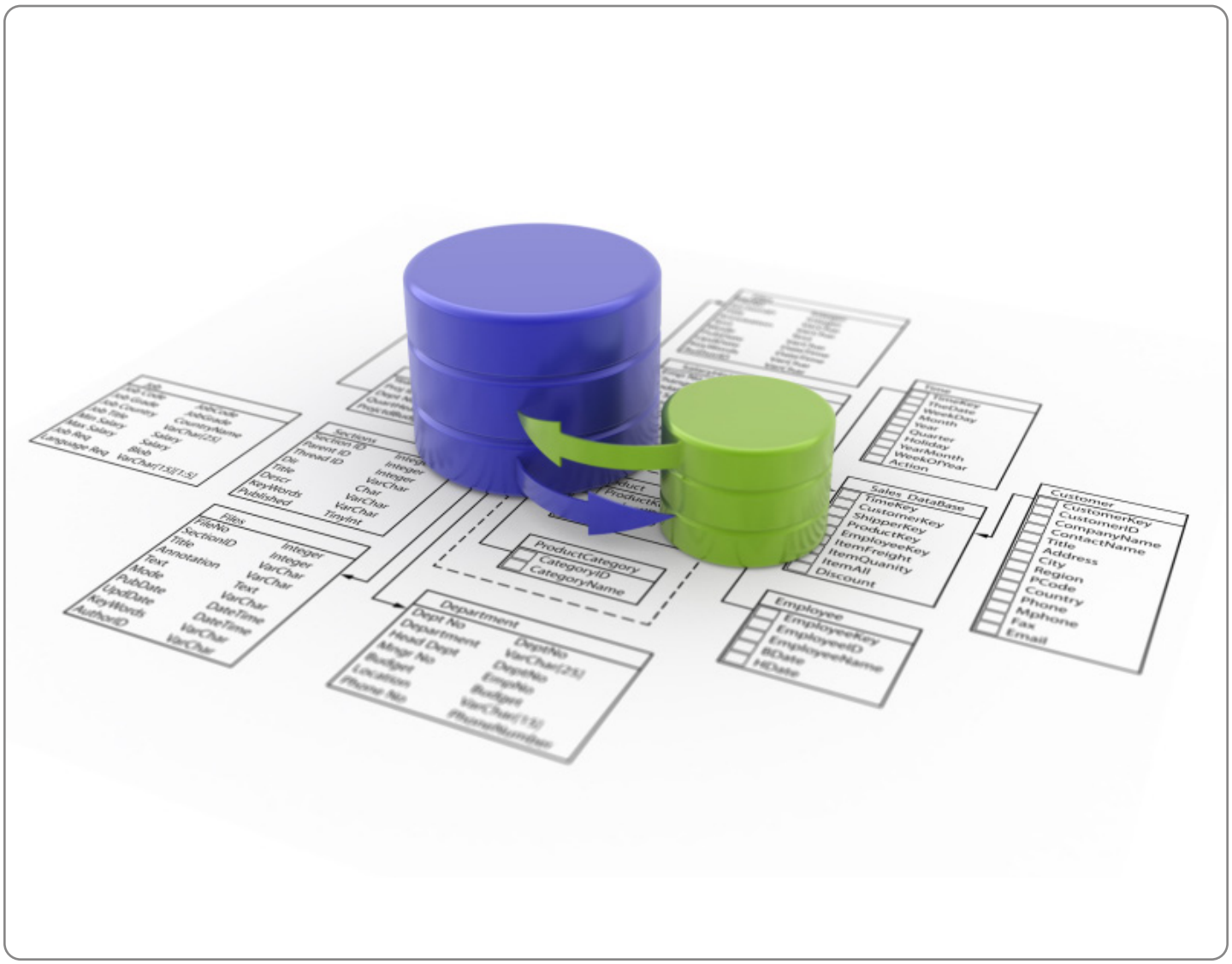
Integrating Behavioral Data & Online Survey Research



A Joint White Paper by Semphonic and iPerceptions

semphonic

iperceptions

Executive Summary

Behavioral analytics and online survey research, often referred to as Voice of Customer (VoC) analytics, are both established disciplines in today's enterprise online environment. As research tools, however, they've been largely siloed from each other. That's unfortunate because the combination of actual behavior trails along with survey respondent information has a powerful synergy that can answer many research problems unavailable to either.

The common-place wisdom that behavioral analytics provides the "what" and VoC analytics provides the "why" is an oversimplification. It's more precise to say that the combination of behavioral and survey response data helps the analyst protect against errors caused by misinterpreting cause and effect – allowing for more accurate and more incisive analysis.

In this white paper, we'll begin with the basics of integration between behavioral and VoC tools. We'll show how integration fundamentally changes the research program – and how this impacts thinking about sample size. We'll show how important it is to increase take-up/completion rates and the methods that can accomplish that. Finally, we'll showcase some real-world examples of analysis using the combination of behavioral and VoC data and how the two combine to drive powerful and actionable learnings not available from either source used individually.

Introduction to the Companies

Semphonic is the largest independent web analytics consultancy in the United States. Its practice is focused on helping large enterprises across a variety of industries to use web analytics and online measurement to improve their online operations on marketing. A full-service consultancy, Semphonic helps its clients with web analytic implementation and infrastructure issues, developing sophisticated reporting systems, and the actual analysis of behavioral data using advanced methods and techniques. Its clients run a variety of tools including Omniture, WebTrends, Unica, Coremetrics and Google Analytics as well as data warehousing and statistical analysis tools dedicated to online behavioral research. Semphonic has offices in Boston, New York, Washington D.C. and Portland and is headquartered in the San Francisco Bay Area.

Semphonic's focus has always been on behavioral data and analysis. Founded by experts from the credit card database marketing industry, the company's view has always been that no fact about a customer or visitor on the web is more significant than what they actually do. The relentlessly anonymous nature of most data on the web has made this behavioral focus essential in web analytics. And over the course of more than a decade, Semphonic has refined a set of techniques for taking web behavioral patterns and tying them to broader issues of site efficiency, customer intent, and marketing success.

iPerceptions is a leading web-focused Voice of Customer analytics provider, whose unique online intercept survey has been implemented across thousands of websites. Its webValidator Continuous Listening solution, 4Q product suite, Web Analytics Solution Profiler (WASP) and proprietary iPerceptions Satisfaction Index (iPSI) turn millions of data points into easy-to-understand strategic and tactical decision support for its clients. With advanced integration capabilities, real-time reporting, and global satisfaction index, the iPSI, iPerceptions is measuring online satisfaction and providing insight to help optimize the entire website experience. iPerceptions has offices in New York, Toronto, Montreal and London, U.K.

iPerceptions was a pioneer in behavioral and online survey integration and over the years has refined its techniques to provide both survey push and behavioral pull capabilities. The company has developed a flexible, easy-to-use interactive dashboard that enables clients to combine, analyze, and act on a mix of quantitative and qualitative feedback, for additional context and deeper insight. Among its many features, the dashboard automatically tests for statistical significance to avoid erroneous conclusions. iPerceptions clients include such well-known brands as InterContinental Hotels, Mazda, Dell, Harvard Business Review, and Monster Worldwide.

Introduction to the Problem

The analysis of behavioral data and the analysis of VoC data are each, in their own right, powerful and entirely appropriate techniques for measuring the effectiveness of the online channel and improving its efficiency. Each has carved out a significant set of problems for which the techniques are particularly well suited.

Behavioral data has been very effective in guiding tactical improvements to site operations. Functional, Real Estate, Use-Case, Funnel, and Key Correlates analysis are just a portion of the analytic techniques that Semphonic has found successful for helping companies identify usage problems on the website and surface potential fixes. These types of problems and solutions have a direct and measurable impact on behavioral measures of site success including engagement, orders, revenue and leads.

VoC data, on the other hand, has proven particularly valuable for providing insight into the website's impact on customer satisfaction and brand awareness and position relative to the competition. It has also proven effective in helping measure the importance and success of in-page elements and tools. Questions such as whether or not people like or value the images on a website may be nearly impossible to answer behaviorally but can emerge from VoC analysis. Finally, survey research has been the primary means by which marketers tie their understanding of offline behavior and customers to the online world. VoC provides demographic and attitudinal information that puts "flesh-and-bones" around online metrics.

As powerful as each technique has proven to be, they both have limitations. Experienced behavioral or VoC analysts know that there are types of problems very difficult to solve with either method.

One of the most intractable analytic issues with both behavioral and VoC analysis is figuring out how much preexisting intent is responsible for demonstrated behavior. Was a visitor successful on the website because the content was effective or were they simply determined to succeed? Was a visitor satisfied with a website because of their experience or did they simply arrive with a strong disposition to be pleased?

Here's an example. Many sites these days are adding significant social/feedback components to their site. Visitors can comment on articles and start discussion threads. They can review products. They can rate content. They can Tweet their views. In most cases visitors who do these things view more pages and are more satisfied than visitors who do not. Does this mean that social interactions drive page consumption and satisfaction?

Unfortunately, it doesn't.

Correlation is not causation, as virtually every analyst has eventually discovered to their chagrin. Social interactions may be correlated to increased page views or increased satisfaction because visitors who do these things are inherently more engaged. This problem is ubiquitous in analysis

problems and applies just as much to studies of why things fail as why things succeed.

Overview of Benefits to Integration

It turns out, however, that by using a combination of behavioral and VoC data, many problems of cause and effect can be controlled for. As in the example earlier, when it's possible to create a control group of visitors who viewed similar amounts/types of content and then compare the satisfaction of those who used social/feedback components with those who did not, it becomes possible to isolate, with a much higher degree of confidence, the real impact of satisfaction. And when one is able to control for those who used these components, their demographics, attitudes and visit intent, the behavioral impact can be much more sharply defined.

Because problems of cause and effect are ubiquitous, it turns out that nearly every analysis problem in behavioral and VoC analytics benefits from having the data from each combined.

The existence of independent systems of measurement drives value in other areas as well. Many researchers are uncomfortable with the quality of their sampling and unsure how representative their online survey results are of the total online population. Building a behavioral profile of your online sample is a powerful tool for showing that your sample is accurate or identifying places where it isn't. If you find that you are under or over sampling natural search sourced visitors or customer support visitors you can take corrective action. Without behavioral data, you have no means of ever discovering if you have a problem in the first place.

There are also cases where the combination of behavioral and VoC data add value in different phases of an analysis. A behavioral study may reveal that certain pages perform poorly, for example Functional routers (pages whose primary purpose is to move visitors to the right places in the website). But knowing the pages are poor performers doesn't provide insight into why they are not working well. By matching bad routes to visit intent, it's often possible to shed light on why visitors failed in certain places; this can't be done with either system (behavioral or VoC analytics) on its own. Some industry experts have described this process by saying that behavioral analytics tells you the "What" and VoC tells you the "Why." While this is true, it's also important to acknowledge that both behavioral and VoC research can identify significant potential issues in your online relationship with your customers. But that each can have limitations on its ability to tell you the full extent of what the problem is or why it's occurring.

Integration Methods

Methods

There are three methods of combining VoC and behavioral data:

Method 1: Send the VoC data into your behavioral analytics tool

Method 2: Send the behavioral data into your VoC analytics tool

Method 3: Send both behavioral data and VoC data into a full-scale statistical analysis application (or data mart)

Naturally, there are advantages and limitations to each method. By sending VoC data into you behavioral analytics tool, you benefit from the rich set of pathing and web segmentation tools that are available in those solutions. Moving the behavioral data to the VoC tool often gives you easier access to more general statistical analysis methods.

Combining the two in a full-scale statistical analysis application or database gives you almost unlimited analysis possibilities but is, of course, more work.

In some cases, you may consider using more than one method, since behavioral tools are not designed to analyze survey data, and vice versa.

Real-Time and Batch Integration

When iPerceptions sends VoC data to a behavioral analytics tool, as in Method 1, it automatically passes the following variables in real-time:

- Respondent ID
- Survey ID
- Self reported task completion
- Overall satisfaction rating
- Self reported purpose of visit

The variables are passed after each survey is completed in real-time to the behavioral analytics tool, where they are stored as custom variables. Real-time integration is extremely useful and supported by nearly every behavioral tool. Though it provides a limited subset of survey variables, these "golden" questions are the ones you are most likely to use in any combined analysis.

Integrating all survey responses to the visitor, however, can only be done in batch mode. Batch integration is similar to real-time integration in that you pass the Respondent and Survey ID to the behavioral tool in real-time. These become the keys for a batch join of all survey responses. Your survey vendor will then create a feed that moves all of the survey respondents and their associated responses to the behavioral tool. This can be done on a one-time basis but is more typically setup as a regular, scheduled task. Most enterprise behavioral analytics tools will support this type of batch data enhancement.

Impact of Integration

When you integrate behavioral and VoC data, you'll find that you've changed the research program for both your behavioral data and your VoC data. You've opened up a whole new set of possible analysis projects and tasks – many of which will not be explicit or understood when you do the integration.

Changing the research program on the VoC data has a profound impact of your survey research because it impacts one of the most important variables in opinion research – your sample size. This has less impact on behavioral data, because it is rarely sampled. However, if you do most of your behavioral research in a tool that does sample (tools such as GA, Discover, and Explore all sample in at least some cases), you'll to be aware of this on the behavioral side as well.

Sample size is determined by two factors – what questions you want to answer and how much confidence you want to have in the answer you get. When you design a survey, you've laid out a particular research program. It can still be fluid, of course, because sample size is impacted by the degree to which you want to cross-tabulate one or more questions. If you are only interested in frequency counts for a single variable, your sample size only has to be large enough to give you confidence around any one answer. If you want to cross-tabulate a variable like gender or income, then you need to have a large enough sample to give you confidence when you answer questions about the attitudes of say, males, or if you are cross-tabulating multiple variables, males with incomes between 50-75K.

The finer grained your cross-tabulations, the larger the sample necessary to support analysis.

Now here's the real rub. With VoC data, you generally have a pretty good sense of which cross-tabulations matter and you'll rarely be interested in anything finer than a 3-way cross-tabulation with fairly substantial cell percentages. So it's fairly easy to calculate the sample size you need and it's often not that large.

When you integrate your VoC data with a behavioral stream, however, you've suddenly opened up the potential for many, many more cross-tabulations. So many more that you generally can't plan for them and you'll have no reliable method of answering one of the most basic questions in survey research – how large should your sample be?

There is no magic bullet here. Without a definite research program, there is no way of deciding just how large your sample needs to be. At any moment, you may (and probably will) generate research problems for which your sample is inadequate.

So contrary to much of the traditional thinking about survey sample size, we think there's a simple rule to keep in mind. When you're working with integrated behavioral and VoC data and an open research program, size does matter and the larger the sample the better.

Of course, this immediately introduces a real-world constraint. Nobody wants to impact the user-experience of a website by over-sampling. So you will be faced with a genuine trade-off: balancing the value of additional survey data against the cost of collecting it.

Fortunately, this isn't a completely zero-sum game. There are some other factors to consider when you think about survey size: take-up rates and completion rates.

Take-up Rates and Completion Rates

How many people do you have to ask to take a survey to get a survey? Whatever the answer is, it's not one. This ratio between asks and respondents is your take-up rate. In the section earlier, we pointed out that integrating behavioral and VoC data opens up your research program and puts a premium on a large sample size. But you can increase your sample size without asking a single additional person – if you can improve your take-up rate.

If you can change your take-up rate from 1 in 20 people to 1 in 10, you've doubled your sample size without asking a single additional person.

Take-up rates aren't fixed. They are impacted by a variety of factors including when/how you ask, how many questions you ask (and what you say about how long the survey will be), and even visitors past experience with the survey tool you're using.

It isn't all about take-up rates either. For every visitor who starts a survey, how many complete it? And how many give you full and complete answers? The ratio of starts to full surveys is your completion rate and it's another metric that can have a profound impact on your sample size.

Most survey analysis only uses complete surveys. This doesn't change when you integrate with behavioral data. In some cases, you might not even get integration except for finished surveys. But if you can double your completion rate, you've double your effective sample size – without asking a single additional visitor to participate in a survey.

When and How to Ask

How and when you prompt online visitors to provide post-experience ratings and reviews will ultimately affect take-up and completion rates, as

well as the validity of your survey results and the perception of your brand.

Nobody likes being interrupted – so don't interrupt potential customers in the middle of their visit with a solicitation for feedback that was clearly triggered by some action they took. Research shows that 'survey interruptions' yield four to five times fewer respondents than on-arrival invitations and significantly increase the aggravation among all of the visitors involved, whether they participate or not.

By inviting on arrival, you avoid interrupting visitors during their experience. Think of the difference you feel when entering a retail outlet and receiving a polite, "Welcome, is there anything I can help you with today?" Compare this polite greeting to a sales person following you around the store and rushing over once you actually pick up an item.

Moreover, inviting feedback during or at the end of a website visit introduces a negative bias into the results, as these forms of invitations have been proven to garner more negative feedback. The timing affects visitors' motives to take the survey and people are quicker to complain than they are to compliment.

On-arrival invitations produce a representative sample of both positive and negative feedback, as the commitment to provide feedback after the visit, occurs before the experience itself. Further, research has shown that the positive to negative ratio among those who participate in on-arrival surveys is equal to those who don't participate at all.

Survey Length

The number of questions and the amount of time it takes to complete your survey both have a big impact on completion rates. There is a school of thought that it doesn't matter if your completion rate is low provided you get enough respondents to analyze. But as soon as you add behavioral data into your research program, that school of thought goes out the window. Each survey respondent who doesn't finish a survey may be impacting your ability to answer some question that uses the combined data. So there is a real – and unavoidable – trade-off between collecting more information from each respondent and having more respondents.

It's always been easy to add questions to surveys. Everybody wants their own piece of the research and there may not have been a compelling reason to say no. Now there is a compelling reason. Most organizations discover that half of the questions they ask are used rarely, if at all, in real-world analysis. But there has been little incentive to remove those questions. After all, they might be handy someday, right?

Get rid of them.

It isn't just the number of questions. The easier you make it to fill out the survey and the faster and more comprehensible your survey is to complete, the higher your completion rates are going to be.

Avoiding survey fatigue is critical to ensure high completion rates. This can be achieved by reducing the number of screen interactions and ensuring that respondents only see relevant questions. One way to reduce the number of clicks is to present one question at a time and use auto progression technology so that a single choice selection automatically calls up the next logical question, without the need to scroll or click "Next." With the exception of multi-select questions, one screen interaction and no scrolling is the best way to optimize survey responses. It also allows for advanced branching techniques, so that respondents are only asked questions that are relevant to their situation and valuable to you as a company.

Web Analytics Behavioral Sampling

Sampling may be a much bigger issue for the VoC program than the behavioral program, but it isn't completely isolated to that side of the house. There are some very popular behavioral tools that sample including Google Analytics, Coremetric's Explore (for unlimited queries) and Omniture's Discover (depending on your volume).

When a behavioral analytics tool uses a sample, it's generally a pretty massive one. You might see samples like 1:3 or 1:5 in your tool. But while this almost invariably produces a fine sample for behavioral analysis, it does mean that your survey sizes can get chopped down right out of the gate.

In other words, if you join VoC data to behavioral data and you collect 1,000 surveys a month and you're analyzing that sample using a 1:5 sampling rate, you're down to 200 surveys in your actual data even before you start cross-tabulating.

This is disastrous when it comes to analysis. So if you're behavioral tool samples your data, we strongly recommend pulling the data out and analyzing it elsewhere when you do the integration. For tools like Omniture, you can also consider using unsampled approaches to the data such as Data Warehouse or ASI (you can use Discover on the ASI segment).

Examples

The integration of behavioral and VoC data isn't particularly complex or expensive. In most cases, it takes only a few weeks and is quite easy on the budget. But how compelling are the benefits? In this white paper, we've talked about the difficulties in analysis that combining the data can overcome. But this discussion is technical in nature – and while an analyst may immediately grasp the essence of the problem and the power of the solution, that doesn't necessarily translate into a real organizational push to get the job done.

Below are a couple of case studies that illustrate how the integration of behavioral and VoC data drives real value.

Fix or Find – A Common Tool Dilemma

The Problem

VoC data for a leading online retailer showed that frustration with product information was a consumer stated reason for not following through to transaction – a barrier more common than price concerns and one that generated lower levels of satisfaction than other issues for abandoning. Given the size of the barrier and the population's position in the sales cycle, the ROI for a resolution was potentially quite large. Open-ended feedback from the survey data pointed to the understanding of option choices as one of the key elements of this type of failure.

Of course there was a tool designed to aid in option selections that was designed to provide users with the information at the right moment relative to this issue. Was the tool broken? Was it even being used? Was the right next step to fix the tool or make it even more prominent on the site? This "fix or find" dilemma is common when survey data raises issues about some aspect of site performance.

The Analysis

The behavioral data showed that conversion levels were higher among shoppers using the tool. But it couldn't tell us if it was reaching the right population – the shoppers with information concerns about option selection. Causation or correlation? Tool-users might simply be the subset of the population most likely to buy and not the real target audience at all.

By combining VoC and behavioral data, we

isolated the group of visitors facing this barrier and compared their future conversion (0-5 days) depending on whether or not they had used the tool. Subsequent conversion for those that used the tool was almost double those that did not. What's more, behavioral data revealed that a users citing the information barrier as their reason for abandonment were much less likely to have used the tool.

The Solution and Reflections

Being able to identify the impact of a tool, against the audience it was suppose to assist, was the simple value in this case. It led to an easy decision to the Fix or Find dilemma. The tool wasn't broken, the access paths to it were. Note the pattern of interaction between VoC and behavioral data in this analysis. Survey data surfaced the problem. Behavioral data provided the detailed information about the underlying experience. The combination of the two allowed us to set up a control group that actually tested the underlying question (fix or find). This is typical of how the two systems generally interact. Either may surface a problem. Each can be used to deepen the analysis. The combination of the two helps generate a control group that provides a much more reliable answer to the analysis problem.

Tool Effectiveness Study

The Problem

A technology site had developed two different tools that were similar in purpose but quite different in function. The company wanted to know which of the tools was better for users. The site itself was not transactional nor was the function of the tools to drive sales. The tools were also quite difficult to distinguish in words – making it nearly impossible to ask visitors which tool they'd actually used. How could the company decide which tool was better?

This basic problem is repeated over and over in analysis problems. How can you decide which of two tools, experiences, content loops, customer support pages, etc. is better? In some cases, A/B testing may be the best solution. But testing isn't always practical navigationally and, where a tool's performance is primarily measured by satisfaction outcomes it may not even be feasible.

The Analysis

One of the keys to this analysis was the recognition that one answer might not fit all. With any two tools or site experiences, there's a pretty good chance that each will work better with certain types of visitors and visit types.

Our survey data had two critically important pieces of information: visit type and satisfaction. What it didn't have was any information about which actual site experience related the two. On the other hand, our behavioral solution knew exactly which tool the user had used – it even knew if they'd used both. It had all of their previous and subsequent behavior. What it didn't have, was an easy way to decide which of the two tools had worked better.

We tried some behavioral measures (subsequent pages, exits, time, etc.) but they were ambiguous and controversial. It was unclear whether, for example, viewing more pages after the tool-use was good or bad. Sound familiar?

Fortunately, the behavioral solution was already capturing the integrated survey data.

Interestingly, we found that while the two tools produced rather different distributions, the overall satisfaction score they produced wasn't markedly different. However, when we broke this down by visit type we found that while many of the visit types had similar patterns, several showed distinctly different patterns. There were clear indications that for some visit types, each tool was distinctly better.

The Solution and Reflections

Is one tool better than another? This question is incredibly common. It applies to any case where you have multiple types of content that serve similar purposes: this may be video vs. text, alternative product selection methods, different customer pages, multiple search experiences, etc. Nor is this type of analysis limited to pure comparisons. Is it worth paying for third party content? Did a paywall impact satisfaction and engagement? Was a flash experience worth the extra expense of its build? If you've ever tried to answer these questions with either VoC or behavioral data, you know that it can be extremely challenging. In this case, behavioral cues simply weren't sufficient to measure success.

Survey data lacked the essential bridge between intent and satisfaction – the actual user behavior.

But the combination of the two data types transformed the analysis. It gave us dramatically higher confidence in identifying which tool was more impactful and which visit types might be tuned. This is a paradigmatic case of data-driven decision making. Prior to the analysis, it was subjective opinion that ruled – opinions about which tool was better based on anecdotal evidence and gut instinct about whether the difference actually mattered. After the analysis, the organization had the necessary information to easily decide when, where and why it made sense to invest in each tool.

Summary

Data siloes exist everywhere throughout the enterprise. Research is no exception. Online survey research and behavioral analytics have traditionally resided in separate places and been used by separate teams. There have been many industry experts who have touted the potential benefits of integrating these two research streams, but their arguments have been mostly hypothetical and have often misstated the actual benefits.

Nevertheless, most vendors on each side of the issue have developed basic integration methods that make it relatively easy to accomplish a high-level of data integration. There are three integration methods, including the ability to send VoC data into your behavioral analytics tool, send behavioral data into your VoC analytics tool, and send both datasets into a third party statistical analysis application.

Regardless of which integration method you choose, you'll find that the combination of the two data streams provides analysts with a powerful tool for answering questions raised by each system or proposed as a research program. The ability to test your online sample and to create valid control groups that eliminate or reduce the potential to misinterpret cause and effect in each system make the combination particularly useful.

You'll also find, however, that integration creates new demands on the data. It's impossible to foresee all the questions and sub-segments you're likely to be interested in from a behavioral perspective. Having enough VoC data to analyze those segments may mean increasing your sample rates. Inevitably, for any site without very large volumes, there is some trade-off between increased sample sizes and usability. Most enterprises want to limit the number of "asks" they have to do.

Because of this, it's important to address issues around take-up rates and completion rates. By improving these rates, you can increase your effective sample size without impacting your user experience.

What's the bottom line? Few integration opportunities in Web analytics are as simple, compelling and potentially rewarding as the combination of VoC and Behavioral data.

Contact Us

iPerceptions
Info@iperceptions.com
or 1.877.796.3600

Semphonic
info@semphonic.com
or 1.888.517.2171