



June 9, 2008

## Is The Long Online Panel Quality Nightmare Over?

Yes — If The Industry Embraces Several Key Quality Initiatives

by **Brad Bortner**

with Ellen Daley and Heidi Lo

### EXECUTIVE SUMMARY

Today, almost half of the quantitative research in the US is being conducted with online panels. While online panels provide vastly lower costs and increased speed, they also create heartburn and anxiety for market researchers because of the lack of confidence in online panel quality due to selection and respondent bias. In other words, are the hard to identify people on online panels' representative, and are they who they say they are? To date, major associations and vendors providing online panel samples have not generated coherent standards to enhance panels' transparency that would allow buyers to judge for themselves. Thankfully, this is all about to change with two major initiatives introduced by Peanut Labs and MarketTools.

### FASTER, CHEAPER, ONLINE PANEL RESEARCH ALSO DRIVES ANXIETY

Until the Internet, panels were a research backwater, with few market researchers willing to use them for research. The vastly reduced expense of using online panels enticed researchers to use them for all types of research. While many suspected there was a problem with projectability, it was hard to quantify because panel quality had been an enigma — with few identifying quality markers that buyers could trust to lift the fog.<sup>1</sup> With traditional research, users evaluate quality as data is collected, since they are recruiting, verifying, and interviewing. With online panels, users are reduced to just interviewing. All quality checks to ensure that respondents are who they say they are, that no fraud is going on, and that no one is being over-surveyed are in the opaque hands of the panel provider. What are the key online panel quality questions?

- **Who joins online panels?** Online panels are generally composed of people recruited online.<sup>2</sup> Some are recruited directly by the panel owners. Others are recruited through aggregators that recruit from many online sources and supply sample to the panels. Given the basic anonymity of the Internet, this means that it is challenging to know whether panel members are who they say they are. Even proving such fundamental demographics as age, gender, and location can be hard. As a result, it is difficult to know if your results can be projected to the market you care about.<sup>3</sup>
- **Are professional survey takers (PSTs) creating projectability issues?** Some people really like to take surveys. Lots of surveys. Their reasons for this can range from greed — and the potential for harvesting incentives — to just desire to make their opinion known. A lot. These people not only tend to sign up for multiple panels; they also tend to accept lots of survey invitations.<sup>4</sup> Some unscrupulous PSTs might even intentionally disguise themselves in order to take the same survey several times. Even worse is the potential for unscrupulous PSTs to create survey-taking automated programs or bots to harvest incentives.

- **Is voting early and often a way to project to a market?** Another issue is generated by the common use of sample aggregators to stock panels. This means that people — knowingly or unknowingly — may be signed up for the same panel several times because they signed up for what they thought were several panels. If they used different email addresses, deduping these people becomes quite hard, even for panels who are diligently trying. The result is similar to PSTs: The same person may be responding to the same survey several times.
- **How can you weight it out if it's not there?** Some respondent types are not just underrepresented on online panels — they are close to being totally missing. The very poor, the very rich, and technophobes on the consumer side, as well as senior-level decision-makers on the B2B side, tend to be notable in their absence. While many distortions in panel composition can be fixed by weighting them to a known universe, researchers need to keep in mind that weighting extremely underrepresented (or absent) sample is bound to be misleading. For these cases, traditional data gathering methods, while expensive, remain superior.

## THE CAVALRY FINALLY ARRIVES WITH NEW QUALITY INITIATIVES

A solution to resolve these fundamental online panel issues is at long last arriving as a result of several recent initiatives that went live in May/June 2008. At their core, they all attempt some level of identity verification and ensure that the same person does not take the same survey several times, either accidentally or through subterfuge. Some are more comprehensive than others, but all hold out the promise of significant improvements in online panel sample. The first initiative allows buyers of sample to pass it through a machine fingerprinting verification process to ensure that the sample is good. The second initiative involves buying sample from sample vendors that are using a variety of protocols to clean up their sample.

### Initiative One: Virtual Machine Fingerprinting To Enhance Quality At A Machine Level

Machine fingerprinting allows you to check your sample to ensure that the same machine is not represented in your sample multiple times, perhaps under different aliases. It holds promise because it is an automated approach and is therefore relatively inexpensive to execute. It also can be deployed against any online sample. Its downside is that it is machine-based, so corroborating respondents' true identity is still a challenge. These solutions come primarily from two vendors (although only one is market-research-specific).

- **Peanut Lab's Optimus solution.** This approach allows market researchers and panel providers to ensure that the same survey is not taken by the same PC or that too many surveys are taken in a given period. The technology employs an electronic fingerprinting approach that any panel, research firm, or research buyer can license. A machine is fingerprinted by Optimus as the survey is initiated by redirecting clickstream data. Since it is machine-specific, this methodology works even for those respondents recruited through aggregators or when panels are blended

to create larger samples. Peanut Labs also include steps to track survey speedsters and other problematic survey takers so that repeat offenders can be scrubbed. Key backers of this initiative include Knowledge Networks, PSB, Decipher, and ReRez.

- **41st Parameter's PCPrint approach.** This antifraud technology also uses electronic fingerprinting, similar to Peanut Lab's offering. However, 41st Parameter's current focus is on key financial institutions and online merchants. MarketTools has recently integrated them into its TrueSample initiative. 41st Parameter does not currently have a go-to-market offering for the general market research market. Market research firms that are intent on using PCPrint as a standalone product would likely have to do some custom work to integrate it into their sample cleaning protocols.

### Initiative Two: A Belt-And-Suspenders Approach

MarketTools has just released its TrueSample solution for panel sample quality. This initiative has backing from major consumer and technology market research buyers such as senior researchers at Proctor & Gamble and General Mills. Essentially, various panel sample providers sign up to go through the TrueSample protocol. Researchers of all stripes could then be assured that their sample is clean if they buy from a TrueSample provider.<sup>5</sup> Initial measurements show that almost 30% of existing sample can be flawed, with obvious implications for valid research findings.<sup>6</sup> MarketTools is licensing its TrueSample solution to online panel providers and other market research vendors. Its approach is multitiered and focuses on three stages of data hygiene:

1. **Prove that respondents are real.** Using an approach common in mass-target marketing, TrueSample uses a third party (so confidentially is assured and vendors need not fear that MarketTools is poaching sample) to test sample identities against publically available sources that direct marketers use to clean their lists, to ensure that their identities exists, and to match their panel tags. If they do not meet these criteria, the sample is removed.
2. **Prove that they are unique.** MarketTools' second pass at enhancing quality involves deduplication across multiple sample providers through integration of virtual machine fingerprinting and identify comparisons. This looks for matching on a combination of machine fingerprints and identities from the step above. Of course, for this to work, you should be sourcing all of your sample from a TrueSample vendor that would take care of this matching process. This means that by the time the sample reaches this level, respondents have not only been proven to be actual people, but they have also been proven to be unique, at least at a machine level.<sup>7</sup>
3. **Prove that they are engaged.** Finally, TrueSample attempts to measure other fraud or disengagement issues. This can take many forms ranging from the disengagement (as indicated by not showing variation in scaled responses) to a program taking a survey (as indicated by seemingly random responses in super fast time) to professional survey takers whipping through

surveys faster than could reasonably be expected to just badly designed surveys that cause respondent confusion. While this is an interesting quality feature to implement, it is also the hardest, since it is difficult to sort out what is caused by bad survey-taking behavior and what is caused by poorly designed surveys that do not keep a respondent's interest.<sup>8</sup>

### Why Should Sample Providers Play In This Sandbox?

One might fear that online panel providers will attempt to ignore both of these initiatives, much the way US car manufactures did in the '50s with Dr. Demming when he preached the values of statistical process controls, with a "Go away; we are making money."<sup>9</sup> Given the major players backing these initiatives, however, Forrester does not believe that such a wall to quality improvement will stand.<sup>10</sup> Major sample buyers — many market researchers and research providers — will increasingly insist on online panels' cleanliness, and the end of shaky panel quality is clearly in sight. Who would risk buying from a panel that does not embrace one of these quality approaches when they think that as much as 30% of its sample could be contaminated? Buyers want this. Panels will embrace it over time.

## RECOMMENDATIONS

### REQUIRE THAT ANY ONLINE SAMPLE BE CLEAN

There is no longer any excuse for buying dodgy sample. The industry is starting to provide clear ways to buy clean sample with confidence. You can help move the market. The sooner major buyers demand it, the sooner most panels will fall into line.

- **Pay the extra for quality.** It will cost somewhat more, but the increment should be on the order of cents per response, not dollars. Panel sample providers are in a very competitive business, which will likely keep any incremental cost relatively low. It is simply not logical to buy suspect sample for marginally less money when better-quality, verified sample exists, especially since the cost of sample is only a fraction of the cost of most research projects. That would be like saying, "Hey, the answer is wrong, but I got it for 10% less!"
- **Continue to push the vendors.** These initiatives are new — there are going to be challenges and opportunities for you to recognize, and it is not clear whether any sample cleaning approach will become dominant or if there will be several vendors in an oligopoly. This competition is a good thing. Continue to push the vendors to help them prove that their sample is the cleanest. In the long run, that means everyone will win.

## ENDNOTES

- <sup>1</sup> Some researchers have just given up on the idea of projectable research altogether with these methods. See the February 22, 2008, "[Does Declining Research Projectability Matter?](#)" report.
- <sup>2</sup> Very few panels use recruiting offline through traditional random digit dial methods because it is very expensive and makes their sample unattractive from a price perspective for buyers who cannot prove that it is of higher quality.

- <sup>3</sup> Forrester provides a detailed description of all of the issues degrading potential quality with online panels. See the May 1, 2007, "[Online Panels: Cheap And Effective, But Buyer Beware](#)" report.
- <sup>4</sup> comScore created quite a stir when it showed that for some panels, a small percentage of respondents were taking a huge percentage of the surveys. Forrester provides more details about the perils of PSTs. See the January 22, 2008, "[Professional Survey Takers, Still A Threat!](#)" report.
- <sup>5</sup> One research firm, Longwoods International, is pushing a "TrueSample Inside" campaign to differentiate its offering. Its message is to buy its research because the core sample it is using for its analysis is demonstrably cleaner.
- <sup>6</sup> Optimus has shown similar shrinkage numbers. Bear in mind that the actual effects on specific surveys are hard to assess, since it is an open question as to how many duplicate panel members would make it into a specific survey.
- <sup>7</sup> Both Optimus and TrueSample's machine fingerprinting claim additional granularity, since they include information about logon identities as part of their ID. Making this work, however, assumes that different people sign on with different IDs on the same machine. It is hard to assess how often this is the case.
- <sup>8</sup> Currently, this also requires researchers to use MarketTools' zTelligence survey software, aside from doing business with only TrueSample vendors. This may prove to be a hurdle to some buyers who are married to different survey software platforms, but they will still get all of the other benefits of TrueSample if they deal with certified vendors.
- <sup>9</sup> Of course, the Japanese, neglecting to ask if anyone in the US was embracing statistical process control (SPC), took Dr. Demming (the inventor of SPC in the US) to heart. The result was a surge in the Japanese auto market that lasted until this day.
- <sup>10</sup> The specific implementations of these solutions also let sample vendors rest assured that neither Peanut Labs nor MarketTools can illicitly gain access to their sample. Security for such is well defined.