

Ensuring Respondent Quality In Online Research



Written By:

Michiko Kuwahara-Elrod
President
Sigma Validation

Mary Beth Weber
Executive Vice President
Sigma Validation

INTRODUCTION

In this day of faster and cheaper research, some market research professionals are losing sight of the true costs of “faster and cheaper.” Are we sacrificing the quality of our respondent-level data?

When market research firms are in a crunch to deliver “faster and cheaper,” one of the first shortcuts they take is curtailing validation costs. If the research firm validates the survey data, some respondents may need to be removed or replaced, which costs additional time and money. However, we all seem to be forgetting a basic principal of Market Research 101 – “Garbage In, Garbage Out.”

WHY VALIDATE ONLINE RESEARCH?

Validation is important in both traditional and online research. However, this paper will specifically address online research, where third-party validation methods are limited in scope or non-existent.

In this paper, we will discuss why research professionals should validate online research studies, and how they can take steps to ensure the respondent-level quality of the studies being conducted.

As previously mentioned, many research professionals have ventured into the online research world because of the significant time and cost efficiencies that can be achieved through this medium. However, some of the time savings should be invested into ensuring that multimillion/billion dollar decisions are being made based on the highest quality data possible.

1) Professional Respondents

One reason to validate online research is to identify and remove professional respondents. Professional respondents aren’t just repeat respondents. In this paper, we define professional respondents as those people who participate in a market research study only for the offered incentive. They will do whatever it takes to qualify for a study and the incentive – even if it means falsifying their answers.

There are some online panel companies who provide incentives to a respondent regardless of whether or not they qualify to complete the entire study. Then there are others who only provide the full incentive if the respondent does qualify. In the first case, some respondents might purposely answer screening questions so that they receive the incentive without completing the entire study. In the latter case, the respondent might complete the entire study regardless of whether they actually qualify.

To minimize the inclusion of professionals in the final analysis, consider offering some type of incentive to all respondents, regardless of whether they qualify. This will help to ensure that answers from professional respondents (AKA “Garbage”) will not impact the study results.

2) Respondent Overlap

Another reason to validate online research is to ensure that the same group of respondents isn't being contacted over and over again, through the same or different panel companies.

If a panel company has a low number of active panel members, the same individuals may be included each time a study is conducted. In addition, if a study involves more than one panel, and some of the respondents are members of each, the final analysis might include duplicate interviews.

3) Screening Requirements Not Being Adhered To

Validation is also important to ensure that screening requirements are being adhered to.

Use validation to confirm that the specified demographic group was actually contacted; and/or to ensure that certain respondents were excluded from a study, such as those who were recently interviewed regarding a similar product category.

4) Hyperactive Respondents

A company called comScore has a panel of respondents who have agreed to have all of their online activities passively measured. In exchange, these respondents receive some free antivirus-type software.

comScore says that it captures all of these respondents' online activities, including the details of the online surveys they participate in. In analyzing this activity, comScore estimates that 30% of all online surveys are completed by 0.25% of the population. In addition, they assert that this small group takes an average of 80 surveys over a 90 day period – that's about one survey per day – and are members of as many as seven other panels. (Source: <http://www.comscore.com/custom-research/sample.asp>)

If these statistics are accurate, it means that a very small percentage of the population is completing a lot of online surveys. However, is the data necessarily bad?

In the best case scenario, if all of these respondents are answering truthfully and are not just participating for the incentives, some of them will be conditioned. That is, their answers will have some bias just because they are so over exposed to market research studies and many of those studies might be in the same or similar product categories.

In addition, if a company is conducting multi-panel research, the odds are that it will have some of these hyperactive respondents in its study. As previously mentioned, if these respondents are members of the other panels being utilized, their duplicate interviews could be included in the final results. To minimize this potential problem, an online research provider should institute procedures to identify and remove these types of respondents from the research (such as downloading cookies to a potential respondent's computer).

5) Hyperactive Respondents Versus the Research Universe

Some researchers may think they don't have to worry about those 30% of surveys because their panel company has methods in place to purge their panel of these unwanted respondents.

Unfortunately, there is still concern. According to an article by Larry Fisher of Synovate, which appeared in the July 2005 edition of the MRA Alert! Magazine, 100% of surveys are completed by just 20-23% of US adults, regardless of whether the study is random sample or panel based.

This means that although there may not be hyperactive respondents in a study, there may still be a number of repeaters due to a small universe of accessible and willing research participants. Consequently, some type of validation or auditing procedure must be utilized to ensure that a fresh sample of respondents is being reached, and that the panel companies are continuing to increase or refresh their respective panel bases.

6) Worse than Professional Respondents

In a September 2004 white paper by Rachel Krupek of 20/20 Research, Ms. Krupek discusses how to handle paid survey sites. In the paper, she states that almost all paid survey sites encourage members to join more than 250-450 panels; and that some of the sites offer software to help respondents fill out surveys up to 300% faster. (Source: September 1, 2004 White Paper entitled "Handling Paid Survey Sites." <http://www.qualtalk.com/news/wp040901.htm>)

Accordingly, online research needs to be validated to ensure the elimination of not just ordinary professional respondents, but also respondents who are automatically completing surveys without even reading the questions.

7) Professional Respondents = Prize Mongers

In Harvey Lauer's "You Say Evolution, I say Devolution" article that appeared in the July/August 2005 edition of Quirk's magazine, Mr. Lauer asserts that "Technology and economic incentives may be transforming a pool of once-diligent, civic-minded respondents into a horde of game-playing prize mongers who view survey content as a necessary evil – an annoying obstacle to a grand prize, to be dispatched as quickly as possible."

By including this quote, we are not trying to be alarmists, but realists. We are not saying that all members of online panels are these prize mongers that Mr. Lauer is alluding to. However, these types of people exist in online research, and it is important to conduct simple due diligence to ensure that multimillion/billion dollar decisions are not being based on answers from these types of respondents.

8) Professional/Duplicate Respondents in a Panel

One of the panel companies that we spoke with said that they currently identify over 500 respondents a week trying to join their panel using the same or different email address, slight variations in the physical address and other techniques.

These could be members of the same household who innocently wanted to join the particular panel so that they could get the same rewards that their spouse or roommate were receiving. Or, they could be respondents purposely trying to create several fabricated accounts because they are the prize mongers referred to by Harvey Lauer. In either case, it is important to make sure that the panel companies have procedures in place to identify and eliminate these types of respondents from studies, so that the highest quality respondent-level data is achieved.

One of our clients was working with a leading panel company to conduct an In-Home Usage Test. They needed to receive the physical address in order to send out the sample. After receiving the file, they simply sorted it by address and identified a number of duplicate respondents that were not identified by the panel company itself.

In this case, the panel company probably did not create a standardized format for the physical address. For instance, according to a computer, “1275 15th Street” would not match with “1275 15 St.” However, if the panel company standardized the format, both addresses would be registered identically within the computer, and this problem would have been easily avoided.

9) Repeat Respondents & Research Standards

One of our clients was concerned with the over-interviewing of respondents in their online research. As part of their traditional standards, they require that their respondents should have never participated in another study involving their product category. Research professionals know this is difficult to achieve in the traditional world, never mind in the online panel world.

In order to better understand the practices and habits of their online respondent base, this client decided to add a question to their survey regarding the frequency of online survey participation. Through this question, the client found that their respondents, on average, participate in 5-7 online surveys a month. This client was surprised by the high number – they didn’t know what to expect.

comScore asked their panelists a similar question and found that respondents tend to underestimate how often they participate in online research studies. With comScore’s observation-based technology, they can compare their panelists’ actual activity versus reported activity.

By adding a question to their survey, our client is able to have a more realistic view of respondents in the online research world, and now can adjust and adapt their internal research standards and screening specifications based on this new knowledge. Based on comScore’s observation, such findings should be adjusted upwards if they are to be instituted as a maximum monthly participation rate as part of a company’s online research standards.

10) The Caveat of Panels Working with Panels

If the incidence of a study is low, a panel company might not be able to meet the sample requirements using their own membership. In these cases, the panel company may supplement their sample by sending a study invitation to members of another panel company – and they may, or may not alert their client of the need to do so.

The primary asset of panel companies is their membership base. Therefore, a panel company will not be quick to share email addresses and/or physical addresses with a competitor in order to find duplicates between the two respondent bases. If they are not checking for respondent overlap using this simple approach, what are they doing to help ensure that duplicate respondents are not included in the final analysis?

HOW TO ENSURE RESPONDENT QUALITY?

At this point, we have presented legitimate reasons to invest time in validating and/or auditing online research.

Now the question is: “What can market research professionals do to ensure respondent quality?”

In this section, we will make suggestions and recommendations to empower research professionals to ensure the respondent-level quality of their online research. This will require some reinvestment of the time savings that are attained through conducting online research, back into making sure that high quality respondent-level data is realized.

This investment will be well worth the effort. Just remember ... “Garbage In, Garbage Out.”

1) Ask Questions

The first thing to do is ask questions of the online sample provider. Simple questions will provide an objective way to compare and contrast panel companies.

a. How is the panel recruited?

First, find out how the panel company recruited their members. This will help to ensure that the desired demographic audience is being reached.

As an obvious example, to reach men between the ages of 25-54 who drive a Harley Davidson, don't work with a panel company who recruited most of their members from the GoodHousekeeping.com website.

In the same vein, to reach a representative sample of the US population, make sure that the panel company is recruiting in such a way that would be truly representative.

b. How large is the “active” panel?

A panel company may have 5 million members. However, how many of those panel members are active? For instance, how many of those 5 million members have responded to a survey in the past 3 months, 6 months or year?

Once the number of active panel members is learned, a researcher might then ask how many interviews the panel company conducts each week, month or year. By dividing the total number of interviews conducted by the total number of active panel members available, the average can be calculated for the number of surveys that one panel member is expected to complete over a particular time period.

This average can be estimated for each online sample provider, and will provide an objective measure to compare two or more companies. A panel company may have a relatively low membership base, but they may have a much higher percentage of active participants. Looking at this statistic could equalize the playing field in terms of available sample across two or more panel companies.

c. How many new respondents are recruited each month?

If the size of the active panel is known, then ask how many new respondents are recruited each month. At what rate is the active panel membership expected to grow on a monthly, quarterly or yearly basis? This will help to ensure that a fresh sample base is available for ongoing research needs.

What panel management procedures are in place?

As any panel company will attest, there are legitimate companies who use state-of-the-art techniques to ensure that they are providing the most accurate respondents and responses possible. But, as with traditional research, there are other companies who tend not to operate using these same high-level of standards. However, to distinguish the differences, ask additional questions.

d. How often does the panel company invite or allow respondents to participate in market research studies?

There are some panel companies, who send and allow their respondents to respond to more than 4 invitations per day, while others might send 4 invitations per month.

e. What is the panel company doing to identify and eliminate professional respondents within their database?

As mentioned earlier, one panel company uncovers more than 500 respondents per week trying to join their panel multiple times. Fortunately, this panel company has procedures in place to identify and remove these respondents. We recommend finding out what the panel company is doing to control this issue. Make sure that they are doing the basics – like standardizing the physical address to ease the identification of duplicates.

f. If a panel company is supplementing their sample for low-incidence studies with respondents from other panels, how are they eliminating duplicates?

As we discussed, members are a panel company's biggest asset, so they are not quick to provide a competitor with the email and/or physical addresses for their respondents. If they are not checking and eliminating duplicate respondents based on these attributes, what other procedure(s) do they have in place?

What cleaning processes are used after data collection?

A panel company may use all of the best techniques available to ensure that their members are who they say they are. Now the question is what are they doing after the data is collected?

g. Is the panel company checking that the respondents who answered the questions, are the same respondents who were sent the original invitation?

Are they checking for consistency in the demographics to make sure that the respondents' answers correspond to their previous answers regarding age, gender, ethnicity, and other respondent attributes?

h. Is the panel company checking for oddities in the respondent base?

Are they making sure that they're not receiving a number of responses from the same IP address (indicating multiple responses from the same computer)? Are there an unusual number of questionnaires being completed by users with the same unusual domain name?

As an example, Theo Downes-Le Guin of Doxus presented an interesting case study at the Chicago MRA conference in June of 2005. Mr. Downes-Le Guin said that in one of their studies, they found 5 identical response patterns each repeating dozens of times in the data set. These patterns accounted for more than 200 records, out of a total of approximately 850.

Doxus then found that all cases were the result of a single respondent who created more than 200 bogus panel profiles using "unique" email addresses in a domain that the respondent owned. This could have been avoided, or at least minimized, if the panel provider had a mechanism in place to ensure that panelists were entering valid physical addresses upon registration. Regrettably, this particular panel company did not have a simple address verification procedure in place.

i. Is the panel company checking the data for peculiarities in the responses?

As in the example mentioned, Doxus found identical response patterns in their research, which lead to the detection of a potentially data-corrupting issue.

An excellent recommendation made by Mr. Downes-Le Guin, is for research professionals to include decoy screening questions in their surveys to help identify professionals. For instance, include a question regarding the ownership of low incidence items.

In another case study, Doxus included such a decoy question and found that more than 71% of panelists who claimed to be Segway Human Transporter owners reported household incomes of less than \$35k. This vehicle retails at over \$5,000. What are the odds??

What is the panel company doing in terms of trying to identify these kinds of peculiarities?

j. Is the panel company checking the survey time from start to completion; and the time between questions?

As discussed earlier, some respondents use keystroke automation software that can help them complete a survey without ever reading the questions. What techniques does the panel company have in place to ensure that these types of respondents will be identified and their answers removed from the final results?

The panel company should be checking and tracking the time from start to completion and the time in between questions. They should then be able to compare these statistics across respondents to find an average. To ensure validity, ask them to remove any respondents who fall outside an acceptable range.

2) “Inspect” the Online Sample Provider

After asking questions, the next recommendation is to inspect the online sample provider. Inspect a provider by requesting a physical and/or virtual tour of their facilities and programming capabilities; and by conducting random audits of respondents and their surveys.

For instance, one of the panel companies that we visited gave a physical tour of their facilities including their state-of-the-art video conferencing rooms, which help project managers to coordinate their teams working on international studies across the globe.

In addition, they gave us a virtual tour of their programming capabilities including how they can turn on and off simple procedures such as which respondents should be included in a study based on demographics, geographics and other respondent attributes; and whether or not to allow a respondent to go back in a questionnaire after they completed a page or section.

An inspection of the programming for a survey is critical to do prior to the launch of a study. One of our clients was working with a leading panel company on a custom project and later found out that the company added their own questions to the survey and that they also reversed the client-specified preference scales -- the panel company did this for their own internal reasons, without the client's permission.

In traditional research, typically 10-20% of a study's respondents are required to be validated. In online research, why not require the same type of validation? Either request telephone validation for a certain percentage of respondents, or provide the online sample provider with a list of random respondent ID's and request a copy of their questionnaires or their participation history (including interview dates and study topics) over the past 3-6 months. Also request to review the questionnaires that fall outside a certain parameter in terms of the survey completion time or in terms of respondents' answers to decoy questions.

3) Implement Online Research Standards

One of our final recommendations is to implement online research standards. Speak with peers, online research providers and respondents. Develop a reasonable estimate for how often respondents should participate in online studies; and in online studies within a particular product category. Based on this information, some basic standards can be instituted for past participation.

Also, set up basic standards for the panel company regarding the identification and elimination of professional respondents within their database; and removal of duplicate respondents from supplemental sample. Also require notification and approval prior to outsourcing. Set up standards on how the online sample provider should be cleaning the sample after the study to purge suspicious respondents and their data.

Once standards are in place then set up procedures for auditing the online sample providers to make sure that they are in compliance. In addition, keep abreast of changes in the industry and adapt and modify standards as necessary.

IN CONCLUSION: EXPECT AND THEN INSPECT

Expect that online sample providers are appropriately managing the invitation and participation frequency of its members; and that they are using state-of-the-art techniques for identifying professional respondents both before and after data collection. Also expect that they are screening respondents based on the study's specifications.

Expect that respondents are answering questions truthfully. Also expect that respondents are participating in a study because they actually qualify and not only for the offered incentive; and that they are not the same people that were interviewed in the last study conducted through the same or different panel company.

Then Inspect. Inspect online research providers by asking questions and auditing them to confirm that they have best practices in effect for panel management and data cleaning, both before and after the data collection process. Also institute some type of validation or auditing method to randomly inspect survey programming and the respondent-level data.

ABOUT SIGMA VALIDATION

Sigma Validation is part of the New Generation Research Group, which was founded in 1968. The company collaborated with General Foods (now Kraft Foods) in 1978 to create the **Duplicate Number® Search**, a technique for identifying potentially data-corrupting respondent-level issues in traditional quantitative and qualitative market research. Most recently, Sigma Validation created the **Duplicate Email™ Search**, a new technique for helping to ensure the respondent-level quality of online research and Internet studies, including identifying duplicate and professional respondents.

Michiko Kuwahara-Elrod, President

Michiko (Miki) Kuwahara-Elrod assumed overall management of Sigma Validation upon the passing of her husband, John F. Elrod, who created and founded the firm. Her first mission as president was to aggressively update the internal programming and systems for the foundational Duplicate Number Search, so that Sigma Validation could continue to meet the industry's ever evolving needs – especially, in terms of the explosion in online research and Internet studies.

Miki has more than 40 years of business experience, and has been involved with Sigma Validation since its inception. She received her BA from the University of Chicago, along with a Traveling Fellowship from the School of the Art Institute of Chicago. In addition, Miki received her Masters from the University of Illinois through a Teaching Fellowship. One of Miki's most esteemed accomplishments is a manuscript that she created which is housed in the National Archives and can be seen at the Smithsonian.

Mary Beth Weber, Executive Vice President

In her current role at Sigma Validation, Mary Beth Weber works with clients to create new solutions to ensure the respondent-level accuracy of their national market research studies. In her previous positions at TNS, Simmons and REIS, Mary Beth helped clients to understand and utilize the value of syndicated databases and custom research for such things as evaluating sports sponsorship opportunities, choosing advertising placement, creating demographic and psychographic profiles for their customers, and making investment decisions on real estate-based portfolios.

Mary Beth has over 20 years business experience, with 15 specifically in market research, and received her BS and MBA in Marketing from Rutgers College and Rutgers Graduate School of Management, respectively.

SOURCES

Gian M. Fulgoni, "The Professional Respondent Problem in Online Survey Panels Today." MRA Annual Conference, Chicago, June 2005

comScore Website

<http://www.comscore.com/custom-research/sample.asp>, September 2005

Larry Fisher, "Random Sampling v. Panel-Based Research."
MRA Alert! Magazine, July 2005

Rachel Krupek, "Handling Paid Survey Sites" White Paper
<http://www.qualtalk.com/news/wp040901.htm>, September 1, 2004

Harvey Lauer, "You Say Evolution, I say Devolution."
QUIRK'S Marketing Research Review, July/August 2005

Theo Downes-Le Guin, "Satisficing behavior in online panelists: What's a research buyer to do?" MRA Annual Conference, Chicago, June 2005