree"	○	●	○	○	○
I would recommend Company X to a friend	○	○	○	○	○

By including a "checkpoint" in the middle of a table format question, researchers are able to verify respondents' attention and remove straight-liners when necessary. A straight-liner is defined as a respondent who selects the same answer choice (possibly with an auto-fill feature) throughout the survey, forming a straight line down the tables. In this

example, if the respondent does not choose the "Disagree" answer choice, they would be marked as an "inattentive" respondent.

Of course, there are other question based verification methods that you can use in place of, or in combination with the example above:

- **Inconsistency** → We can simply ask for the respondent's zip code both in the beginning and end of survey. If the respondent's answers do not match then they would be classified as a "fraudulent" respondent.

- **"Red Herring"** → Within a list of possible answer choices, a nonexisting choice may be entered. For example, if a respondent claims to have eaten at a restaurant that does not exist, they can be flagged as a "fraudulent" respondent.

- **Opposite wording** → Respondents are asked a pair of similar questions at different points in the survey. For example, "I always use the internet for driving directions" and "I always use paper maps for directions". We are again able to identify "fraudulent" respondents with this verification.

2. Another method of protection which Infosurv has utilized since the company was founded is the use of unique ID's for each survey respondent. To protect the security of our Web based forms, each online respondent is assigned a unique **Personal Access Code (PAC)** embedded into their survey URL in his or her email invitation. Personal access codes can be hidden or visible to the survey respondent:

- **Hidden** – used when working with external sampling partners or other cases where the survey respondents do not need to know a PAC is being used. This eliminates the possibility for survey respondents to tamper with their PAC and allows demographic information to be embedded with confidence.
- **Visible** – used when we want the respondent to know they are using a PAC to access the survey. This gives survey respondents a sense of security and still allows demographic information to be embedded. If a survey respondent changes his/her PAC they will not be able to access the survey.

This methodology protects our clients from ballot-stuffers that may try to complete multiple surveys in order to skew the results. Following the survey administration period, we are able to use the PAC's to create demographic

slices of the results, identify randomly selected prize winners, pass back information to external sampling partners, and trace respondents in appropriate situations.

3. The third form of respondent verification we will discuss is **time-elapsed verification** which is used to catch "speeders" in market research surveys. These respondents are identified as not taking enough time to provide meaningful responses to your market research questions. The basic concept behind this methodology is to verify a valid survey response by examining the elapsed time each survey respondent spends completing the whole survey or a section of the survey. An acceptable range for the survey completion time can be calculated in two ways:

- An educated estimate based on the number of items in the survey. Infosurv typically uses the following time estimates during for this method: 3 close-ended questions per minute, 1 open-ended question per minute, 6 repeated-rating table items per minute.

- Evaluating the average completion time of the sample collected. Once an acceptable range has been defined, any outliers would be removed and considered invalid respondents. For example, if the acceptable range is determined to be 10 to 20 minutes and a respondent completes the survey in 2 minutes, it is more than likely that he/she provided invalid data.

Conclusion:

Professional researchers should always have tools accessible to hedge against data manipulation. It is the responsibility of each firm to ensure their clients are receiving valid data for which critical business decisions are often based upon. The methodologies explained in this white paper are just a few of the current strategies used to avoid invalid data and we are confident that as technological capabilities grow, more methods will arise. At least two measures should be taken in every market-research survey to protect data integrity. Continually taking these steps to protect client data will increase confidence in survey results across the industry and lay the foundation for more research in the future.

infosurv

About the Author

Chris Gwinner

Chris Gwinner joined Infosurv in January 2004 with the goal of helping Infosurv grow in terms of service quality, technical capabilities, and market research prowess. He currently serves as an Infosurv Senior Project Manager, overseeing projects for many of Infosurv's top government and Fortune 500 clients. Chris manages Infosurv's project management, technical, and analytical staff to assure the smooth operation of every Infosurv project and adherence to Infosurv's standard operating procedures.

During his term as an Infosurv Senior Project Manager, Mr. Gwinner has overseen survey projects for such recognized organizations as the United States Army, Turner Broadcasting, Accenture, MasterCard, Dell Financial Services, Merck Pharmaceuticals, The Weather Channel, Grainger, Caterpillar, LA Fitness, and Centex. He has also helped dozens of other corporate, government, academic, and non-profit organizations achieve superior results through the use of advanced research methodologies.

During his 4 years conducting online employee satisfaction, customer satisfaction, and market research surveys, Mr. Gwinner has assisted in the development a number of proprietary processes including The Infosurv Experience™. He and his team have consistently run projects smoothly and on-time for all Infosurv clients resulting in Infosurv's average overall client satisfaction rating of better than 9 out of 10. Clients consistently remark how their experiences with their Infosurv Project Managers exceed their expectations in terms of professionalism, responsiveness, and technical expertise.

Before joining Infosurv, Mr. Gwinner served as a business analyst at Delta Air Lines in Atlanta, GA. At Delta Air Lines, he conducted in-depth customer service analysis projects and process improvement studies for the busiest airport in the world. His specialties included customer service research and statistical analysis.

Mr. Gwinner can be contacted at chris.gwinner@infosurv.com.

infosurv

Find Out More

For more information on conducting better, faster, affordable research, and to read other useful topics in this series, please contact:

www.infosurv.com

888-262-3186

980 Hammond Drive, Suite 720
Atlanta, GA 30328

Subscribe to Infosurv Insider

Enjoy the latest in modern marketing research trends, case studies, and how-to guides.

The *Infosurv Insider* is a series of online case studies, white papers, and hot trends in modern marketing research edited and authored by recognized research practitioners and academic thought leaders.

Your route to inside information on better, faster, affordable research, the Infosurv Insider is available free to qualified subscribers—such as marketing and HR professionals.

Sign up now at www.infosurv.com/insider.