Whitepaper

Evaluating Sample Suppliers in the Digital Age

By JD Deitch Ph.D., Cint
and Alyssa M. Haskins, The Harris Poll

Additional contributors:
E. Paul Johnson, The Harris Poll
and Michelle Gosney, The Harris Poll

Cint Accelerating insights.

The Harris Poll
Harris Insights & Analytics LLC, A Stagwell Company

Introduction
Pg. 03

4
Challenges in the digital age

6
Intent of Research

Key Research Questions

Methodology

7
Accuracy of the Data

8
Representativeness

Fraud Levels

9
Participation Rates

Comparability

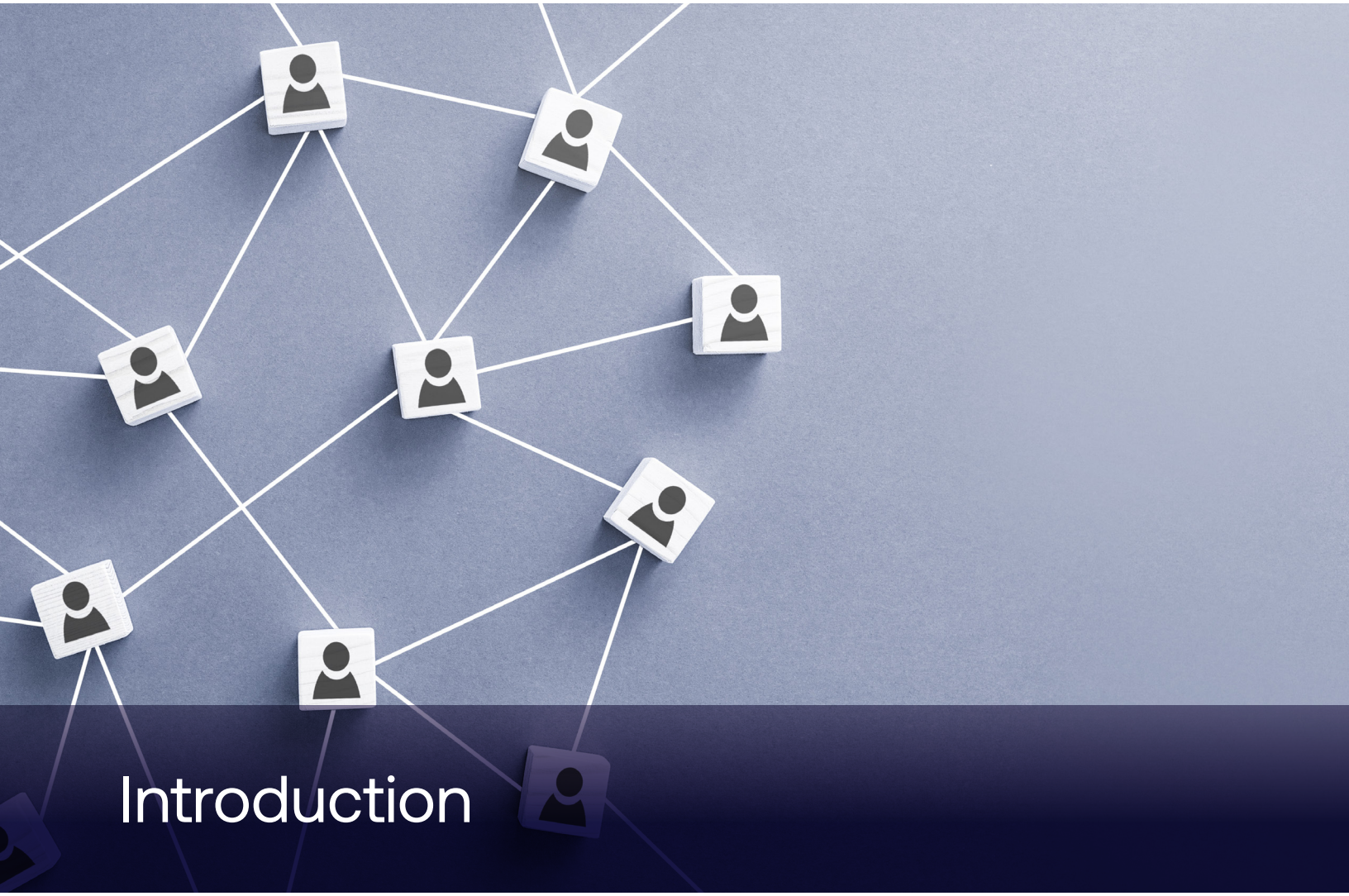
10
Discussion

11
Summary of Results

12
Best Practices of Evaluating Sample Suppliers

Conclusion
Pg. 13

Contents



Introduction

Sample management in the digital age presents a new set of complex challenges. While the benefits from a highly networked supplier landscape are numerous, the scale and diversity of sources is extraordinary. With these, the industry has gained greater liquidity and transparency in the ecosystem. Programmatic (i.e., computer driven) delivery and intelligent routing are now the dominant mode of delivery. This means sample can come from anywhere and be redirected anywhere. While sample suppliers understand this, sample buyers are far less aware despite the significant implications for their work. To shed light on this topic, The Harris Poll conducted a study to evaluate current and potential sample sources and blends from various providers on measures of both Quality and Comparability. The study found, not surprisingly, that the different samples had varying results on these evaluations and results were mixed across panel type as well (i.e., API, double opt-in, etc.).

Challenges in the Digital Age

It is truly remarkable how the supply of respondents has been transformed over just the past 5 years, much less since the dawn of the online sampling era some 25 years ago! The industry has evolved from a balkanized collection of proprietary research-only panels to a rich supply ecosystem providing access to diverse respondents at an unprecedented scale.

Today's sample ecosystem is virtually fully networked via application programming interfaces (APIs), which have led to significant benefits for suppliers and buyers alike. Like water over uneven ground, respondents can flow to available surveys, creating greater overall liquidity, improving market transparency around pricing, and leading (in most cases) to an improved respondent experience. The automation implied by API delivery has also created significant efficiency gains for research and project managers. Taken to its logical end, automation can ensure that once the targets are selected, projects can be launched, fielded, and closed without the intervention of a human, even if the parameters change in field.

With these advantages come predictable downsides. This vast, networked, automated ecosystem requires us to have similarly robust automated tools to detect and mitigate fraud and disengagement. Systems need to be built to accommodate the idea of consistent source blends for tracking and normed studies. We cannot let the fancy tech cause us to lose sight of the fact that we are trying to represent known populations, which means we need to think about sample bias. Even this can be a struggle with suppliers who don't "speak research". These downsides are manageable though. There are huge strides that have been made in machine learning algorithms that detect bad respondent behaviors. And the combination of survey data collection systems and effective communication between buyers and suppliers helps to mitigate issues around sample blends and other research concerns.

Finally, despite the remarkable changes, many researchers remain unaware of what the modern supply landscape looks like. They thus remain susceptible to myths they knew to be true a dozen years ago, or still have a mindset with terminology that doesn't always apply. The most pernicious of these myths is that double opt-in panels are the 'gold standard' and real-time or programmatic sampling methods are bad. "Programmatic" simply means the provision of sample using automation. There is nothing wrong with double opt-in panels, but there is nothing intrinsically better about them in the programmatic era, either. Indeed, communities and other non-traditional providers

are almost universally better at optimizing the user experience with more profiling data, better targeting, greater efficiency, and more stuff to do.

Another topic that creates concern is related to the concept of response rates and random sampling. These days, the majority of sample delivered to surveys is not via invites, but programmatically. By simply turning up at some website or application, the respondent is effectively signaling her/his availability for a study, at which point one is presented—which is the opposite of how it "should" work per sampling theory. This means that metrics like response rate no longer make sense. Beyond this, each supplier has its own intelligence that decides which opportunity or survey to present to the user, meaning that the selection of the respondent for a given study could be conditional on other factors which may then impact the validity of the data in that study.

Ultimately, the truths around sample have not changed in the face of these great technological advancements. Cint has enshrined four fundamental principles in its Quality Charter:

1. Respondents must be real, that is not fraudulent.
2. Respondents must be unique, that is not duplicates.
3. Respondents should be engaged, that is they should provide accurate and complete answers to questions.
4. Respondents in the sample must be representative of a known population. As with most research in the industry, we are still using convenience samples rather than probability-based samples, yet there is a lot of work that has been done on sample design and post-hoc weighting that can help with this.

"Another topic that creates concern is related to the concept of response rates and random sampling.

The Harris Poll Study

Intent of the Research

The intent for this research was to evaluate current and potential sample sources and blends from various providers on measures of both Quality and Comparability. With a rapidly changing marketplace, this annual review is a critical part of the vetting process for sample providers and offers a solid starting point and framework for decision making on sample source selection, along with other factors such as: pricing, ease of use/management, and feasibility.

Key Research Questions

To assess Sample Quality The Harris Poll study focused on accuracy of the data, representativeness, fraud levels, and participation rates. External benchmark measures were utilized to gauge the accuracy of each sample overall and within several topic areas. Representativeness was evaluated based on the amount of post-stratification weighting needed to align each sample on known population demographics post field. Fraud levels for different types of fraud, including digital fingerprinting and in-survey metrics for inattentiveness were also reviewed for each sample. Participation rates, to gauge the level of professional respondents, were measured by comparing the number of surveys taken per month.

For Sample Source Comparability, in addition to the external benchmark measures, custom questions were included to help inform on optimal replacement options in the event feasibility issues occur for a given provider on a study.

Methodology

The Harris Poll interviewed 7,644 US adults age 18+ across nine different non-probability sample provider blends from June 9, 2021, through June 21, 2021. The sample sizes for each ranged from n=835 to n=856.

Respondents were asked a range of questions on topics including Sense of Security, Sense of Community, Altruism, Political Views and Engagement, News Consumption, Opinion Elite & Other Engagement, Chronic Health Conditions, Health Related Behaviors, Internet & Social Media, Shopping Habits, and Demographics.

External benchmark comparisons to the following sources were used for over 50 of the survey items: Census Pulse Week 32: June 9-21, 2021, NHIS 2018, NHIS 2019, Census Voting and Registration Supplement 2020, CPS Volunteer Supplement 2019, CNN Politics, PEW NPORS 2020, PEW ATP Wave 74 Online Harassment, PEW March 13-27 and April 4-18 2017, PEW ATP Wave 13 December 2015, PEW 2018 National Survey of Latinos, Current Population Survey (CPS) 2020, and Experian.

Additional custom questions were added to round out the topic areas for comparisons.

Data were weighted separately by sample blend to population proportions from the Current Population Survey (CPS) 2020 for Education, Age by Gender, Race/Ethnicity, Region, Household Income, Household Size, and Marital Status. Data were also weighted as a total combined data set to the same figures to provide a representative Total across all samples.

Accuracy of the Data

Using the MAE (Mean Absolute Error) for the benchmark measures, the top four most accurate blends from this research were Sample G, Sample H, Sample D, and Sample F. The blends with the greatest differences from the benchmarks overall were Sample C, and Sample B. The chart below shows strengths and weaknesses across samples when the MAE is broken down by topic area.

MAE (Mean Absolute Error)	Sample A	Sample B	Sample C	Sample D	Sample E	Sample F	Sample G	Sample H	Sample I
Overall	3.85%	4.26%	4.16%	3.67%	4.07%	3.71%	3.64%	3.66%	4.06%
Demographics	1.86%	1.79%	1.77%	1.66%	1.94%	1.66%	1.86%	1.74%	1.85%
Sense of Security	5.61%	5.59%	5.50%	4.94%	5.10%	4.32%	4.78%	4.01%	4.75%
Sense of Community	5.78%	7.03%	6.00%	6.76%	6.18%	7.21%	6.78%	5.98%	6.24%
Altruism	7.32%	4.97%	10.57%	13.04%	8.47%	13.98%	13.03%	8.36%	7.49%
Political Views and Engagement	5.37%	6.26%	7.51%	5.49%	6.16%	3.84%	5.59%	5.09%	6.77%
News Consumption, Opinion Elite, and Other Engagement	9.88%	10.85%	8.93%	9.04%	8.33%	9.21%	8.18%	9.80%	9.82%
Chronic Health Conditions	5.32%	7.47%	5.80%	3.79%	4.99%	3.29%	3.46%	4.05%	6.36%
Health Related Behaviors	8.33%	9.19%	8.77%	7.04%	8.22%	5.71%	6.98%	7.59%	8.09%
Internet and Social Media	4.56%	5.40%	6.05%	5.67%	5.25%	5.64%	4.18%	4.48%	5.96%
Shopping Habits	8.56%	10.93%	10.19%	8.19%	11.24%	13.38%	9.11%	10.84%	9.64%
Mosaic Codes (Experian Respondent Classification)	0.85%	0.71%	0.68%	0.83%	0.83%	0.84%	0.83%	0.79%	0.91%

The color coding above shows distance from the benchmark measures for each topic (row) across samples. Dark red indicates the sample that was the furthest from the benchmarks and dark green indicates the sample that was the closest to the benchmarks for each topic.

Representativeness

Representativeness was gauged using the weighting efficiency of the post stratification weights for each sample. All providers had access to real-time dashboards during field allowing for the comparison of the incoming completes against the weighting targets for each demographic variable. Sample F had by far the highest weighting efficiency at 92.3%. The other samples were much closer to the average, with Sample D, Sample G, and Sample H, coming in just above the average.

Sample A required a target adjustment for an overage of Males 65+ and Sample C required a target adjustment for a shortfall of Less than High School Education. Target adjustments are utilized to minimize any extreme weight factors in the RIM weighting process. Since these two samples exhibited proportions outside of the standard parameters allowed for weighting on the factors mentioned, adjustments were needed. Had the adjustments not been made, the weighting efficiencies would have been even lower.

Weighting Efficiency	Sample A	Sample B	Sample C	Sample D	Sample E	Sample F	Sample G	Sample H	Sample I
Overall Weighting Efficiency: *(average: 59.8%)*	49.0%	53.5%	54.5%	60.9%	52.3%	92.3%	60.7%	62.4%	52.4%

Fraud Levels

Overall fraud level measures included Digital Fingerprinting Fraud Score Failure, as well as several in-survey metrics for inattentiveness: Real Answer Failure, ISQ Failure, and Manual non-engaged respondent identification. Samples G/H had lowest overall fraud level at 7%, with Sample F coming in a close second at 8%. Sample A, Sample D, and Samples B/E had the highest percentage of fraud, all coming in over the 18% average. Sample C and Sample I were just below the average for the overall measure.

When looking at the in-survey inattentiveness measures only, Sample H still outperformed most with 5%. Sample F and Sample G also still came in below the average of 11%, at 7% and 8%, respectively. Sample A, Sample B, and Sample E performed better with the in-survey metrics only, coming in much closer to the average. However, Sample D still came in much higher than the average at 21%. Sample C and Sample I came in just slightly above the average for the inattentiveness measures only.

Fraud Levels	Sample A	Sample C	Sample D	Sample B & E*	Sample F	Sample G & H*	Sample I
Percent Fraud Overall: *(average: 18%)*	32%	15%	25%	22%	8%	7%	14%
Percent Fraud - Inattentiveness measures only *(average: 11%)*	12%	12%	21%	B: 15% E: 12%	7%	G: 8% H: 5%	12%

**Overall fraud measures only reportable at the overall provider level, not separately by blend.*

Participation Rates

Sample F had by far the lowest number of surveys taken per month by panelists. Sample A was also below the average. The other samples were relatively similar on number of surveys, all just above the average.

Number of Surveys	Sample A	Sample B	Sample C	Sample D	Sample E	Sample F	Sample G	Sample H	Sample I
Average Number of Surveys Taken per Month by Panelists: *(average: 28.45)*	25.64	30.45	32.85	30.20	31.36	12.77	33.15	30.91	30.01

Comparability

Based on the comparability of responses across all metrics in total (benchmark and non-benchmark), Sample E, Sample C, and Sample H, came up the most frequently as good replacement options for instances where feasibility issues arise, and there is a need to shift sample allocation to other providers for tracking work. Sample A, Sample D, and Sample I came up as replacement options in only a few instances, and Sample F did not come up as an optimal replacement for any of the other sources. Sample B was not included in this portion of the analysis as it was a blend being phased out of use. The chart below is an example of information to use as a starting point for determining possible replacement options and would be used in conjunction with information on other factors such as feasibility and pricing for a given study.

Replacement For:	Sample A	Sample B	Sample C	Sample D	Sample E	Sample F	Sample G	Sample H	Sample I
Option 1	Sample E	Sample E	Sample E	Sample A	Sample C	Sample G Sample H	Sample E	Sample C	Sample H
Option 2	Sample D	Sample C	Sample H	Sample H	Sample H	Sample E	Sample D	Sample E	Sample C
Option 3	Sample C	Sample H	Sample I	Sample E	Sample A	Sample C	Sample I	Sample I	Sample A

Discussion

Although this annual research is just a snapshot in time of a small subset of the overall panels, based on the comparisons across metrics, and additional feedback on process and usage from colleagues, we've created a report card for the providers included in this research. This is meant to serve as a guide for decision-making by research experts for consumer studies, with the caveat that the generalizations below may not be applicable to all scenarios, so other factors may need to be considered.

In the interest of collaboration, The Harris Poll also took the opportunity to share the results (blinded) of the research with each of the providers and noted strengths as well as areas where there is an opportunity for improvement.

Report Card	Sample A	Sample C	Sample D	Sample B & E*	Sample F	Sample G & H*	Sample I
Accuracy of Data							
Representatives							
Fraud Levels							
Participation Rates							
Good Replacement Option							
Feasibility							
Pricing							
Overall Ease of Use / Management**							

Excellent	Good	Fair	Needs Improvement

*** Ease of Use/Management is defined by attentive project management - communications, pacing, minimum over quotas, balancing of targets, etc.*

Summary of Results

As exemplified by the research here, not all sample providers are the same and it is important to be an informed consumer. It is critical to consider a wide range of factors that are measurable, such as accuracy of data, representativeness, fraud levels, participation rates, and comparability. However, it is also imperative to include experiential factors as well, such as feasibility, pricing, and overall ease of use/management, in addition to the specifics of the project under consideration. For example, Sample D excels on accuracy of data but needs improvement on feasibility, so this sample might not be an optimal choice for a new tracking study or a study with a low incidence rate population of interest. However, it may be a good option for a small, high incidence rate, ad hoc study, given that it excels on both accuracy of data and pricing, provided there is ample time in the project cycle to account for fraud removals and replacement. Sample F performed very well on most of the metrics but did not come up as a good replacement option for other sources, and comes in higher on pricing, so may only be appropriate for instances where other sources are not feasible or where several sources would need to be included to achieve the project goals. Samples G and H performed very well across the board on all the metrics included in the analyses. The attention to continued quality is evident from these results as well as day-to-day experiences with this provider.

Best Practices for Evaluating Sample Suppliers

The goal of this research was not to show that data quality and representativeness vary by suppliers. This should hardly come as a surprise! Rather it was to understand how these factors operate at key moments in this research process. The aim of sharing this work is to help fellow researchers understand how to be informed buyers and, specifically, where to look for issues at which points in the research process.

The Harris Poll takes a whole process approach to data quality and representativeness with focus on all aspects of the research process. To produce a representative view of a population, relying on weighting alone is not enough. Alignment with sampling and appropriate survey design are necessary for the minimization of bias (error) inherent in survey research.

Sampling

- Utilize purposive sampling techniques that allow for optimal efficiencies and alignment with the population of interest
- Monitor the demographic profiles of incoming completes in real-time and work with sample providers to align with known population proportions during field to minimize the amount of weighting needed

Survey Design

- Ensure alignment of screener items, qualification criteria, and quota section designs
- Make sure question formatting is appropriate for accurate measurement and is mobile-first
- Be cognizant of survey length and make appropriate recommendations to clients to maximize benefits of data and maintain quality responses
- Use thorough in-survey fraud detection techniques and manual review of data, to identify inattentiveness, inconsistencies, illogical responses, duplication, and fraudulent activity, and provide a feedback loop to suppliers

Weighting

- Employ post-stratification weighting for final alignment with the population of interest to alleviate potential remaining errors that may affect data quality, enabling representative and projectable results
- Engage in rigorous internal research on optimal calibration techniques

The technological advances we have seen in online sample provision have increased the scale, diversity, liquidity, and transparency of sample supply in our ecosystem. Yet they have also increased the prevalence of fraud, disengagement, questionable practices, and — simply and objectively — changes to how we need to work. To do online research is to accept these implications. Yet this does not mean we are adrift in a sea without principles. As this paper shows, a holistic view of the research process that includes both sample and survey design allows us to adapt our methodology to the realities of online sampling in the digital age. A structured approach gives us a framework to ensure solid research outcomes and be informed buyers as well.

About Cint

Cint is a global software leader in technology-enabled insights. The Cint platform automates the insights gathering process so that companies can gain access to insights faster with unparalleled scale. Cint has the world's largest consumer network for digital survey-based research, made up of over 160 million engaged respondents across more than 130 countries. Over 3,200 insights-driven companies use Cint to accelerate how they gather consumer insights and supercharge business growth.In June 2021, Cint acquired Berlin-based GapFish – the world's largest ISO certified online panel community in the DACH region - and in December, completed the acquisition of US-based Lucid – a programmatic research technology platform that provides access to first-party survey data in over 110 countries. Cint Group AB (publ), listed on Nasdaq Stockholm, has a rapidly growing team across its many global offices, including Stockholm, London, New York, New Orleans, Singapore, Tokyo and Sydney. (www.cint.com)

About The Harris Poll

The Harris Poll is a global market research and consulting firm that strives to reveal the authentic values of modern society to inspire leaders to create a better tomorrow. It works with clients in three primary areas: building twenty-first-century corporate reputation, crafting brand strategy and performance tracking, and earning organic media through public relations research. One of the longest-running surveys in the U.S., The Harris Poll has tracked public opinion, motivations and social sentiment since 1963, and is now part of Stagwell, the challenger holding company built to transform marketing.

For more information on how Cint can help you with your market research studies please contact your account manager or info@cint.com

For more interest about The Harris Poll Study, please contact Alyssa Haskins, alyssa.haskins@harrispoll.com

The Harris Poll
Harris Insights & Analytics LLC, A Stagwell Company

Cint Accelerating insights.

Cint is a global software leader in technology-enabled insights. The Cint platform automates the insights gathering process so that companies can gain access to insights faster with unparalleled scale. Cint has the world's largest consumer network for digital survey-based research, made up of over 160 million engaged respondents across more than 130 countries. Over 3,200 insights-driven companies use Cint to accelerate how they gather consumer insights and supercharge business growth.In June 2021, Cint acquired Berlin-based GapFish – the world's largest ISO certified online panel community in the DACH region - and in December, completed the acquisition of US-based Lucid – a programmatic research technology platform that provides access to first-party survey data in over 110 countries. Cint Group AB (publ), listed on Nasdaq Stockholm, has a rapidly growing team across its many global offices, including Stockholm, London, New York, New Orleans, Singapore, Tokyo and Sydney. (www.cint.com)

cint.com