



Editor's *Vue*

Nikita James Nanos, CMRP

Apart from our new look in 2005, *Vue Magazine* also has some new features. Two new columns have been launched to reflect the diverse MRIA membership. **Bob Collins**, in his "RU What", is a new regular feature on issues related to research users (RU) in Canada. Another new column from **John Ball**, the "RAC Report" focuses on research agency issues. Together they round out a stellar lineup of regular columnists.

I would like to take this opportunity to thank our regular columnists for their hard work and commitment to the success of our Association and magazine.

Thank you...**Natalie Gold** (The QualCol), **Ruth Lukaweski** (B2B Reporter), **David Lithwick** and **Enrico Codogno** (CI Corner) and **Stewart Hemerling** (i.on Research).

Sex & fire...do I have your attention?

Our feature this month is from the venerable **Michael Adams**. In his article "Sex and Fire," Michael explores the Canadian and American views on religion, homosexuality and authority.

Following Michael is an article by **Mike Farrell**. Mike shares insight in the role of research in the creation and launch of "stupid.ca", a cutting edge anti-tobacco campaign, targeted at Ontario youth.

Benoît Gauthier in "Are all samples of telephone numbers created equal?" reveals the findings of a very interesting study which compares the quality, efficiency and precision of the samples from five sample providers. The findings, related to the specific sample sources, have been masked but this is a must read for practitioners.

If you want to share your experience, a book or product review, or your opinions, please do not hesitate to contact me directly. *Vue Magazine* is considering submissions on a full range of topics. Our next deadline for authors and advertisers is February 20th. Please feel free to forward your ideas and articles to me by e-mail nnanos@sesresearch.com or telephone (613) 234-4666.

Until next month,

Nikita James Nanos, CMRP
Editor-in-Chief

are all samples of telephone numbers created equal?

by Benoît Gauthier

All samples of telephone numbers are not created equal: service quality varies from supplier to supplier, but, most importantly, the efficiency and precision of the samples are also variable. Key criteria are service quality and pricing, efficiency, avoidance of out-of-geography numbers and inclusiveness.



MOST COMPANIES ACTIVE IN MARKET research and in telephone surveys use specialized suppliers to secure samples of telephone numbers. These specialists collect information about the use of telephone exchanges (the first three digits in the seven-digit number that follows the area code); they attach geographical information to listed numbers; they develop algorithms to draw lists of random numbers that are efficient and well targeted.

For the United States, there are five companies offering what appear to be similar sampling products. This article analyses the quality of the products and services offered by these five companies in the U.S. (in alphabetical order): Affordable Samples Inc., ASDE Inc. (Survey Sampler), Genesys Sampling Systems, Scientific Telephone Samples and Survey Sampling International LLC. The underlying research was funded by ASDE Inc.

but conducted in an independent manner, using an approach that avoided assessment biases. The information and suppliers referenced have been randomized for anonymity.

THE CRITERIA

In this research, three facets of sample supplier "quality" were measured:

- the efficiency of the samples, i.e., the proportion of the numbers supplied that were ringing residential numbers;
- the precision of the samples, i.e., the proportion of the numbers supplied that were within the pre-determined geographical area and the proportion of the geographical area represented in the sample;
- the quality of service, i.e., the initial telephone contact, the advice received, the ordering process, the availability of the product sought, the delivery of the sample order and the price.

THE APPROACH

As one of the companies being assessed sponsored the study, it was crucial to put in place a research approach that would minimize the risks of bias. A four-prong strategy was implemented.

First, one consultant was hired to select two geographical areas (counties) in the United States and to purchase samples for these areas from each of the five suppliers, as a "mystery shopper". Completely RDD samples of 1,000 numbers were purchased in March 2004 from the five companies for each of Johnson County, Iowa (an urban setting) and Nicholas County, Kentucky (a rural setting). The first consultant made a number of notes regarding the quality of service using a systematic observation grid. This consultant also combined all telephone numbers received from all five suppliers into a unique list; this list

was sent anonymously to a fieldwork company. The consultant and the fieldwork company did not know each other's identity.

Second, in late March 2004, the field house dialled each telephone number received from the first consultant without knowledge of which sampling company supplied the telephone number. Each number was dialled up to six times to determine whether it was resi-



dential and whether it fell within the geographical limits of the county. Once the fieldwork was completed, the field house received the code of the sample supplier(s) that had delivered each number – still without knowledge of which company corresponded to which code. This field house performed quantitative analyses of efficiency and precision of the samples.

Third, a statistician, unknown to the first two, received the reports from the first consultant and from the field house and prepared a synthesis of the results – still not knowing which sample supplier corresponded to which supplier code in the data sets.

Finally, an audit firm reviewed the entire research process to ascertain whether the protocols were followed and to weight the risks of bias in the research process. Its conclusions were positive.

THE RESULTS

Efficiency of the samples

In this study as in many market research studies, residential households were the target. Everything else being equal, a sample containing more residential numbers would be preferable to one containing fewer – the former would produce lower data collection costs.

Table 1 presents key results from the

vided a higher proportion of in-service numbers than other suppliers in their urban sample and a higher proportion of residential numbers than other suppliers in their rural sample.

Pre-dialling improved the position of supplier 'd' for in-service and residential numbers, bringing it up to supplier 'y' numbers and bettering them in the urban sample but not in the rural sample. However, this was achieved at the cost

Table 1 – Field work results

	What's better?	Supplier 'd'	Supplier 'l'	Supplier 'y'	Supplier 'x'	Supplier 'h'
Johnson County, Iowa (urban)						
% in service in whole sample	Higher is better	57	69	75	69	61
% residential ¹ in whole sample	Higher is better	44	57	58	56	52
% in service in pre-dialled sample	Higher is better	84	79	84	–	75
% residential ¹ in pre-dialled sample	Higher is better	74	65	66	–	66
% in service in excluded sample	Lower is better	21	9	16	–	8
% residential ¹ in excluded sample	Lower is better	4	0	0	–	0
Nicholas County, Kentucky (rural)						
% in service in whole sample	Higher is better	48	45	64	46	60
% residential ¹ in whole sample	Higher is better	43	42	61	41	46
% in service in pre-dialled sample	Higher is better	65	54	78	–	60
% residential ¹ in pre-dialled sample	Higher is better	63	50	76	–	56
% in service in excluded sample	Lower is better	10	6	6	–	7
% residential ¹ in excluded sample	Lower is better	2	1	1	–	2

Note: suppliers randomized for anonymity and did not offer pre-dialling of numbers.

¹ Includes confirmed residential numbers, residential answering machines, no response (no response being potentially residential)

dialling operation carried out by the field house on each sample. The most effective sample would be one:

- with the highest proportion of in-service numbers in the sample or, alternatively, the highest number of in-service numbers in the pre-dialled sample (an automatic filtering service was offered by four of the five suppliers);
- with the highest proportion of residential numbers in the sample or, alternatively, the highest number of residential numbers in the pre-dialled sample;
- with the lowest number of in-service or residential numbers in the sample excluded, based on pre-dialling as these should have remained in the sample (i.e., lowest false negative error).

The results of the study indicate that different companies provided samples of different efficiency. Supplier 'y' pro-

vided a higher proportion of in-service numbers than other suppliers in their urban sample and a higher proportion of residential numbers than other suppliers in their rural sample.

Globally, the efficiency of the supplier 'y' sample was significantly better than that of other suppliers in the rural sample and in the original (before pre-dialling) urban sample. Pre-dialling saved the day for other suppliers in the urban sample.

Precision of the samples

Telephone number samples must be precise in two ways:

- first, they should avoid including numbers that fall outside the geographical area of interest (which adds to survey costs but can be recouped by confirming the geography at the beginning of the questionnaire); and,
- second, they should avoid includ-

Table 2 – Percentage of numbers ascertained falling outside the county

	Supplier 'd'	Supplier 'l'	Supplier 'y'	Supplier 'x'	Supplier 'h'
Johnson County, Iowa (urban)	1	0	0	2	1
Nicholas County, Kentucky (rural)	0	6	7	7	0

Table 3 – Percentage distribution of samples numbers by exchange for Nicholas County, Kentucky

Telephone exchanges	Percentage distribution of the numbers delivered					Ideal distribution according to...	
	Supplier 'd'	Supplier 'l'	Supplier 'y'	Supplier 'x'	Supplier 'h'	Supplier 'd'	Supplier 'y'
234	–	–	–	–	–	0.1%	–
289	100%	91.7%	94.7%	90.9%	100%	96.0%	91.5%
383	–	6.0%	2.7%	3.7%	–	0.1%	5.9%
484	–	2.2%	2.0%	5.4%	–	3.2%	2.5%
987	–	–	0.5%	–	–	0.2%	0.3%
(606)247	–	–	–	–	–	0.4%	–

Table 4 – Differences between the samples received and the ideal distributions

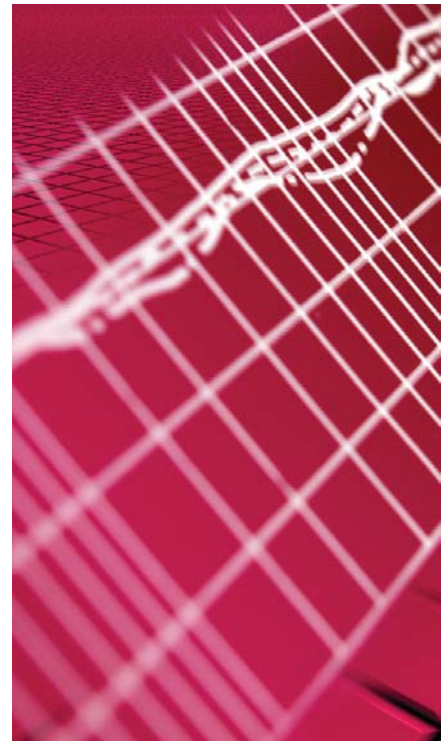
	Supplier 'd'	Supplier 'l'	Supplier 'y'	Supplier 'x'	Supplier 'h'
Compared to the ideal distribution from supplier no. 1					
Johnson County, Iowa (urban) ERRONEOUS	1%	1%	1%	1%	1%
	92	35	210	17	15
Nicholas County, Kentucky (rural)	3%	0%	0%	0%	3%
	27	55	10	44	27
Compared to the ideal distribution from supplier no. 3					
Johnson County, Iowa (urban)	0% ¹	0%	0%	0%	0%
	166 ²	69	26	174	173
Nicholas County, Kentucky (rural)	9%	0%	0%	0%	9%
	113	0	21	14	113

¹ Percentage of the county telephone numbers in exchanges completely excluded from the sample; 0% means less than 0.5%
² Mean square distance between the distribution of sample telephone numbers and the ideal distribution.

Table 5 – Average rank achieved in each service area (a low value is a better rank)

	Supplier 'd'	Supplier 'l'	Supplier 'y'	Supplier 'x'	Supplier 'h'
Initial contact (# of people before reaching the person responsible, minutes to obtain the person responsible, quality of interpersonal contact)	3.3	1	2	4	2
Order and advice (quality of advice, limits noted voluntarily, limited noted after prompting, payment modes, ease of payment by cheque, delivery promised)	2.8	2.5	1.2	2.5	3
Product (product range, availability of the product sought, availability of pre-dialling of numbers, available file formats, delivery modes, information available in the sample file)	1.2	1.8	1	2.3	2.2
Order fulfilment (delivery within promises, product conform to specification)	3	3	3	1	2
Price (order price, minimum cost for small samples, cumulative cost for additional samples within the same study)	3	3	1	3.5	4.5
Global average	2.7	2.3	1.6	2.7	2.7

Note: within each of the five service areas, between two and six criteria were observed; suppliers were ranked from best to worse (1 being best) on each criterion with tied performances being given the same, upper rank; average ranks are reproduced in this table; these averages are averaged globally, thereby giving equal value to each service area in the final figure.



ing numbers that are within the chosen geographical area (which is an error that cannot be recouped during the survey and that can create statistical bias).

Table 2 reports on the first of these criteria. Suppliers 'd' and 'h' did best in terms of avoiding bleeding at the limits of counties. The others did as well in the urban setting but were less precise in the rural area. Where the information could be obtained from household members, it was found that 6% to 7% of numbers fell outside of the rural county limits.

It is more difficult to determine whether valid telephone exchanges were excluded from the samples – the second criterion. The problem lies with exchanges that are not entirely subsumed within the area; if, for example, 50% of the telephone numbers in an exchange belong within the county and 50% outside, one can choose to include the exchange (but increase the risk that numbers outside the area will be included in the sample) or to exclude it (but reduce the coverage and precision of the sample).

Table 3 describes the choices made by the five suppliers with regard to the



simpler of the two counties used here (Johnson County's table is too large to be presented here but its results are similar). Suppliers 'd' and 'h' delivered numbers only in one exchange – which, arguably, encompassed at least 91% of Nicholas County¹. Only supplier 'y' provided sample numbers in four exchanges, even though the three suppliers who provided information to that effect indicated that there were at least four exchanges present in this county.

Table 3 also shows what two suppliers consider to be the ideal distribution of exchanges for Nicholas County. How supplier 'd' developed this distribution is unknown but supplier 'y' developed this distribution based on zip codes for the county (from ZipInfo 2003) and information on the match between zip codes and exchanges (from Axiom and Telcordia). As the next table shows, establishing this ideal distribution is no simple task and can lead to diverging conclusions.

Table 4 reports the quality of the match between the samples received and the ideal distributions from suppliers 'd' and 'y'. Based on the latter ideal distribution, supplier 'y' provided the

most inclusive urban sample, whereas supplier 'l' produced the most inclusive rural sample, followed by suppliers 'x' and 'y'.

The ideal distribution from supplier 'd' (the only other one available) indicates that supplier 'y' produced the most inclusive rural sample. Meanwhile, it suggests (incorrectly, as will be shown) a poor performance from supplier 'y' on the urban sample; however, these conclusions are ill-founded since the ideal distribution of supplier 'd' was incorrect for one important exchange (335): it claimed that there were only two households in the exchange, whereas this study contacted 73 valid numbers in that exchange alone. Therefore, supplier 'y' was correct in including this exchange in its ideal distribution.

Quality of service

Most aspects of the quality of the service provided are qualitative in nature. It is possible, nonetheless, to rank suppliers from best to worse. Table 5 reports the average rank achieved by each sample supplier within each of the service areas observed.

Supplier 'y' came out as the leader in

service quality. It shined in the areas of product offering, ordering, professional advice and pricing. Supplier 'l' was stronger in the initial contact.

CONCLUSION

So, are all samples of telephone numbers created equal? The clear answer is no: service quality varies from supplier to supplier, but, most importantly, the efficiency and precision of the samples are also variable. Some suppliers emphasize efficiency at the expense of precision. Others take a more balanced approach. Some suppliers have developed methodologies that make their samples markedly more effective than their competitors – and apparently with a continuing concern for precision.

Benoît Gauthier, is the President of Circum Network Inc., he can be reached at (819) 770-2423 or gauthier@circum.com.

¹ Three suppliers provided estimates of the numbers of households listed under each exchange in the county and the proportion of each exchange falling within the county.