

Cyborgs vs. Monsters:

Assembling Modular Mobile Surveys to Create Complete Data Sets

By Edward Paul Johnson, Director of Analytics, SSI; Lynn Siluk, Vice President, Marketing Sciences, Gongos Research;
Sarah Tarraf, Research Director, Gongos Research

This research was presented at the 2013 CASRO Online Research Conference

We live in a world where the supply of consumer opinions is eroding to ever-increasing demand. Add in the mindset shift brought about by smart devices, and respondents expect even more control over the survey experience to stay involved in the research process. They want shorter, more relevant surveys on topics they enjoy. Research buyers, on the other hand, want the most “bang for their buck” and often try to squeeze as much content as possible into a single survey (Suresh and Conklin, 2010). While researchers generally have final say in questionnaire design, respondents have important weapons to fight back—namely abandon rates, satisficing behavior, and future participation. One way to help solve this conflict is by breaking surveys into modules.

Kantar and Lightspeed jointly proposed modularizing mobile surveys at the 2012 CASRO Online Research Conference to give more control to the respondents (Johnson et al., 2012). The modular approach breaks down a typical 20+ minute survey into modules that are 5 to 8 minutes long. At the end of each module the respondent is rewarded for the time spent so far and given an opportunity to exit the survey or to continue (for an additional incentive). Johnson was able to show that these modular surveys reduce respondent fatigue and the perception of survey length. This finding supports the theory that respondents desire and enjoy more control over the survey experience.

Modularization can be done in two different ways: within-respondents and across-respondents. The within-respondent modularization method has the same respondent answering all modules. This is the preferred

Respondent matching involves looking at the individual survey modules and finding the respondents in each module that most closely resemble each other and then combining them as if they were “one” respondent.

method of modularization for researchers because it leaves no holes in the data. It does exclude those that only complete a few modules. As these partial respondents are given incentives this type of modularization incurs costs that do not directly provide any data for the researcher.

Across-respondent modularization is more efficient because it utilizes all respondents who take any modules. However, end users of the data can’t help but have concerns over large amounts of missing data. Particularly, advanced research techniques, such as segmentation, driver analysis, or even simple cross-tabulations and correlations—analytic staples of market research—cannot easily be computed with missing data.

Kantar addressed the missing data problem at the 2012 CASRO Online Research Conference in conjunction with uSamp (Smith et al., 2012). They conducted an across-respondent modular test where the test groups were

only allowed to do one module and compared them to a control group that had no modularization. They then used hot-deck imputation to create a synthetic data set that filled the gap of missing data between modules. Because the correct hook variables to link answers in one module to another did not exist within this synthetic data set, the hot-deck imputation did not yield results similar to the control group. After incorporating auxiliary data from a previous study as hook variables, they got connecting hooks that better preserved the correlations existing in the control group.

Methods

Leveraging a three-way partnership that included a global beverage manufacturer, Gongos Research and SSI set out to compare synthetic data sets (cyborgs) to a stitching algorithm that takes the partial surveys and fuses them together to make complete respondents (monsters). In particular we want to find out how effective the module environment was at engaging respondents and whether the data fusion techniques would allow for advanced analytics.

We used SSI's proprietary panels to test the two methodologies. The panelists were first screened via a six-minute screener survey to identify smartphone owners willing to take a survey on their mobile device and ensure the topic of the survey applied to them (consumers of frozen carbonated beverages who frequent convenience stores). Those who qualified were then randomly divided into four different groups as shown below.

- **25-minute Control Online Survey**
- **Online Modular Survey (3 modules)**
- **Mobile Modular Survey (3 modules)**
- **25-minute Control Mobile Survey**

The control online and control mobile surveys were a standard web survey. No stitching or hot deck imputation was needed for these groups. The mobile surveys were not done in an app environment, but were

programed to accommodate the smaller screen of the mobile device. All of the mobile surveys verified that the user agent string contained a mobile device to make sure that the respondent was accessing the survey through the appropriate medium.

The modular survey designs randomly assigned each respondent to a starting module. Once the starting module was completed, respondents were given the opportunity to exit the survey with a partial incentive or complete an additional survey module to receive an additional partial incentive. Because respondents were able to exit the survey after each module, we included questions from every section within each module rather than breaking subjects or sections of the survey into their own modules. Also, we identified key variables ('hook' variables) we thought were important and always asked them in the starting module. For example, a battery of 17 potential reasons for drinking frozen carbonated beverages was split into 6 'hook' items that were always shown in the starting module and the other 11 were evenly split across the three modules. Likewise, the 25% of the essential non-grid questions were always shown in the starting module while the other 75% were assigned to a specific module. The design resulted in a starting module that was double the length of subsequent modules a respondent could complete. This incentivized them to complete multiple modules so that we would have less missing data.

Many advanced analytics require no missing data. In this experiment we are trying to group respondents using a cluster analysis, and then build a segmentation algorithm to accurately classify respondents into their assigned cluster. We compared two different techniques for filling in the holes created in the data. One technique is to create synthetic data (cyborgs) while the other makes assumptions that certain people are identical (monsters). Both are described in detail below.

Cyborgs

Hot deck imputation involves intelligently estimating what the respondent "would have said" had they



answered the question. Much like manufacturing mechanical prosthetics for missing limbs, we generate synthetic data for the missing parts of the data set and deliver results based on cybernetic (part organic and part mechanical) respondents. The key to hot-deck imputation is to impute data based on the characteristics of the individual (the equivalent of molding the implant to fit the individual patient) rather than using random or mean imputation (which would be like mass manufacturing implants and just hoping that they would fit). Hot-deck imputation preserves the total sample size of respondents who completed at least one module of the original survey. Gongos Research determined that hot-deck imputation (assigning the respondent an answer from a randomly selected similar or “hot” respondent) would be an effective imputation technique to deal with missing data in segmentation analysis.

Monsters

Respondent matching involves looking at the individual survey modules and finding the respondents in each



module that most closely resemble each other and then combining them as if they were “one” respondent. This process results in a complete data set with a sample size smaller than the total sample size of respondents who completed at least one module of the original survey. This process is like creating a Frankenstein monster—stitching together the legs and arms of one partial respondent to the torso and head of another. The obvious challenge in stitching respondents together is making sure you have respondents who are similar to each other. You don’t want to stitch a small head to a body with long arms! Ultimately, the desired result is an organic creation with no artificial or synthetic body parts derived from estimation.

Findings

We asked two satisfaction questions, “How satisfied were you with your experience in this survey?” and “How interesting was the topic of the survey to you?”, to every respondent both in the screener and the follow-up survey. We used the screener survey as a baseline and looked at the difference in each experimental cell to its baseline. The results are shown in Table 1 below. In almost every instance the follow-up survey’s average satisfaction score was lower than that group’s corresponding screener survey satisfaction score, likely because it was significantly longer and more involved than the screener survey. While none of the differences are statistically significant, in every instance the modular survey designs performed better than the control survey design.

Table 1. Satisfaction Ratings

	Control Online	Control Mobile	Modular Online	Modular Mobile
Average Difference Satisfaction	-0.06	-0.14	-0.03	-0.02
Average Difference Topic	-0.01	-0.07	0.02	-0.01

Another good indication that the module design was successful in keeping respondents engaged is that a majority (72%) who complete at least one module ended up taking all three modules. This finding underlines the importance of incorporating within-respondent modularization. Because we allowed within-respondent modularization and many respondents chose to do multiple modules our final data set had significantly fewer holes.

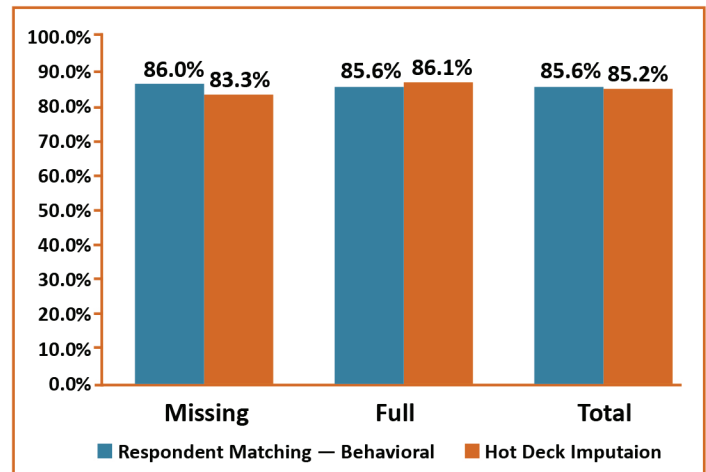
We also found that both techniques (hot-deck imputation and respondent matching) filled in the missing data in an effective manner to allow for cluster analysis. The results of the online control group were interpreted first, revealing an optimal three-segment solution. These results were then used as the baseline for the comparison of the clustering results across the mobile modular groups.

To ensure that the resulting cluster algorithms were similar we took the algorithms from two mobile modular data sets and applied them to the control data. We then compared those results to the segment results from the control data algorithm to see if the classification rate was high enough to confirm that the algorithms defined the same clusters. Both of the methods for filling in the missing data create very similar segmentation algorithms to the control algorithm. The respondent matching algorithm reclassified the control group with 90% accuracy. This rate was not statistically different from the hot-deck imputation algorithm, which reclassified the control group correctly 88% of the time.

We also wanted to know how cluster assignments worked in reverse, so we took the algorithm from the control group and applied it to both the respondent matching and the hot-deck imputation data. In particular, we wanted to know if the reclassification rate was better or worse among the group with missing data. Figure 1 shows the results of this analysis. The respondents for whom we used imputation or matching reclassified at the same rate (83%-86%) as those with the full data set. This indicates that the hot-deck

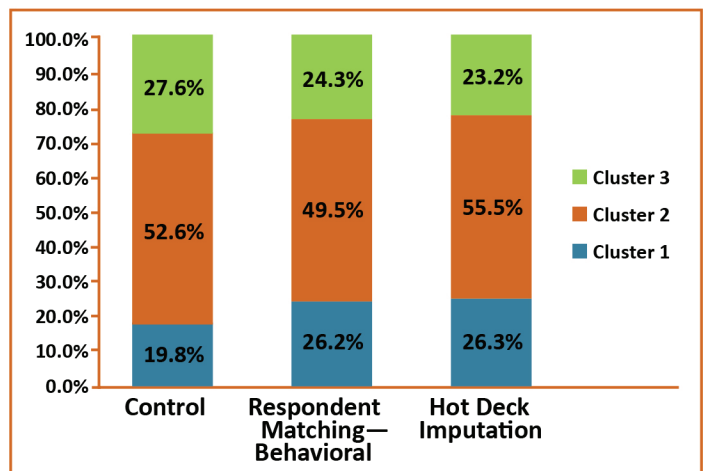
imputation and respondent matching algorithms are doing a good job informing the segmentation algorithm.

Figure 1. Reclassification Rates of Control Algorithm to the Test Group



Despite these two findings, we did see a skew in the segment size especially with cluster 1 (Figure 2). As we have eliminated the possibility of a difference in algorithms with high reclassification rates, we are left with a difference in the make-up of the respondents. While many mobile device users said they were willing to take the survey, many tried to access the survey through a standard computer and were not allowed to participate. We theorized that this caused the skew in the segment sizes across the online control and mobile modular samples.

Figure 2. Reclassification Rates of Control Algorithm to the Test Group



Conclusions

Highly positive outcomes were revealed during each complex phase of our research. Modular survey design requires both critical thinking and buy-in across all stakeholders; and data fusion techniques hold the key to link modern consumers with our need to engage them in research.

Both techniques to deal with the missing data worked very well for segmentation. They each have their own advantages. Giving a data set to a client that includes hot deck imputed data and running standard significance testing can inflate the Type 1 error rate, creating 'false positives'. Some potential ways to adjust for this inflation include weighting the data back down to the proper sample size, adjusting the significance level, or confirming the results with a separate study. The benefit of hot deck imputation is that no data is thrown away and you do not make an assumption that two different people are really one. With respondent matching, our sample size is slightly smaller, as gaps in the data are filled with another respondent's answers. This allows for a complete data file with no missing data for a client, without the risk of an inflated Type 1 error.

One aspect of modularization, however, that will benefit from further testing is the appropriate number and length of the modules. For example, can the modules be as small as one or two questions? These approaches need to be tested for both across-respondent modularization and within-respondent modularization. They also need to be tested across different sample sources such as consumer email lists rather than online panels. The research here uses online panelists that are already primed to survey taking and have a higher tolerance for long surveys regardless of the medium.

The hook variables used to deal with the missing data also need additional testing. There are a multitude of potential hooks that could be used. An example is data from previous wave surveys which help decrease the need for hooks inside the survey itself in tracker

research. Panel profile data could potentially be used to create the hooks for those that have complete profiles. Even passively collected cookie tracking data can be used to link similar respondents with their permission. This data can create the links or matches before the survey commences and drastically increase the efficiency of the modules and decrease survey length.

While the near-term goal is to satisfy the needs of both end-users of research and respondents, there is greater promise at play here. One of the reasons we are rigorously putting mobile to the test is the inevitability of the future of communication. It's no secret that much of our communications are untethering at a very rapid pace. When you add global into the mix, you have a world that is increasingly connecting through customizable smart devices and apps. Mobile surveys, in theory, could become a catalyst that forces us to give up the mass collection of large amounts of data from one person at a single sitting that we as researchers have become accustomed to with PC-based surveys.

If we keep pushing harder and positively—both with how we engage consumers in our research and how to synthesize the data on the backend—we will strengthen our place in a world being influenced by social networks, apps and gamification. Once both sides experience the real advantages of the modularization of surveys, it will be hard to turn back the clock.

Works Cited

Suresh, Nallan and Conklin, Micheal. "Designed to Engage." *Quirks*. July 2010.

Smith, Renee, Kotzev, Ivailo, Miller, Chuck, and Kachhi, Dinaz. "Modularizing Surveys to Meet Today's Respondent Challenges." *2012 CASRO Online Conference, March 2, 2012, Las Vegas, NV*.

Johnson, Alex, Kelly, Frank, and Stevens, Sherri. "Modular Survey Designs for Mobile Devices." *2012 CASRO Online Conference, March 2, 2012, Las Vegas, NV*.

About the Authors



Lynn Siluk

Lynn and her team of eight marketing science professionals fuel Gongos Research's analytical thinking and data-driven insights to enable clients to gain an innovative edge in the marketplace. Beyond statistical analysis, study design and reporting, Lynn consults on high-level projects, offering client teams different ways to approach complex methodologies and evaluate data. She joined Gongos Research in 2006 as Research Director, Analytics, and prior to that she spent 17 years with Ford Motor Company. Lynn holds an MS in Statistics from Oakland University and BS in Psychology and Statistics from Central Michigan University.



Edward Paul Johnson

Paul specializes in online sampling design and advanced analytic methods. Paul's research-based approach to identifying and recommending solutions positions him well to investigating panelist behavior; oversee data processing efforts; improve online data integrity and provide advanced analytic support to clients. Paul also actively pursuing a research agenda on survey research methods and applications of conjoint methodology and mobile data collection. His research on research efforts have been featured regularly at ESOMAR, AAPOR, CASRO, MRA and Sawtooth events. Paul holds an MS in Statistics and BS in Actuarial Science from Brigham Young University.



Sarah Tarraf

Sarah conducts a variety of statistical analyses, ranging from basic correlations and significance testing to advanced statistical modeling, to provide insight to business clients regarding customer needs and opinions. In addition, she provides statistical consulting to internal and external clients. Sarah has experience with a variety of statistical software packages, including SPSS, Stata, Latent Gold, XLStat, and SAS, as well as Microsoft Excel and Access. Specialties include: Statistical analysis, including descriptive statistics, basic inferential statistics, and advanced statistical modeling; Survey research and evaluation, including instrument development, data collection (CATI, web, face-to-face, and mail), statistical analysis, and report writing. Sarah earned a BA in Economics from Michigan State University and a Master's from Wayne State University.