DC-AAPOR & Washington Statistical Society 2017 Summer Preview/Review Conference

Program and Abstracts

July 24, 2017
Bureau of Labor Statistics Conference Center
2 Massachusetts Avenue, NE, Washington DC

Thank you to our conference sponsors!

Gold Sponsors

IMPAQ International LLC · RTI International · Westat

Silver Sponsors

AIR American Institutes for Research · GALLUP

PROGRAM

8:00-8:30	**Registration**	
8:30-8:40	**Welcoming & Opening Remarks**	
8:40-10:10	**Session 1A:** **Incorporation of Mobile Technology into Data Collections** (Rooms 1 & 2)	**Session 1B:** **Health-Related Issues and Perspectives** (Rooms 7 & 8)
	Session Chair: *Rebecca Medway, American Institutes for Research* "When can we call?" Experiment to Assess SMS Text to Prompt Response Across Cultures - *John Lee Pratt Holmes, Qatar University Social and Economic Survey Research Institute* Examining Predictors and Correlates of Consent to Receive Text and Facebook Message Reminders - *Caitlin Deal, American Institutes for Research* Dropdown Response Options in Mobile Surveys - *Elizabeth Nichols, U.S. Census Bureau* Does Typographic Cueing Improve the Processing of Information from Survey Questions on a Mobile Device? - *Brian Falcone, U.S. Census Bureau*	Session Chair: *Doug Williams, Westat* Insure My Beating Heart: Exchange Participant Characteristics and Satisfaction with the System - *Zac Auter, Gallup* The Effects of Racial Priming on Support for Funding of Nutritional Assistance Programs - *Charles Dahan, Stanford University* Measuring Attitudes on Abortion – Time for Another Look - *Kate Stewart, Conway Strategic* Human Trafficking Prevalence Estimation: the USA and Globally - *Davina P. Durgana, SIT Graduate Institute DC*
10:10-10:35	**Coffee Break & Posters (Reception area & Hallway)**	
	Measuring Changing Household Complexity and Diversity by Race and Ethnicity from 2000-2010: A New Approach - *Laurie Schwede, U.S. Census Bureau* Complex Households and the Undercount of Young Children in the 2010 Census - *Eric Jensen, U.S. Census Bureau* Changing Immigration Laws and Their Impact on Attitudes within Qatar's Sponsorship System - *John Lee P Holmes, Qatar University Social and Economic Research Institute* Comparison of Model-Based to Design-Based Ratio Estimators - *James R. Knaub, Retired U.S. Government* National Health and Nutrition Examination Survey (NHANES) 24-hour Dietary Recall Sample Weights: Day of the Week Adjustment - *Te-Ching Chen, National Center for Health Statistics* An Experiment with Varying Survey Cover Letter Paper Color and its Effect on Survey Response Rates by Mode - *Amy L. Djangali, IMPAQ International* Reducing the Number of Pages of a Mail Questionnaire in a Mail/CATI Mixed Mode Survey: How Does It Affect the Response Rates and Costs of Each Mode? - *Michelle Cantave, IMPAQ International*	

10:35-12:05	**Session 2A:** **Strategies in Survey Design and Data Collection** (Rooms 1 & 2)	**Session 2B:** **Advancements in Weighting Adjustment and Sampling** (Rooms 7 & 8)
	Session Chair: *Jessica Holzberg, U.S. Census Bureau* Trading Precision for Reliability: Time Frames and Estimation Strategies in Answering Questions on Alcohol Consumption - *Meredith Massey, National Census for Health Statistics* Time after time: Exploring the Impact of Fixed-start, Variable-length Reference Periods - *Maura Spiegelman, National Center for Education Statistics* Exploring Sampling Techniques to Reduce Respondent Burden - *Yijun Wei, George Mason University and USDA National Agricultural Statistics Service* Smartphone GPS Applications as a Mode of Travel Survey Data Collection - *Joann Lynch, Resource Systems Group*	Session Chair: *Chris Moriarity, National Center for Health Statistics* I Conducted a Nonresponse Follow-up Survey; Now What Do I Do? - *Phillip Kott, RTI International* Evaluation of Nonresponse Weighting Adjustment Methods for a National Survey of People Living with HIV - *Davia Spado, ICF* Geosampling and Drones for Household Estimation - *Karol Krotki, RTI International* Redirected Inbound Call Sampling (RICS) – Pilot Test Results and Caller Reactions - *Sarah Dipko, Westat*
12:05-1:20	**Lunch on Your Own**	
1:20-2:50	**Session 3A:** **Polling, Elections, and Political Attitudes** (Rooms 1 & 2)	**Session 3B:** **Confidentiality Pledge Assessments** (Rooms 7 & 8)
	Session Chair: *Anna Brown, Pew Research Center* Hanging up on F2F? Mode Comparison for Polling in Eastern Europe - *Marta Churella, Department of State* Latinos in the 2016 Election: Was There a Trump Effect? - *Ana Gonzalez-Barrera, Pew Research Center* Early Voters and Late Deciders: Vote Choice by the Timing of the Vote and the Vote Decision - *Christopher Fleury, GfK* Detecting Politically Motivated Tampering With Workers' Labor Income In Survey Data - *John Angle, The Inequality Process Institute* Behind the Badge: A Nationally Representative Survey of Police Officers - *Renee Stepler, Pew Research Center*	Session Chair: *Jennifer Edgar, Bureau of Labor Statistics* Do They Read It? Using Paradata to Evaluate the Extent to Which Respondents Attend to Confidentiality Pledge Language - *Casey Eggleston, U.S. Census Bureau* Do People Understand It? Cognitive Interviewing Assessment of Confidentiality Pledges for Household Surveys - *Stephanie Willson, National Center for Health Statistics* Do Establishments Understand It? Cognitive Interviewing Assessment of Confidentiality Pledges for Establishment Surveys - *Cleo Redline, National Center for Education Statistics* Exploring Respondents' Perceptions of Data Confidentiality and Enhanced Cybersecurity - *Amy Swallow, Bureau of Labor Statistics* One Size Fits Most? Lessons Learned by Using Multiple Methods to Study Confidentiality Pledges -

		Heather Ridolfo, USDA National Agricultural Statistics Service
2:50-3:15	**Coffee Break & Posters (Reception area & Hallway)**	
3:15-4:45	**Session 4A:** **Evaluation of Survey and Census Issues through Pretesting** (Rooms 1 & 2)	**Session 4B:** **The Use of Internet and Non-Probability Panels** (Rooms 7 & 8)
	Session Chair: *Stephanie Willson, National Center for Health Statistics* Is a Proxy Response Good Enough? Using Paired Cognitive Interviews to Assess the Accuracy of Proxy Responses - *Mary Davis, U.S. Census Bureau* Bilinguals as Cognitive Interview Respondents: Identifying Potential Data Quality Problems and Refining Recruitment Criteria - *Lucia Lykke, U.S. Census Bureau* Conducting Cognitive Interviews with Young Children - *Mahi Megra, American Institutes for Research* Multi-use Field Testing: Examples from the 2017 Census of Agriculture Dry Run - *Jaki McCarthy, USDA National Agricultural Statistics Service*	Session Chair: *Jonathan Katz, U.S. Census Bureau* Using Online Panel Surveys to Estimate Population-level Health Statistics - *Rosalynn Xin Yang, Westat* Who Won't Respond to Your Online Probability Panel Survey? - *Alexandra Brown-Breslin, Federal Reserve Board of Governors* A Deeper Dive on the Digital Divide - *David Dutwin, SSRS* Methodological Considerations in the Use of Web Probing for Questionnaire Evaluation - *Stephanie Fowler, National Institutes of Health* Implementation of a Youth Health Survey via the Internet - *Richard Lee Harding, ICF*
4:45-4:50	**Closing Remarks**	
5:00	**Happy Hour – The 201 Bar, 201 Massachusetts Ave NE (2nd & D St)**	

ABSTRACTS – CONCURRENT SESSIONS

SESSION 1A: Incorporation of Mobile Technology into Data Collections

Presentation Title	**"When can we call?" Experiment to Assess SMS Text to Prompt Response Across Cultures**
Presenter Name	**John Lee Pratt Holmes**
Presenter Email	**johnleeholmes@gmail.com**

This paper examines a split sample experiment that alternates a Short Message Service (or SMS) advance text contact across seven language and three socioeconomic groups surveyed by wireless telephone in Qatar. Qatar's unique advantages include: a small size that facilitates population enumeration, universal cellular coverage, a multicultural respondent base, and a legal environment that permits such public contact techniques. Two years ago Qatar University's Social and Economic Survey Research Institute (SESRI) began conducting its own quarterly survey by phone to measure variation in quality of life and consumer confidence among its Qatari, high income expatriate ("white collar") and low income migrant ("blue collar") populations. Conducted using a random sample drawn from cellular telephones in Qatar, the survey has high response rates compared to surveys in the U.S. and Europe. Nevertheless concern about survey fatigue within a small population and the increasing use of avoidance features on smartphones has led SESRI to pre-emptively encourage positive response. To appreciate whether such a tactic would have the desired effect, advance text contact has been launched via a split sample experiment. Half of the sample receive an advance text sent out using Qatar University's SMS system and half do not. We measure whether this leads to increased blocking of SESRI's contact number, variance in contact and cooperation rates, and differences in the composition of the respondent pool by language group and respondent type (Qatari, white collar, blue collar).We believe this experiment has value in verifying whether such a method changes the composition of those who participate in the survey to better match the actual population. Because of the multiple language groups, it is applicable to broader cross-cultural survey quality concerns for other countries that may also seek to leverage widespread use of cellular and smart phone technology to understand popular preferences.

Presentation Title	**Examining Predictors and Correlates of Consent to Receive Text and Facebook Message Reminders**
Presenter Name	**Caitlin E Deal**
Presenter Email	**cdeal@air.org**

Past research shows different people respond to different types of survey reminders. Email reminders may be more effective than telephone calls for some individuals, while the reverse may be true for others. As a result, surveys often try to use a variety of approaches to contact sample members when possible. However, these efforts may be hampered by missing contact information or lack of participant agreement to be contacted in certain ways that require prior consent, making it impossible to use these contact methods for all of the sample members. If these methods are successful at increasing response, and certain types of sample members are more likely to receive them, then using these additional contact methods may actually have the potential to bias survey results even if response rates ultimately increase as a result of these contact methods. This study will examine mechanisms of consent to be contacted in particular ways and subsequent effects on nonresponse bias using a recent web survey of community college students. Sampled students were asked if they would consent to receive text or Facebook message reminders as part of a follow-up survey. Seventy-two percent of respondents agreed to be contacted by text, and twenty-five percent agreed to be contacted by Facebook message. Key survey outcomes and demographics will be examined as predictors of the likelihood of consenting to each of these methods. We will determine whether these predictors are different for the two contact methods. In addition, we will assess whether respondents who consent to be contacted by text or Facebook also provide higher quality data overall by looking at item nonresponse and straightlining. Finally, we will incorporate responses to the follow-up survey to conduct a nonresponse bias analysis.

Presentation Title	**Dropdown Response Options in Mobile Surveys**
Presenter Name	**Elizabeth Nichols**
Presenter Email	**lizabeth.may.nichols@census.gov**

Dropdown response fields or "pick-lists" are a type of automated design where the user can click or touch on a single field to bring up a list of possible choices. For long lists, the choices are in a scrollable frame. When the user selects a choice, the list disappears, and that choice is displayed back in the single field on the screen. For online surveys, dropdowns are often used when there are space constraints or when there is a long list of response options, such as when someone selects a state in the U.S. On smartphones, the design functions somewhat differently on different devices, depending on operating systems. In the Apple iOS system, the list displays at the bottom of the screen, with a type of spinner that only

shows a few possible choices at a time; while the other systems, like Android, function more similar to the way a PC would handle these fields. This talk will present findings from an experiment comparing the usability and measurement error associated with the drop-down response options using different designs.

Presentation Title	**Does Typographic Cueing Improve the Processing of Information from Survey Questions on a Mobile Device?**
Presenter Name	**Brian Falcone**
Presenter Email	**brian.falcone@census.gov**
As more people use mobile devices to complete Web surveys, it is critical to verify that current Web survey design practices carry over to mobile Web. One set of design practices involves typographic cueing, i.e., using different font styles and sizes to distinguish between the different parts of a question. A common practice is to bold the question stem to draw attention to it and help ensure that it is read first. Another common practice is to italicize instructions to facilitate filtering of information and allow them to be recognized quickly as instructional text. However, typographic cueing has not been tested in mobile Web surveys that typically present one question per page due to the smaller size of mobile device screens. It is possible that the reduction in the amount of information that can be presented at once might reduce or eliminate any benefits of this practice. This talk will present findings from a laboratory experiment (n=30) varying two factors: (a) typography of question stems (bolded or not) and (b) typography of instructions (italicized or not). The study questionnaire contains seven questions adapted from national surveys and will be implemented using a native smartphone app. To assess respondent performance with each of the four designs, we compare completion times, response accuracy when answering questions based on vignettes, and post-survey recall of the keywords in the questions. We expect that bolding the question stems will not affect completion time or accuracy because only one question is presented at a time. By contrast, we expect that italicizing the instructions may still assist in filtering which could lead to shorter response times but possibly worse accuracy and recall by providing a cue to respondents that this text may be scanned or skipped rather than carefully read.	

SESSION 1B: Health-Related Issues and Perspectives

Presentation Title	**Insure My Beating Heart: Exchange Participant Characteristics and Satisfaction with the System**
Presenter Name	**Zac Auter**
Presenter Email	**zac_auter@gallup.com**
In 2008, Gallup and Healthways initiated a partnership to track and understand the key factors that drive well-being. The Gallup-Healthways Well-Being Index provides an in-depth, nearly real-time view of Americans' well-being. The Gallup-Healthways Well-Being Index is a nationally representative survey of adults currently living in the United States interviewed by telephone every day, 350 days per year. Since this study was initiated in 2008, Gallup has measured the insurance status of adults currently living in the United States as part of the Gallup-Healthways Well-Being Index survey. Gallup's longstanding trend on this important metric provides the ability to analyze the insurance rate prior to during and now following the implementation of the Affordable Care Act (ACA). In addition to tracking the insurance status of those currently living in the United States, Gallup closely monitors the percent of newly insured Americans that obtained their insurance through the exchanges, and public reaction to and support of the ACA including the perceived impact of the policy on Americans and their families. This presentation will provide researchers important insights into how the uninsured rate has changed over time and how satisfied or dissatisfied those who obtained their insurance through the exchanges are with the healthcare system. We will also provide details about the demographic characteristics of those who obtained their insurance through the exchanges, compared with those who obtained their insurance through other sources including their self-reported health status (excellent, good, fair or poor), age, gender and race/ethnicity.	

Presentation Title	**The Effects of Racial Priming on Support for Funding of Nutritional Assistance Programs**
Presenter Name	**Charles Dahan**
Presenter Email	**cdahan@stanford.edu**
The Supplemental Nutrition Assistance Program is one of the federal government programs offering the highest return on investment. While the majority of recipients are white, the media disproportionately portrays recipients as African-American. Additionally, conservative political attacks on the program regularly feature both implicit and overt racial cues. This study investigates three research questions: the relationship between racial priming and preferences for SNAP funding; the relationship between the trajectory of the economy; and two types of primes--religious participation and individualism—affect support for SNAP funding. Additionally, the effects of priming for race are tested to measure whether white respondents report increased perceptions of fraud perpetuated by recipients, and tests a thesis posed by	

Congressman Paul Ryan that parents who utilize school lunch subsidies care less about their children. The study utilizes a randomized online survey experiment with a convenience sample of undergraduates, with respondents receiving an actual New York Times article using sympathetic episodic framing to profile the effects of food stamp cuts on recipients. The race of the individuals profiled is altered to prime perceptions of the race of food stamp recipients. I find evidence that each independent variable—the media treatment, racial prejudice, ideology, and ethnocentrism--significantly affects support for food stamp funding, while racial prejudice, ideology, and ethnocentrism—but not the treatment-- affect the length of time an individual should be eligible for receiving food stamps, perceptions of fraud in the program, and support for Paul Ryan's contention that parents utilizing food assistance programs care less for their children.

Presentation Title	**Measuring Attitudes on Abortion – Time for Another Look**
Presenter Name	**Kate Stewart**
Presenter Email	**Kate@conwaystrategic.com**
The traditional abortion question which has been asked for decades by pollsters would have us believe that the country has barely moved on this issue since abortion was legalized in 1973. When attitudes on abortion are reported in the media, we most often hear that views on abortion are "closely divided" and have been "stable" for decades. Given the level of debate around abortion policy nationally, the dramatic rise in state laws on abortion access, and threats to shutdown the federal government over Planned Parenthood funding, it is time we take another look at how we are measuring abortion attitudes. Most major pollsters ask Americans to indicate whether their beliefs align with one of the following four categories: Abortion should be legal in all cases; abortion should be legal in most cases; abortion should be illegal in most cases; or abortion should be illegal in all cases. Some surveys also ask whether individuals identify with the label "pro-choice" or "pro-life" and increasingly have found that sizeable segments of the public, in particular younger Americans, identify as both. For example, Public Religion Research Institute recently released a survey, How Race and Religion Shape Millennial Attitudes on Sexuality and Reproductive Health, and found nearly half of millennials (age 18-35) identify with both "pro-life" and "pro-choice" labels (27 percent) or identify with neither label (22 percent). Only about one-quarter of millennials identify exclusively as "pro-choice" (27 percent) or as "pro-life" (25 percent). The paper will examine the conventional abortion measurement tools, raise questions about whether we need to review how we measure attitudes on abortion, and begin to provide ideas for ways to create tools which more accurately tap into the complex and, at many times, conflicting attitudes the public holds on abortion.	

Presentation Title	**Human Trafficking Prevalence Estimation: the USA and Globally**
Presenter Name	**Davina P. Durgana**
Presenter Email	**davina.durgana@sit.edu**
This session will present the current efforts to employ best practices in data analysis and protection in the United States and globally to estimate human trafficking prevalence. The speaker is Senior Statistician of the Global Slavery Index and works with original survey data from Gallup World Poll on modern slavery in over fifty countries. She is also a Forbes 30 under 30 for Science and the 2016 ASA Statistical Advocate of the Year for her work on vulnerability modeling and human trafficking. She will present her current and ongoing work to employ multiple methods to determine human trafficking prevalence and will discuss the field of applied statistics in human trafficking and the significant difference that good data science is making in this field.	

SESSION 2A: Strategies in Survey Design and Data Collection

Presentation Title	**Trading Precision for Reliability: Time Frames and Estimation Strategies in Answering Questions on Alcohol Consumption**
Presenter Name	**Meredith Massey**
Presenter Email	**wnx6@cdc.gov**
Asking questions with a shorter time frame has been shown to result in more accurate estimates of respondent behavior. This is because the longer the timeframe, the harder it is for respondents to remember specific events. However, asking respondents about their behavior within a shorter time frame may not provide accurate estimates of their long-term behavior. Therefore, survey researchers must make trade-offs between reliability and precision when designing questions to measure behavior. This paper discusses the relationship of question time frame to the estimation strategies used by respondents. Findings are derived from a study done by the Collaborating Center for Questionnaire Design and Evaluation Research, National Center for Health Statistics on survey items aimed at measuring behavior related to alcohol consumption. The study was a cognitive interview evaluation of alcohol questions intended for use on the National Health and Nutrition Examination Survey and the National Health Interview Survey. Ninety interviews were conducted. Results	

show that respondents use different strategies to estimate their alcohol consumption for 30 day versus 12 month recall periods. These strategies and their potential effect on survey estimates will be discussed in more detail.

Presentation Title	**Time after time: Exploring the Impact of Fixed-start, Variable-length Reference Periods**
Presenter Name	**Maura Spiegelman**
Presenter Email	**maura.spiegelman@ed.gov**

Reference periods are important both for maximizing respondent recall and balancing seasonal effects when collecting data on event counts. For some event types, respondents may naturally encode their experiences relative to a fixed starting point (for example, the beginning of a school year), and survey questions that use this fixed starting point can yield more accurate responses than reference periods with variable starting points and fixed lengths (for example, within the past two weeks). However, counts can only accumulate over time, so true values depend on when a question is answered; with fixed starting points, late responders are instructed to consider a longer reference period than early responders. This paper analyzes data from two surveys that use reference periods with fixed starting points and variable lengths: the 2010 School Survey on Crime and Safety (SSOCS), conducted February through June, and the 2012 National Household Education Surveys (NHES), conducted January through August. Both surveys ask respondents to count events that have occurred since a fixed starting point, for example, the number of unplanned fire alarms or the number of days a child has been absent from school since the beginning of the school year. Event counts will be examined both as straightforward counts and transformed to counts over a fixed, standardized reference period. Finally, similar early- and late-responders will be matched in order to determine whether response propensity explains differences in event counts or rates of events between these two groups. Results are expected to differ by survey item, for example, with conscientious early NHES responders indicating higher rates of attending school meetings and activities and lower rates of child absences than late responders, suggesting that transforming counts to rates may not be appropriate in all instances. This discussion of tradeoffs between recall, measurement error, and response propensity will be valuable to surveys that use fixed start, variable length reference periods.

Presentation Title	**Exploring Sampling Techniques to Reduce Respondent Burden**
Presenter Name	**Yijun Wei**
Presenter Email	**ywei2@gmu.edu**

The USDA's National Agricultural Statistics Service (NASS) is exploring sampling approaches that allow for coordination of multiple samples drawn within a year across the population in an effort to control the respondent burden. Most of these sampling techniques, which are both design-based and model-based approaches, utilize permanent random numbers (PRN) for the purpose of achieving the amount of desired overlap within a survey or between different surveys and, therefore, help reduce the respondent burden. However, little discussion comparing design-based and model-based approaches or examining a combination of these two approaches appears in the literature. Using a simulation study, we investigate different sampling strategies (a combination of sampling design and estimator) that utilize both design-based and model-based inferences. Our simulations are based on data from several USDA surveys. Results are presented.

Presentation Title	**Smartphone GPS Applications as a Mode of Travel Survey Data Collection**
Presenter Name	**Joann Lynch**
Presenter Email	**joann.lynch@rsginc.com**

Household travel surveys (HTS) are conducted in major metropolitan areas in the United States (as well as internationally). A typical HTS collects both demographic data and detailed travel diary data from all household members through a two-part survey process. The modes of data collection for an HTS were almost exclusively paper (mail-back) or telephone through the early/mid-2000's, at which time data collection became primarily web-based, while continuing to offer telephone response. Recently, some metropolitan planning organizations (MPOs) and state departments of transportation have adopted new technologies for HTS data collection using smartphone-based applications to collect both the survey data and the GPS travel diary data. By using a smartphone-based GPS application as the primary means of data collection for the majority of households (those that own smartphones), planning organizations can obtain more accurate and detailed spatial, temporal, and survey-specific data, particularly when compared to traditional data collection modes. Smartphone-based GPS applications also provide the opportunity to collect longitudinal and/or continuous data due to decreased respondent burden and administration costs. From 2014-2016, RSG conducted three comparable HTS efforts for North American MPOs including the City of Calgary (Alberta, Canada), the Puget Sound Region (Washington, USA), and the Triangle Region (North Carolina, USA). Across the three regions, a total of 11,800 households participated in a one-day travel diary either online or over the phone. For each MPO, a panel of households (approximately 6.4%) also completed a multi-day GPS study using RSG's smartphone-based GPS application, rMove™, on their personal smartphones in addition

to the one-day diary. The panel results provide a direct comparison of traditional and smartphone GPS data collection modes at the household-, person-, and trip-levels. Data collection modes are analyzed based on the data quality, data quantity, respondent burden, and costs associated with each methodology.

SESSION 2B: Advancements in Weighting Adjustment and Sampling

Presentation Title	**I Conducted a Nonresponse Follow-up Survey; Now What Do I Do?**
Presenter Name	**Phillip S Kott**
Presenter Email	**pkott@rti.org**

It has become a standard practice for a government statistics agency to conduct a nonresponse follow-up (NRFU) survey on the unit nonrespondents to an official survey. What to do next is not clear. A NRFU survey was conducted for the National Pilot of the Residential Energy Consumption Survey (RECS), an addressed-based sample survey of potential primary residences enumerated by web and mail rather than in person as has been done in previous RECS surveys. Virtually all unit (i.e., whole-record) nonrespondents to the National Pilot were sent a short mail questionnaire containing 18 key items from the full survey. Here, we first compare two ways of adjusting variables collected on the NRFU for unit nonresponse. In one, only the weights for respondents to the full National Pilot survey were adjusted to compensate for nonresponse using a calibration weighting procedure that assumes response to be a logistic function of variables known for the entire sample. In the other, only the NRFU-survey respondents' weights were adjusted for nonresponse using an analogous calibration weighting scheme, while weights for the respondents to the full survey were not adjusted. The resulting two national estimates for many of the NRFU variables were then compared. When the two were significantly different, the latter estimate was treated as unbiased and added as a calibration variable when adjusting (a second time) for unit nonresponse to the full sample. When they were not significantly different, both were deeded unbiased, and the mean of the two added as a calibration variable when readjusting for nonresponse to the full sample. The theory behind this practice and its repercussions are discussed.

Presentation Title	**Evaluation of Nonresponse Weighting Adjustment Methods for a National Survey of People Living with HIV**
Presenter Name	**Davia Spado**
Presenter Email	**davia.spado@icf.com**

In 2015, the Medical Monitoring Project (MMP) was redesigned to sample individuals diagnosed with HIV in order to provide national and local estimates of various characteristics of persons living with HIV. Weight adjustments for use with this new stratified multistage sample design needed to be developed. We compared weight adjustments generated from a weighting class model with those from a propensity weighting model. Two components of nonresponse were accounted for: non-contact and nonresponse among those contacted. Weight adjustments methods were applied separately to each to allow for differing predictors by nonresponse component. The propensity method has greater flexibility to incorporate continuous variables and more than two variables into the nonresponse adjustment. However, this flexibility led to only minor improvements in model fit. We examine how the magnitude and variance of key estimates differ between methods at each stage of adjustment. Adjusted weights from both methods resulted in similar weighted estimates with small differences in variances.

Presentation Title	**Geosampling and Drones for Household Estimation**
Presenter Name	**Karol Krotki**
Presenter Email	**kkrotki@rti.org**

Listing housing units in a sampled segment is often a laborious, lengthy, and difficult task, especially when maps are of poor quality or not available. Researchers have recently looked into using GIS and drones to not only map out and sample the geographical units but also to implement a low-cost listing operation to prepare for the final stage sampling, selection of households. Our initial attempt to use drones for listing and estimating housing units and number of households was conducted on a very limited basis in Galapagos Island. That study identified room for improvement in the use of drones and Geo-listing for surveys in developing countries. Furthermore, in that study we did not have access to on-the-ground listings so we were unable to evaluate the accuracy of the listings based on the drone technology. In an attempt to replicate, expand, and improve the methodology in a standard setting of a developing country, we conducted a followup study in Guatemala. In this paper, we discuss the use of Geo-sampling supported by drone aerial imagery to enumerate and sample respondents in randomly sampled Grid-cells. We compare listings, for the same areas, based on drone imagery with those implemented using on-the-ground listing. We also describe improved systems for extracting data from the aerial images to provide more accurate information for updating dwelling unit and household counts. For example, we

have made progress in identifying multi-story units and units that contain multiple households.

Presentation Title	**Redirected Inbound Call Sampling (RICS) – Pilot Test Results and Caller Reactions**
Presenter Name	**Sarah Dipko**
Presenter Email	**sarahdipko@westat.com**

This paper presents results of testing a new non-probability sample source of inbound calls, in essence a “river sample” of callers. Inbound callers interested in completing the survey can be connected with IVR surveys or to telephone interviewers. A 14.5 percent response rate (AAPOR RR3) was reported for the earliest test of RICS as a sample source for research purposes, with a survey conducted using IVR (Levine, Krotki and Bobashev, AAPOR 2016). The paper describes a ten-day RICS test conducted in July 2016. Live interviewers were used to conduct the study in order to better understand the circumstances and reactions of the callers in the RICS sample. What types of calls were being made when callers reached the survey request? How was the survey request and experience (for respondents) perceived by the callers? Assumed and observed rates are presented, and characteristics of the sample are compared to national estimates for demographics, geographic representation, and health measures. Cost and management issues are addressed such as out of scope calls (repeat calls from identical phone numbers or those from outside the United States) and significant fluctuations in the volume of the inbound call flow to interviewers. The survey experience for inbound callers is examined using survey data, paradata, and case notes combined with insight from an interviewer debriefing held after the study concluded.

SESSION 3A: Polling, Elections, and Political Attitudes

Presentation Title	**Hanging up on F2F? Mode Comparison for Polling in Eastern Europe**
Presenter Name	**Marta Churella**
Presenter Email	**ChurellaMU@state.gov**

Face-to-face interviewing has been the main mode for high-quality survey research across Eastern Europe. As cell phone penetration increases across the continent, possibilities for moving to dual-frame telephone polling with high levels of coverage have improved. The Office of Opinion Research conducted multi-mode experiments in Poland and Ukraine to examine the feasibility of moving from face-to-face to telephone polling. Both the in-person and phone polls were designed to be nationally representative, used the same questionnaires, were conducted during the same time period, and had similar sample sizes within country (n=1,000 in Poland and n=1,500 in Ukraine). Different field firms administered each survey. This paper will explore mode differences and examine data quality variance across a variety of metrics including attitudinal and demographic differences, weighting efficiency, response dispositions, length of fieldwork, overall costs and general representativeness of the national populations. A key research question is to gauge the stability of trend data from in-person to phone, and whether stable political trends will be interrupted by a change in mode.

Presentation Title	**Latinos in the 2016 Election: Was There a Trump Effect?**
Presenter Name	**Ana Gonzalez-Barrera**
Presenter Email	**agonzalez@pewresearch.org**

The 2016 election placed a great focus on the record 27.3 million Hispanic eligible voters and their political values and attitudes. At the beginning of his campaign, Donald Trump made provocative comments about Mexican and Muslim immigrants and Hispanics in general. The issue of immigration was also a focal point of Trump’s campaign. The 2016 presidential election drew strong responses among the Latino community overall, but did Trump’s comments motivate Latino voters more than in past election years? This analysis will be based on a number of Pew Research Center surveys done prior to the election, media exit polls done on the night of the election and the latest data from BLS and Census Bureau on civic engagement. Findings highlight that although a majority of Latino voters were engaged in the election and had discussed Trump’s comments about Hispanics; this does not appear to have had a strong effect on voter engagement or voter turnout. A month before the election, 69% of Latino registered voters said they were absolutely certain they would vote on this election, compared with 77% who said so in 2012. It is likely that in 2016 the number of Hispanic voters reached a record number, but voter turnout rate might have remained at levels similar to past elections or increased slightly, with only about half of those eligible to vote casting a ballot.

Presentation Title	**Early Voters and Late Deciders: Vote Choice by the Timing of the Vote and the Vote Decision**
Presenter Name	**Christopher Fleury**
Presenter Email	**christopher.fleury@gfk.com**

The 2016 presidential election provided many voters with the option of voting before Election Day. At the same time, the unique dynamics of this election – including the historic unpopularity of the pairing of major party candidates – meant that many voters were polarized early on into one camp or another while others were reluctant to commit to either camp. The confluence of these developments provides an opportunity to explore the interplay between when voters decided on a candidate and actually cast their ballot. This presentation explores these questions using data from the GfK Pre-Election Poll, which was conducted using GfK's KnowledgePanel® from November 4-8, 2016.

Presentation Title	**Detecting Politically Motivated Tampering With Workers' Labor Income In Survey Data**
Presenter Name	**John Angle**
Presenter Email	**inequalityprocess@gmail.com**

There is no reason to think that the labor income data in public use samples of the March Current Population Survey (CPS), a series dating back to March 1962, has been tampered with. So that series presents an opportunity to test the ability of a mathematical model of labor income statistics, the Inequality Process, to detect simulated tampering with labor income data. There are autocratic regimes that assign the production of surveys similar to the March CPS to a propaganda role. A recently elected would-be autocrat not yet able to rule through violence and fear may want a government statistics office to report gains in labor income to maintain political support. Such a regime's statistics office may wish to continue issuing public use samples of surveys like the March CPS to maintain a façade of integrity while it tampers with labor income data. There is a more convincing way of detecting tampering with data than scoffing at large increases in labor income. The Inequality Process (IP) implies functional invariances in labor income statistics. These have approximate mirror images in March CPS labor income statistics. Some forms of tampering with labor income data in a survey like the March CPS are evident in multiple violations of IP invariances.

Presentation Title	**Behind the Badge: A Nationally Representative Survey of Police Officers**
Presenter Name	**Renee Stepler**
Presenter Email	**rstepler@pewresearch.org**

Fatal encounters between black Americans and police in recent years have fueled a national debate over police tactics, methods and training. During the Spring and Summer of 2015, the Pew Research Center, working with the National Police Research Platform, conducted a nationally representative survey of 7,917 police and sheriff's officers employed in departments employing 100 or more sworn officers. While national surveys have explored public views about the police, race and the use of force, this survey of police officers helped to fill a knowledge gap on how police view these recent events and what they see as the key issues and concerns facing their profession. Overall, the survey found that police view their jobs are harder now following a series of fatal encounters between blacks and police. Large majorities of officers also say relations between police and blacks in their community have grown more tense and that officers in their departments are now reluctant to carry out some police duties. The project also found striking differences in the views of police and those of the general public on a range of issues, as measured in a companion Pew Research Center survey of the general public. This presentation will discuss in detail the main findings for our survey of police officers and the general public.

SESSION 3B: Confidentiality Pledge Assessments

Presentation Title	**Do They Read It? Using Paradata to Evaluate the Extent to Which Respondents Attend to Confidentiality Pledge Language**
Presenter Name	**Casey Eggleston**
Presenter Email	**casey.m.eggleston@census.gov**

As part of an interagency effort to evaluate the effect of potential changes to confidentiality pledge language used by agencies of the Federal Statistical System, the Bureau of Labor Statistics (BLS) and the Census Bureau collected indirect evidence about respondent reactions to variations of confidentiality pledge language incorporating new information about cybersecurity monitoring. In addition to the explicit measures included in cognitive testing and a web survey, paradata (specifically, eye movements and response times) was also collected to capture responses that may have been subconscious or that respondents did not self-report. Overall, both eye tracking and response time data indicated that, while participants did read the confidentiality language, they did not spend a significant amount of time on it and did not especially seem to fixate on the new versions of the pledge that contained information about cybersecurity monitoring.

This presentation summarizes the detailed findings and discusses implications for communicating with respondents.

Presentation Title	**Do People Understand it? Cognitive Interviewing Assessment of Confidentiality Pledges for Household Surveys**
Presenter Name	**Stephanie Willson**
Presenter Email	**zex8@cdc.gov**
This presentation describes how respondents interpret and understand different versions of confidentiality pledge language, focusing on household survey respondents in particular. The confidentiality language we tested incorporated requirements of the Cybersecurity Enhancement Act of 2015 into the standard CIPSEA language already common among Federal household surveys. Cognitive interviewing methodology was used to evaluate the new pledge. Fifty face-to-face, in-depth interviews were conducted; 30 were conducted by the U.S. Census Bureau and 20 by the National Center for Health Statistics. Interviews were designed to explore respondents' interpretation of and reaction to the new pledge language. Because the new cybersecurity language introduces the idea that government data will be more closely monitored than in the past, concerns existed that respondents may react negatively to the new pledge. However, while some did express cautious attitudes about what the new language meant, the preeminent finding was that respondents had generally neutral or even positive interpretations of and reactions to the new pledge language. Some believed that a high level of scrutiny of Federal data systems by the government already existed. Others believed increased monitoring was necessary and important as cyber threats have become more common. Additionally, various versions of the pledge were tested by each agency. It was found that respondent attitudes were influenced by specific wording. For example, versions of the pledge that used the word "protect" invoked mostly positive reactions while versions that used the word "monitor" tended to invoke more negative reactions. This and other specific findings will be reviewed in more detail.	

Presentation Title	**Do Establishments Understand It? Cognitive Interviewing Assessment of Confidentiality Pledges for Establishment Surveys**
Presenter Name	**Cleo Redline**
Presenter Email	**Cleo.Redline@ed.gov**
This paper discusses cognitive research that explored how establishment respondents understand and react to proposed revisions to agency confidentiality pledges, given that the Federal Cybersecurity Enhancement Act of 2015 has been enacted. Using similar protocols, four government agencies (Bureau of Labor Statistics, National Agricultural Statistics Service, National Center for Education Statistics and the U.S. Energy Information Administration) conducted 102 cognitive interviews with participants from different types of establishments, including private and public companies, farms, and schools. Participants were probed to determine if they understood the intended meaning of the new language so that agencies could assess the potential impact on their future data collections. Participants generally understood the proposed language. The research identified differences in interpretations of and reactions to the proposed text by establishment and respondent types. For example, farmers were more likely to be concerned about the new language, while energy companies preferred wording that specifically mentioned DHS. This paper will cover the testing approaches used, summarize the results and detail the advantage of conducting cognitive testing with unique respondent groups.	

Presentation Title	**Exploring Respondents' Perceptions of Data Confidentiality and Enhanced Cybersecurity**
Presenter Name	**Amy Swallow**
Presenter Email	**swallow.amy@bls.gov**
Protecting respondent confidentiality is a top concern for survey organizations, but determining how that protection is communicated to respondents can be challenging. Ideally, confidentiality pledges are easy to understand and reassure respondents that their data are kept confidential. Many government surveys have used the same confidentiality language for many years as part of the Confidential Information Protection and Statistical Efficiency Act (CIPSEA). However, the passage of the Cybersecurity Enhancement Act of 2015 resulted in a change to the way data are protected, requiring revisions to the confidentiality pledges. To better understand the impact of the revised confidentiality pledge, the Bureau of Labor Statistics (BLS) conducted research on people's comprehension of, and reaction to, both the current and revised pledge. A total of 864 participants were recruited from Amazon's Mechanical Turk to complete an online survey. Participants answered an initial question about how they believe the government keeps their data confidential, and then they were randomly assigned to read either the current or revised confidentiality pledge. Participants then answered a series of questions designed to assess their attitudes and comprehension of the pledge. We systematically coded participants' initial perceptions, attitudes, and comprehension of the pledge, as well as the types of concerns participants have about how the government keeps their data confidential. We found that participants' initial perceptions of how the government keeps data confidential varied widely. In contrast, participants did not seem to focus on either version of the	

pledge, with very few differences in comprehension or reactions by pledge version. However, differences in the level of concern about confidentiality participants reported led to group differences in data quality and comprehension of the pledge. We explore the types of concerns participants had about their data confidentiality, and discuss the implications this may have for future data collection as federal cybersecurity systems continue to evolve.

Presentation Title	**One Size Fits Most? Lessons Learned by Using Multiple Methods to Study Confidentiality Pledges**
Presenter Name	**Heather Ridolfo**
Presenter Email	**Heather.Ridolfo@nass.usda.gov**
In 2016, a workgroup consisting of representatives from ten statistical agencies was established to study respondents' expectations and reactions to changes in the federal confidentiality pledges. The expected outcome of this research was to develop a universal confidentiality pledge that would be used in all statistical agencies. This paper will describe the differences among the various studies (e.g., respondent type, methodology, and pledge wording), then highlight common findings. First, the analysis of paradata and web survey results showed that although respondents read the confidentiality pledges, they did not spend a lot of time doing so, and often could not recall information provided. Second, the cognitive interviewing assessments showed that participants understood that there would be additional cybersecurity monitoring, but there was inconsistency in the interpretations of who would be doing the monitoring and what would be done with their information. Finally, initial reactions to changes in the confidentiality pledge were neutral or positive, but those reactions shifted upon further reflection. Reactions to pledge changes also varied across survey populations, with the agricultural population having the strongest negative reaction. Ultimately, the panel concluded that a one-size-fits-all pledge was not possible. This presentation will highlight the careful balance that must be struck between a) transparency and clarity and b) respondents' expectations for confidentiality.	

SESSION 4A: Evaluating Survey and Census Issues through Pretesting

Presentation Title	**Is a Proxy Response Good Enough? Using Paired Cognitive Interviews to Assess the Accuracy of Proxy Responses**
Presenter Name	**Mary C. Davis**
Presenter Email	**mary.c.davis@census.gov**
Conducting a survey interview with a single household member who can accurately provide proxy reports for all other household members can offer cost benefits relative to conducting a survey interview with each member of the household individually (Blair, Menon, and Bickart, 1991). This research evaluates the feasibility of proxy interviewing through the use of paired cognitive interviews conducted with related or unrelated members of the same household, using an approach described by Blair et al. (1991). This research uses two recent cognitive interview projects (a survey on social attitudes and behaviors, and a survey on computer and internet use) from which we provide guidelines for conducting paired cognitive interviews to evaluate the feasibility of using proxy response, including recruiting and scheduling procedures for paired interviews with related and unrelated respondents, and potential challenges. This research also presents suggestions for analyzing data from paired cognitive interviews in order to determine the feasibility of using proxy interviews. This includes quantitative analysis, such as by calculating the rate of matching or near-matching responses, and analysis of qualitative feedback, such as identifying characteristics of questions that are easier or more difficult for proxies to answer, and identifying characteristics of household members or relationships that facilitate or impede the ability to accurately respond by proxy. Finally, this presentation will discuss overall benefits and shortcomings of using paired cognitive interviews to evaluate the usefulness of proxy interviewing, and potential applications for this research design.	

Presentation Title	**Bilinguals as Cognitive Interview Respondents: Identifying Potential Data Quality Problems and Refining Recruitment Criteria**
Presenter Name	**Lucia Lykke**
Presenter Email	**lucia.lykke@census.gov**
Examining whether testing survey materials with bilingual vs. monolingual respondents is effective is a recent area of interest in multilingual survey methodology. Studies on this topic investigate whether and how bilinguals and monolinguals differ in their understanding of the cognitive interview task or compare understanding of key survey concepts between monolinguals and bilinguals. Although the research shows that bilinguals can help to uncover survey translation issues as well as monolinguals in some cases, there is little research that systematically examines within-group variation among bilinguals' responses in multilingual survey research. Further, it is unknown whether measures other than survey comprehension problems may be valid measures of data quality from cognitive interviews. In this study, we use	

data collected from 27 cognitive interviews with bilingual and monolingual Spanish speakers to test a newly translated housing survey to examine whether and how participants use English in a non-English interview setting and comprehension issues on 15 specific survey items. We compare these measures across self-reported English language proficiency and language preferences before and after the interview. We consider the implications of using bilinguals of varying levels of English proficiency and language dominance for testing survey questions, with future implications for the recruitment of bilingual cognitive interview participants.

Presentation Title	**Conducting Cognitive Interviews with Young Children**
Presenter Name	**Mahi Megra**
Presenter Email	**mwmegra@air.org**
As a result of the Every Student Succeeds Act, education policy at the federal or state level may require schools and districts to survey children as young as second grade (aged 7-8) about school climate or classroom practices. However, many methodologists urge caution in surveying young children. In a review of the literature, Borgers et al. (2000) discourage surveying children from ages 4 to 8 because they may have limited language abilities, interpret questions quite literally, be more likely to respond in a way to please adults, and have shorter attention spans compared to older children or adults. If younger students will be surveyed as required by federal or state policy, existing surveys that are used for older students may need to be adapted by shortening the survey and simplifying the language. Cognitive interviews may be useful in this adaptation process. In such interviews, participants are asked to read survey questions and think aloud as they respond, reporting to an interviewer the cognitive processes they use in answering the questions. These reports can easily be used to improve survey questions. However, young children may not be developmentally able to explain how they arrived at a particular response. To date, no research has examined the utility of conducting cognitive interviews with young children. This paper will examine the effectiveness of cognitive interviews with 2nd-8th grade students in eliciting meaningful feedback that can be used to improve survey questions. The sample will include 35 students who were administered surveys about classroom practices and personalized learning in the fall of 2015 and the fall of 2016. Using qualitative analyses, we will examine interviewer probes and think-aloud responses to understand if younger children can self-report on their cognitive processes and if particular probes are more or less effective at eliciting useful feedback.	

Presentation Title	**Multi-use Field Testing: Examples from the 2017 Census of Agriculture Dry Run**
Presenter Name	**Jaki McCarthy**
Presenter Email	**jaki_mccarthy@nass.usda.gov**
Many large data collections will run a field test prior to operational field testing. This is typically used as a smaller scale dry run of the upcoming survey. The 2017 Census of Agriculture is the USDA National Agricultural Statistics Service's largest data collection, and is conducted once every 5 years. Prior to the 2017 COA, a field test is conducted as part of the multi method questionnaire evaluation. This large field test is used as a dry run of the questionnaire and data collection processes, but can also be used for other purposes. The 2015 field test was a crucial dry run, but was also used to produce records for the initial donor pool of records for imputation, compare the impact of alternative versions of the questionnaire on data quality and response rates, evaluate NASS's ability to accurately identify a pool of records for a reduced short form, test a new online system for web reporting, and test the ability of a pre-survey contact to help identify census nonrespondents. A field test is a rare opportunity that should be leveraged in multiple ways to make improvements to the subsequent data collections. Results from the 2015 COA field test and how they were used to benefit the COA will be discussed.	

SESSION 4B: The Use of Internet and Non-Probability Panels

Presentation Title	**Using Online Panel Surveys to Estimate Population-level Health Statistics**
Presenter Name	**Rosalynn Xin Yang**
Presenter Email	**RosalynnYang@westat.com**
Despite the increasing use of non-probability samples within survey research and previous studies comparing results between probability and non-probability samples, it is not yet clear under which circumstances appropriately weighted non-probability samples can yield valid inferences of population-level statistics. Last year we presented preliminary results from a study exploring the differences in health-related outcomes between a national probability mail survey and 2 non-probability web surveys. Several key estimates, (such as the proportion of respondents who have never smoked) were similar across methodologies, but others (especially those associated with use of the internet) had large discrepancies that were not corrected by propensity weighting. The current study extends this analysis, using additional data, into a	

more detailed and systematic review of the differences in population-level estimates of health outcomes between probability and non-probability sampling designs. We will compare results from the Heath Information National Trends Study (HINTS), a national probability-based mail survey, to online panel surveys from four different vendors, using a range of sampling methods including probability, quota-based, and convenience. We compare estimates from different sampling methodologies to eight national benchmark estimates, including estimates of health-related outcomes from the National Health Interview Survey (NHIS); and demographic estimates from the American Community Survey (ACS); and Current Population Survey (CPS). These analyses will contribute to clarifying conditions under which online panel data can be used for valid inferences concerning health information knowledge and health behaviors.

Presentation Title	**Who Won't Respond to Your Online Probability Panel Survey?**
Presenter Name	**Alexandra Brown-Breslin**
Presenter Email	**alexandra.m.brown@frb.gov**

Understanding how online probability panels affect data quality is an area of growing research interest. To contribute to this research, our work uses the Federal Reserve Board's "Consumers' Use of Mobile Financial Services" (Mobile, https://www.federalreserve.gov/econresdata/consumers-and-mobile-financial-services-report-201603.pdf) survey to explore the connection between online probability panels, unit nonresponse and attrition. The Mobile survey was conducted for five years, with each subsequent year after the first containing both a new sample and a resampled group that can be analyzed as a panel dataset. Because this survey also contains information on those individuals who were invited to take the survey but did not start and those that started the survey but did not finish, we can explore both unit nonresponse and attrition. After a general exploration into who does not opt in when invited to join an online probability panel and a description of the respondents and non-respondents in the new and panel samples in our data, we employ a latent class analysis model to better understand the distinct groups of individuals that exist within our survey non-respondents. Our research will follow the work of Lugtig, Das, and Scherpenzeel (2014) . Understanding these distinct groups will contribute to the body of research surrounding nonresponse and attrition in online probability panel surveys. It will also help us understand the underlying biases that may exist and should be noted while preparing our reports.

Presentation Title	**A Deeper Dive on the Digital Divide**
Presenter Name	**David Dutwin**
Presenter Email	**ddutwin@ssrs.com**

According to many public data, the percentage of U.S. households with internet access lies between 86 and 88 percent. If those who have internet access differ from those who don't, then surveys relying only on respondents with internet access may suffer from coverage biases. Such biases might be attenuated if an explicit correction for internet access could be made. In this paper we explore differences between respondents with and without internet access and then quantify potential biases related to internet coverage using data obtained from a dual frame RDD probability sample. We first analyze and document 16 major publically available surveys including the ANES, GSS, and many others on internet access against demographic, technographic, behavioral and political and privacy variables to note which, controlling for demographics, remain as significant drivers of Internet use. We find 542 variables to substantially predict internet use; code them into meaningful categories to better understand the nature of the digital divide; and utilize a random-forest approach to whittle these down to the essential 25 substantial variables. We then field a dual-frame telephone survey with those variables and construct an internet propensity model via both random forest and logistic regression, based on a final core set of 12 predictors. We apply the propensity adjustments to respondents with internet access and compare resulting estimates to the full dual frame sample on a collection of outcomes to explore the adjustment model's utility. We find, in short, that there is a wide chasm between internet and non-internet households, and as such, significant coverage biases. Our model is robust, however, accounting for well over 50 percent of the variance in internet use, and reduces bias significantly in all tests.

Presentation Title	**Methodological Considerations in the Use of Web Probing for Questionnaire Evaluation**
Presenter Name	**Stephanie Fowler**
Presenter Email	**stephanie.fowler@nih.gov**

Over the past several decades, methods for questionnaire development, evaluation, and testing have largely involved the conduct of cognitive interviewing. However, because of the time and resources required to conduct traditional lab-based cognitive interviews, several major shifts in this approach have led to internet-based systems that facilitate access to a wide range and large number of respondents, quickly and cheaply. These challenges and opportunities have led survey researchers to consider fundamentally new approaches to questionnaire pretesting, and in particular, the development of web probing as a component in our toolbox. Therefore, the present investigation examines 3 major aims related to web

probing. First, we examine the utility of web probes, relative to standard cognitive interviews, using the National Health Interview Survey and the Health Information National Trends Survey by examining differences in mode administration on gross productivity (i.e., length of text string entered in response to a web probe), usefulness of responses (interrater coding), and comparison of overall themes (thematic coding). Second, results may vary depending on whether probes are concurrent or retrospective (Willis, 2005). Therefore, using the same questionnaires and metrics described above, we will report the results of 2 web-based experiments that assigned respondents to either embedded probing (i.e., probes immediately following the targeted question) versus retrospective probing (i.e., debriefing). Finally, given the number of web-based platforms available to researchers, we examine the inclusion of web probes within qualitatively different web-based platforms comparing Mechanical Turk to other internet non-probability panels that incorporate varying approaches to recruitment and respondent contact. Finally, we synthesize the results of these studies to produce a set of integrated conclusions concerning the practice of web probing.

Presentation Title	**Implementation of a Youth Health Survey via the Internet**
Presenter Name	**Richard Lee Harding**
Presenter Email	**richard.l.harding@icfi.com**
The Florida Youth Survey project is a composite of three school-based surveys administered using paper-based instruments annually since 2002. ICF recently conducted a web-based pilot of the 2017 Florida Youth Substance Abuse Survey with high and middle school students in 10 counties. The goal of the study was to determine the impact of the administering the survey via a web-based system and to examine the impact of a series of methodological tests, including the use of skip logic and changes to the standard response options. Our findings on student's responses and teacher impressions as well as challenges and recommendations for internet administration are presented.	

ABSTRACTS – POSTERS

Poster Title	**Measuring Changing Household Complexity and Diversity by Race and Ethnicity from 2000-2010: A New Approach**
Presenter Name	**Laurie Schwede**
Presenter Email	**lschwede1@yahoo.com**

Research on household/family complexity and race/ethnic diversity is critical for policy makers, but the scope of such studies is limited by: lack of a standardized definition of household/family complexity; dataset sizes that cannot produce reliable results for the smallest race/ethnic populations; and emphasis on family over household units (with nonrelatives). We present results on changing household/family complexity without these limitations by proposing a new complex household definition and associated typology and by using the largest U.S. population datasets: the 2000 and 2010 censuses. Complex households increased from 21% to 23% from 2000-2010. We compare and contrast the changing complexity of households/families on key variables for the overall population and for eight race/ethnic groups: Hispanics and non-Hispanic: American Indians/Alaska Natives; Asians; Native Hawaiians/Other Pacific Islanders; Whites; Blacks; 2+ Races; and "Other Race." We compare trends, highlight differences, explore links to the persistent miscount of young children ages 0-4, and identify takeaways for demography and methodology.

Poster Title	**Complex Households and the Undercount of Young Children in the 2010 Census**
Presenter Name	**Eric Jensen**
Presenter Email	**eric.b.jensen@census.gov**

The estimated undercount for young children aged 0-4 in the 2010 Census was 4.6 percent. Prior research shows that household structure is related to census undercount. In this poster, we use an alternate approach to analyzing household structure and apply that approach to the undercount of young children. First, we classify households in the 2010 Census based on the different type of relationships within the household, using a new complex household typology that includes a wider variety of household types. Next, we provide a descriptive profile of children in the 2010 Census by the complex household typology. Finally, we analyze data on households that responded positively to one of the undercoverage probes about children across this typology. This shows us which types of households expressed uncertainty about whether to include a child on their census form.

Presentation Title	**Changing Immigration Laws and Their Impact on Attitudes within Qatar's Sponsorship System**
Presenter Name	**John Lee P Holmes**
Presenter Email	**johnleeholmes@gmail.com**

This paper analyzes the variation in ways that different communities in Qatar perceived the likelihood and effects of changes to the state's immigration laws. Criticism of GCC countries' labor laws from international regimes has become focused on Qatar in the context of its hosting of the upcoming 2022 FIFA World Cup. This has led to external pressure to change or eliminate the "Sponsorship" legal framework that regulates all non-citizens working in Qatar. Qatar University's Social and Economic Survey Research Institute (SESRI) has been measuring changes in expectations regarding workers' rights and the sponsorship system via phone and in person interviews since 2010. After an announcement last year that changes to these laws were forthcoming and with the promulgation of the labour laws on December 2016, the survey followed up with a battery of questions measuring understanding of and beliefs about whether changes will be meaningful once implemented. We measure the degree of support by nationals and other groups for the new changes, including how these changes will affect their businesses, social fabric and expectations for an improved international reputation. In addition to the social capital projects, the quarterly surveys also encompass a series of questions about consumer expectations and multiple quality of life measures. The surveys are conducted in seven languages covering the country's three major socio-economic groups: Qataris, high income expatriate ("white collar") and low income migrant ("blue collar") workers. As a consequence we are able to use multivariate analysis to understand underlying relationships between attitudes and support for or opposition to the system and its modification. This study illustrates the power of survey research to help better understand complex social processes generally, and labor migration in particular, issues of increasing importance in the current world economy.

Presentation Title	**Comparison of Model-Based to Design-Based Ratio Estimators**
Presenter Name	**James R. Knaub**
Presenter Email	**jamesRknaub@gmail.com**

The concepts of design-based and model-based ratio estimation, especially for the classical ratio estimator (CRE), are reviewed and compared. Ratio estimation is useful for energy, and natural resource official statistics - agriculture, soil science, and forestry - as well as in social science. Organization/business statistics uses may involve simple econometrics, as well as survey statistics. There are numerous applications. Notably, ratio estimation is very often useful for highly skewed establishment survey populations where, per K.R.W. Brewer, there should be at least as much implicit heteroscedasticity as for that of the CRE. Meaningfulness might be enhanced by understanding this comparison. The model-based and design-based interpretations of the CRE, their corresponding concepts of variance and bias, with relation to sampling and estimation, are reviewed, and familiar extensions of these estimators are also considered. Simple random sampling, cutoff, stratified, cluster, and unequal probability of selection methodologies, and some history are of interest. Even if a regression model is not explicitly considered, this review considers the role it plays, regardless.

Presentation Title	**National Health and Nutrition Examination Survey (NHANES) 24-hour Dietary Recall Sample Weights: Day of the Week Adjustment**
Presenter Name	**Te-Ching Chen**
Presenter Email	**tchen3@cdc.gov**

The What We Eat in America is conducted as a partnership between the U.S. Department of Agriculture and the U.S. Department of Health and Human Services. Under the partnership, the National Center for Health Statistics is responsible for the survey sample design and all aspects of data collection. All National Health and Nutrition Examination survey (NHANES) participants are eligible for up to two 24-hour dietary recall interviews, one in the Mobile Examination Center (MEC) and another by telephone interview 3 to 10 days later. The NHANES participants are selected using a national probability design. To increase the number of participants for specific demographic groups, a multi-stage, unequal probability of selection design is implemented. Sample weights are constructed that encompass the unequal probabilities of selection, as well as adjustments for non-participation. Because food intake can vary by weekdays and weekends, special weights accounting for the day of the week of the food intake and component nonresponse for the 24-hour dietary recall are created for each 2-year data cycle. Beginning with NHANES 2011-2012, the NHANES sample design has included an oversample of Asian Americans. In part as more sampling domains, very small race-age-day of the week sample weight adjustment cells were found when creating dietary weights for NHANES 2011-2012 and 2013-2014. The purpose of this project is to compare alternative dietary weights for the 1st dietary recall based on different groupings of day of the week. Differences in dietary intakes between weekend and weekday interviews are assessed. While initial results confirmed known patterns of weekdays and weekend intakes, estimated differences between weekday and weekend intake using alternative weighting approaches are small. For example the difference average total energy intakes by weekends/weekdays among adolescent females ranged from 190 to 221 kcal across weighting methods, with p-values from t-tests ranging from 0.02 to 0.04.

Presentation Title	**An Experiment with Varying Survey Cover Letter Paper Color and its Effect on Survey Response Rates by Mode**
Presenter Name	**Amy L. Djangali**
Presenter Email	**adjangali@impaqint.com**

One tool employed in the effort to increase response rates to mailed surveys is colored paper. The color may make the materials stand out among other mail, leading to more recipients noticing the mailing and subsequently completing the questionnaire. Results are mixed, however, with some studies showing little or no effect of questionnaire paper color while another meta-analysis found a significant effect. IMPAQ conducted a randomized experiment to investigate the effects of paper color for a statewide employer survey. Unlike previous research, we varied the color of the cover letter, not the color of the questionnaire. The experiment was conducted as part of a multimode survey of employers for an evaluation of a state workforce program. Employers (N = 5,893) were randomly divided into four groups. There were two mailings sent out, an initial mailing with a copy of the questionnaire and a Web survey link, and a second mailing about a month later with a reminder and another copy of the questionnaire and the Web survey link. Half of the sample received white cover letters for both mailings (white/white), and the other half were evenly divided between three groups – blue/blue, white/blue, or blue/white. Our analyses will examine the effect of the cover letter color on overall response rate, wave response, response mode, and length of time from mailing to response.

Presentation Title	**Reducing the Number of Pages of a Mail Questionnaire in a Mail/CATI Mixed Mode Survey: How Does It Affect the Response Rates and Costs of Each Mode?**
Presenter Name	**Michelle Cantave**
Presenter Email	**mcantave@impaqint.com**

Previous research examining the effect of reducing the perceived length of paper questionnaire by reducing the number of pages through formatting, without altering the questionnaire content, is mixed. In addition, to the best of our knowledge, no prior work has examined whether reducing the perceived length of the paper questionnaire in a mixed mode study generates higher response rates and reduces costs for both modes, not just the mail component. The Fee-for-Service Consumer Assessment of Healthcare Providers and Systems Survey (FFS CAHPS) is a sequential, mail, then centralized telephone follow up mixed mode design. From 2011 to 2015, the FFS CAHPS paper questionnaire was a 16 page booklet with additional instructional pages. In 2016, we reduced the booklet to 12 pages and simplified the instructional language (the number of questions remained the same). Confirming the previous literature, we observed an increase in the number of mail surveys received. Though the increase cannot be fully attributed to the reduced pages, no other large-scale changes were made to the survey that would account for such a sharp increase in the mail response and for a halt to the previously downward trend in the response rate. We will present our findings on the increased response rates and decreased costs in both modes, as well as an analysis of the response rate increases by demographic characteristics.