

AAPOR

AMERICAN ASSOCIATION FOR PUBLIC OPINION RESEARCH

66th Annual Conference



Public Perception & Societal Conflict

2011 Conference Abstracts

May 12 – 15, 2011
Arizona Grand Resort
Phoenix, AZ

Thursday, May 12, 2011

1:30 p.m. - 3:00 p.m.

Concurrent Session A

Administrative Records and the 2020 Census

Possible 2020 Census Designs and the Use of Administrative Records: What is the Impact on Cost and Quality?

Frank Vitrano (Frank.A.Vitrano@census.gov)

While administrative records have been supporting decennial censuses for decades, we are for the first time considering their use in our most large and costly operation—non-response follow-up. We believe that using administrative records to supplement non-response follow-up is the single biggest potential cost saver for the 2020 Census. Early 2020 Census planning has focused on operational design work in three areas, which are the major cost and quality components for the census: Establishing Where to Count; Enumeration Activities; and, IT and Field Operational Infrastructures. Establishing Where to Count options all include continual updating of the address list during the decade but vary from a Full Address Canvassing to Targeting Address Canvassing to No Address Canvassing just prior to enumeration in 2020. Enumeration Activities vary from Traditional +Internet approach to a Multi-mode Contact Option to a Visit then Records Options to Records then Visit to Near Paperless to a Full Administrative Records option. The idea will be to model, analyze, and test different approaches to conduct or support these major census functions. Six design alternatives have emerged from combinations of the different operational options. This panel discussion will describe and compare the alternatives but will focus on the alternatives that include administrative records. It will also include a discussion about potential coverage and technical issues associated with using administrative records in lieu of collecting data directly from the respondent.

Expanded Use of Administrative Records in the 2020 Decennial Census: Can the Non-Technical Issues Be Resolved Early in the Decade?

Nancy Potok (npotok@gwu.edu)

Even as the Census Bureau researches the technical feasibility of increased use of administrative records in the 2020 decennial census, it must also resolve a number of non-technical issues. Building stakeholder support for a major design change will be a key component of successful implementation. A few examples of areas in which census stakeholders are likely to take an interest include any new privacy implications that may be raised by use of administrative records, whether administrative records are timely enough to represent place of residence as of Census Day (April 1), assuring the availability of records from other agencies, and measuring the quality of source records. Examination of these and several other areas will be better informed after the Census Bureau conducts research during 2012-2015. However, in order to reduce risk and avoid major design changes late in the decade, the Census Bureau needs to address many of the non-technical issues prior to mid-decade. How these issues are resolved could have a profound impact on many potential uses of administrative records throughout the Federal statistical system. This panel discussion will expand upon and illustrate these concerns, as well as some of the other issues surrounding expanded use of administrative records.

Moving toward Integrated Data Sets—Can we Overcome the Challenges?

Sally Obenski (Sally.M.Obenski@census.gov)

There are several challenges that will need to be addressed as the Census Bureau moves toward greater use of integrated data sets. These challenges can be grouped into four basic types. The first has to do with the way the Census Bureau and other federal agencies are organized: programs are funded in order to produce products, and this means that there is little incentive for them to provide the resources needed for cross-cutting initiatives like integrated data sets. The second set of challenges to be discussed is operational in nature. Working with administrative records in the context of the decennial census entails the processing of enormous files on tight schedules and complex handoffs of analyses and products in a secure environment. Technical hurdles like developing a measure of error for administrative record data will pose operational difficulties as well. The third set of challenges lies outside the Census Bureau: complex acquisitions of federal program records are required for this initiative, and past experience has shown that there can sometimes be substantial definitional differences between agencies. For example, work underway on Child Care Subsidy data from just three states demonstrates that criteria for eligibility can be fluid and can vary substantially by locale. The last type of challenge to be discussed concerns communicating the value of integrated data sets. First and foremost, we must hammer home that keeping an individual's data confidential is sacrosanct---which brings us to the old adage that "The Census Bureau does not care about YOU!" It's also important to communicate the importance of large, high quality samples for data integration. In particular the use of the fledgling, newly funded American Community Survey for federal program eligibility will need to be discussed.

Replacing Census Reports with Administrative Data: Is the Public Ready and Willing?

Eleanor Singer, Nancy Bates, and John Van Hoewyk. Eleanor Singer, ISR, University of Michigan (esinger@isr.umich.edu)

Among the guiding principles for the 2020 census are reducing costs and field time while maintaining quality. One means to achieving this goal is using administrative records in place of personal-visit nonresponse follow-ups. To what degree might potential privacy and confidentiality concerns be a barrier to this change in the 2020 census design?

Prior research on this topic between 1996-2000, suggested that a substantial part of the public had serious reservations about data sharing among federal agencies. A Joint Program in Survey Methodology (JPSM) Practicum survey conducted in 1995 found that a majority of respondents approved the use of administrative records to supplement or improve the decennial census, and a slim majority approved using administrative records as a direct substitute for data collection (Singer and Presser 1996). A Westat Inc. replication of the 1995 JPSM survey found little change in attitudes toward data sharing between the two surveys (Singer, Van Hoewyk and Presser 1997). However, beliefs about privacy and trust in government had declined since 1995. And by 2000 only a minority approved the use of administrative records by the Census Bureau.

The current paper revisits these results by analyzing data collected as part of the 2010 JPSM Practicum. We explore three areas. First, we compare the 2010 survey responses to those in previous surveys. Second, we examine current predictors of favorability toward administrative records use, including trust in government, trust in the Census Bureau, privacy concerns, and fear of data misuse, together with political party, race, age, education, and other demographics, and ask whether these predictors vary depending on the agency whose records are to be used and by the reason given for their use. Finally, we ask how likely attitudes are to predict behavior.

Historical Perspectives on Survey Research and Polling

The Dissemination of an Innovation: Survey Research 1936-1959.

Tom W. Smith, NORC (smitht@norc.uchicago.edu)

Modern public opinion research started in the United States in the mid-1930s primarily with the surveys of Crossley, Gallup, and Roper. Public opinion research soon began to disseminate to other countries. This paper looks at the initial development of organizations, journals, and conferences to form the new field and the special role that World War II and the United Nations had in internationalizing the innovation of public opinion research.

Trends in Media Polling: Public Opinion Polls, Pre-Election Polls, and Trial-Heat Questions since 1970.

Chase Harrison, Harvard Business School (charrison@hbs.edu); Marc Maynard, Roper Center for Public Opinion Research; University of Connecticut (marc.maynard@uconn.edu); Richard Clark, Carl Vinson Institute of Government; University of Georgia (clark@cviog.uga.edu)

In recent years, researchers have focused significant attention on the impact that media-reported pre-election polling has had on elections, campaigns, and voters. Although public opinion polls are relevant tools for measuring opinions, attitudes, and beliefs in general, they often become the most salient and controversial when used as tools for forecasting elections. Many argue that changes in both the technologies available for survey research, combined with changes in the nature of modern media, have caused a particularly large focus on pre-election trial-heat measures compared to more general issue polling. We provide original quantitative data and analysis which can illustrate and illuminate some of these trends.

In this paper, we describe and analyze trends in media polling in the United States between 1970 and the present. We document the increase in media polling during this period, and study the influence of several factors on the increase in media polling. In particular, we attempt to disentangle two separate but related factors concerning media polling: the growth of public opinion polling in general, and the increase in pre-election polling. We separately analyze trends in the number of surveys conducted, number of survey questions reported, and on the nature of organizations sponsoring and conducting polls. We focus specifically on the use of pre-election “trial-heat” questions, as well as questions about politics more generally and questions about issues. We provide descriptive data on the growth in the number of public opinion questions reported, and we test several multivariate models.

Data are derived from a comprehensive database of national public opinion polls developed and maintained by the Roper Center for Public Opinion Research. Our analysis focuses on national public opinion polls conducted by or for media outlets from 1970 through 2010.

The New Deal Realignment in Real Time

Helmut Norpoth, Stony Brook University (hnorpoth@notes.cc.sunysb.edu); Clara Suong, Stony Brook University (clara.h.suong@gmail.com); Andrew Sidman, John Jay College (asidman@jjay.cuny.edu)

Between 1937 and 1952, the Gallup Organization probed party identification in over 150 polls. We examine this largely untapped pool of survey data to conduct a “real-time” analysis of the electoral realignment that is largely attributed to the Great Depression and commonly called the “New Deal

realignment.” With rare exceptions, students of this realignment have relied on the National Election Studies, which did not begin asking about party identification until 1952. Using such latter-day data required reconstructions of partisanship that are troublesome, given faulty memory and life-cycle changes. We find that it was not so much the Great Depression or the New Deal as the wartime experience and the postwar prosperity that gave the Democratic Party its overwhelming hold on the American electorate for the next three decades. The 1948 election plays the critical role, not the 1932 or the 1936 election. Whatever gains the Democratic Party had reaped in party identification by 1936 were short-lived. The generation that contributed the most to the Democratic ascendancy is the one that came of age in the 1940’s, not the one that did in the 1930’s.

Public Agenda in México During 2007-2011.

Paul Francisco Valdes Cervantes, Parámetro Investigación (paul.valdes@gmail.com); Nelson Arteaga, UAEM (arbnelson@yahoo.com)

We present results on the major social concerns perceived by public opinion in Mexico. We present the evolution of the issue of public insecurity and economic situation as well as a partisan identification and approval of authorities using data from a series of seven national telephone surveys done in that period.

Public insecurity and economic situation have been the two main concerns of Mexicans in recent years, particularly we present data from ten measurements taken between October 2007 and February 2011. The main problems perceived by society are reflected on the public agenda, which does not necessarily coincide with the agenda of media and the political agenda. For public agenda is meant the ranking of main problems perceived by a society. In this essay we obtained measuring the public agenda through the open question: What is the main problem that the country faces right now?

Public safety has been present as the main concern of Mexicans in five of the seven measurements, and is generally perceived negatively on performance of the authorities by a majority of Mexicans that have telephone at home. The perception on the economic situation is another main concern of Mexicans. It has consistently been perceived as the second issue of Mexicans except two of the seven measurements (July 2009 and February 2011).

What is the social impact of the deterioration in the perception of public security and economic situation of the country? The survey results tell us that the growing perception of violence and insecurity; on the other hand the perception of the deterioration in the economic situation has contributed to decrease levels of trust towards politics and institutions, and in particular has produced significant costs in the approval of the performance of the government party.

Methodological Briefs: Weighting & Design Issues in RDD/Cell Phone Surveys.

Do Cell Phone Mostly and Equal Cell/Landline Telephone Users Differ By Frame and Should We Care?

Kathleen Thiede Call, University of Minnesota (callx001@umn.edu); Donna D McAlpine, University of Minnesota (mcalp004@umn.edu); Timothy Beebe, Mayo Clinic (Beebe.Timothy@mayo.edu); Michel Boudreaux, University of Minnesota (boudr019@umn.edu); Brett Fried, Minnesota Department of Health (Brett.Fried@state.mn.us); David Dutwin, SSRS (ddutwin@ssrs.com)

To address concerns about representativeness, many telephone surveys are supplementing landline (LL) with cell phone (CP) sample frames (dual-frame). Cell samples cost more, thus, it is important to assess

when they should be used, and what specific type of cell users should be included: (1) screen for cell phone only (CO) households, (2) include dual users who rely mostly or equally on their cell phones, or (3) interview the entire CP frame. Some suggest that screening for CO is simpler and cheaper, however, coverage bias will increase if screened out cases are systematically different than cell users reached by landline. Our primary goal is to compare cell mostly (CM) populations and cell equal (EQUAL) populations (equally likely to receive calls on LL and CP) from the LL versus the CP frame. Comparisons will be made across demographic characteristics and a range of health outcomes. Our second goal is to estimate the level of coverage bias introduced by screening for CO households.

We draw on two dual-frame sources of data. The 2009 Minnesota Health Access Survey, an ongoing survey of health insurance and access that accepted all CP completes (9,811 LL; 2,220 CP). The 2010 Minnesota Survey on Adult Substance Use focuses on prevalence of drug and alcohol use and patterns of treatment that accepted only CO and CM cases. Preliminary analysis is based on the first 10,036 cases (8,054 LL; 1,982 CP); the total (December, 2010) will be approximately 16,000 cases.

These data demonstrate that cell-phone users that are sometimes screened out of cell samples in dual frame protocols (CM and EQUAL) are systematically different in terms of demographics, health insurance, and risky health behaviors than their counterparts captured in the landline frame. We provide evidence that these differences will not be completely mitigated by common-practice weighting adjustments.

Comparison of RDD and Address Based Cell Phone Only Sample Frames.

Courtney N Mooney, Arbitron (courtney.mooney@arbitron.com); Ryan McKinney, Arbitron (ryan.mckinney@arbitron.com)

The growing incidence of cell phone only (CPO) households has caused the potential for coverage bias issues for survey researchers that have historically relied on RDD landline sample frames. In order to address this issue, Arbitron has fielded two different CPO sample frames: RDD telephone based and address based frames, to supplement the traditional RDD landline sample frame.

Both types of CPO sample frames present unique operational challenges to survey researchers. Additionally, the coverage of the two frames can differ due to the mode and contact methods that are employed. This presentation will evaluate the differences and similarities between the two frames by exploring respondent characteristics in four major US markets. In particular, it will examine which frame is more in line with CPO universe estimates, race and age comparisons between the two frames, and the success of both frames when seeking respondents from high density ethnic areas.

The Impact of Cell Phones in Landline Households for an RDD Telephone Survey

Andrea Werner-Leonard, Population Research Laboratory (andrea.werner-leonard@ualberta.ca)

Cell phone only households are increasing across the world. The rate of cell phone only households in the United States is more than double that of Canada. As of 2008, the province of Alberta had the highest rate of cell phone only households (11.5%) compared to the national average of 8.0%. This paper presents the results of a Random Digit Dialing (RDD) telephone survey of 402 adults conducted by the Population Research Laboratory (PRL) at the University of Alberta in the spring of 2010 in Edmonton, Alberta. The study asked landline household participants how many and what type of cell phones or other telecommunication devices were used in the household; how cell phones were used at home; intention to cancel landlines in the next 12 months; use and type of telephone screening practices; the importance of participating in public opinion surveys; and survey completion preferences. The analysis explored whether households with cell phones differed from those without cell phones on various demographic

characteristics such as income, age, sex, marital status, home ownership and education to determine if there were coverage issues in the RDD sample. The results of this study reveal some significant differences among certain demographic variables that are inconsistent with the literature (for example, household size and age) which raises questions for future studies. Finally, the paper attempts to explain why the rate of cell phone-only households is lower in Canada than the United States and many other countries around the world.

Who's There? Comparing Respondents from a Telephone Survey to a Mail Survey

Andrew Zukerberg, NCES (andrew.zukerberg@ed.gov); Stacey Bielick, American Institutes for Research (sbielick@air.org)

Decreasing telephone response rates have caused survey researchers to explore alternate modes of data collection. The increasing accuracy and availability of address sampling frames have led researchers in the public and private sector to reconsider mail self-administered surveys as a viable alternative to telephone RDD studies. Mode effects in surveys have been well documented. One aspect of the effect that is worthy of exploration in the transition of a survey from telephone to mail is whether or not the change in mode draws in a different respondent from the household. The National Household Education Survey (NHES) has been conducted by telephone approximately every two years since 1991. As a result of falling response rates and concerns about coverage in the list assisted RDD sample, the survey began a redesign in 2009. The first step in this redesign was a feasibility test of a two phase mail survey. The NHES consists of a series of rotating modules, most of which are focused on a reference child. Past research has shown that the household informant plays a role in the quality of data collected. As a result of this, telephone interviewers were instructed to ask that the adult in the household that is most knowledgeable about the selected child serve as the respondent. The mail self-administered questionnaire requested that someone knowledgeable about the child respond, however, the request is likely to be less salient to the respondent in a self-administered survey compared to an interviewer-administered survey. This paper will examine the characteristics of reporters in the phone and mail surveys. Changes in the household reporter could have profound effects on measurement. Understanding differences in who is likely to respond by phone compared to mail can also lead to the development of better contact approaches for mail surveys.

Living in a Digital World: A Comparison of Consumer Ownership, Attitudes and Behaviors of Technology by Sample Frame

Rhonda Daniel, Consumer Electronics Association (rdaniel@ce.org)

The National Center for Health Statistics (NCHS) reported that the number of cell phone only households in the U.S. was 24.5 percent at the end of 2009. It is expected that this percentage will continue to increase as cell phone penetration increases and households look for ways to cut costs in this challenging economic environment. The rising incidence of cell phone only households increases the impact of noncoverage bias on survey results using the traditional RDD landline frame and research has shown that younger adults are more likely to cut the landline cord. As young adults (18-34 years old) are more likely to be early adopters of technology, one would expect there to be differences on ownership and usage of various technologies, especially new, emerging technologies.

In order to determine where the most bias exists using an RDD landline sample frame, the Consumer Electronics Association (CEA) conducted three dual sample frame surveys in 2010 to understand how the inclusion of an RDD cell phone sample impacts survey results for consumer electronics ownership and usage as well as attitudes on government involvement on technology design. This paper will present the findings of these three surveys as well as recommendations for future research.

Exploring Health-Related Experiences and Access to Care: Differences Between Online and Telephone Survey Administration.

Michelle McEvoy Doty, The Commonwealth Fund (mmd@cmwf.org); Jordon Peugh, Knowledge Networks (jpeugh@knowledgegenetworks.com); Jeffery Shand-Lubbers, Knowledge Networks (jlubbers@knowledgegenetworks.com)

In an era of health reform where decisions concerning health care delivery are changing, researchers and policy makers need to rapidly gauge public opinions and needs. Online survey administration using a probability-based panel of respondents offers the possibility of getting results quickly without sacrificing integrity of data. This paper compares results from nationally representative surveys of adults using two different data collection modes: online and telephone.

To explore the comparability of data collected online versus by telephone, we fielded identical questions on access to care, experiences with health care delivery, satisfaction with quality of care, and demographics to a national landline RDD sample of 2,501 U.S. adults and a national probability-based web panel sample of 1,039 U.S. adults, using Knowledge Network's KnowledgePanel®. The purpose of this paper is to highlight where results from the two modes align well; to assess where responses do not converge; and explore potential reasons for discrepancies.

Using these two methodologies, we found considerable similarities on responses to behavior/experience related variables, such as cost-related access problems and ease in getting care after-hours. There were 15 to 20 percentage point differences, however, on attitude and opinion questions. We compare how responses to opinion questions vary by income, race-ethnicity, insurance status, and other key demographics.

In addition to an overall comparison of these modes we explore the impact of cell phone only households (which are included in the online panel and not in the RDD sample) and the probability that certain sub-groups such as those with a chronic condition, low-income respondents, and racial/ethnic minorities may be excluded with RDD samples.

The paper will discuss implications for using the different methodologies for conducting health policy research and consider how these methodologies differ on other items such as design effects, response rates, time in field, and cost ratios.

Response Rate Calculation In Mobile Surveys: Higher Precision In Cases Of Unknown Eligibility

Bella Struminskaya, GESIS - Leibniz Institute for the Social Sciences

(bella.struminskaya@gesis.org); Ines Schaurer, GESIS - Leibniz Institute for the Social Sciences

(ines.schaurer@gesis.org); Wolfgang Bandilla, GESIS - Leibniz Institute for the Social Sciences

(wolfgang.bandilla@gesis.org); Siegfried Gabler, GESIS - Leibniz Institute for the Social Sciences

(siegfried.gabler@gesis.org); Sabine Haeder, GESIS - Leibniz Institute for the Social Sciences

(sabine.haeder@gesis.org); Lars Kaczmirek, GESIS - Leibniz Institute for the Social Sciences

(lars.kaczmirek@gesis.org)

As the growing mobile-only population is a source of coverage error in telephone surveys, research organizations increasingly use dual sample frames, which include landline and mobile phone numbers. The inclusion of mobile numbers in the sample brings certain challenges with respect to the calculation of response rates. AAPOR's general recommendation is to adjust standard disposition codes from RDD surveys. However, there are certain problems, such as ambiguous operator messages that vary

unsystematically by operator (AAPOR Cell Phone Task Force 2010). Such outcomes have to be treated as cases of unknown eligibility when calculating response rates. A possible partial solution to this problem is to determine whether a mobile number is working or not by using the existing technologies available in the GSM network. This possibility is of particular value for cases where non-working numbers cannot be eliminated by means of standard interviewer contact. Identifying non-working numbers from the pool of numbers with unknown eligibility would allow for the calculation of response rates more precisely, since the non-working numbers can be treated as ineligible.

In the present study, the outcomes of interviewer contact attempts of mobile numbers fielded as part of a RDD telephone survey (such as successful contacts, non-contacts and reasons for unsuccessful attempts (N=24691)) are compared with the results of the network query of mobile subscriber's status (active, absent, unknown, etc.).

The results show that about 8% of numbers with unknown eligibility are non-working numbers and could be treated as non-eligible. The analysis of unclear operator messages in detail reveals that about 14% of contacts prove to be invalid.

The study shows that a proportion of cases of unknown eligibility can be identified as non-eligible.

Therefore, response rates are underestimated and would be more accurate with the proposed method.

Methods to Improve Web Surveys

Analyzing Paradata as a Form of Post-hoc "Remote" Usability Testing: Initial Results From the 2010 Census Quality Survey (CQS)

Kathleen T. Ashenfelter, U.S. Census Bureau (kathleen.t.ashenfelter@census.gov); Matthew Jans, U.S. Census Bureau (matthew.e.jans@census.gov)

Collection of Census data by Web survey is anticipated to be an important component of Census 2020. To prepare for this, the 2010 Census Quality Survey (CQS) followed-up a sample of 2010 Census respondents with an internet survey (among other modes) The Web survey collected a variety of server-side paradata from respondents as they completed the online CQS Survey. These included measures of how long it took to complete the each question and the entire survey, which items were changed (including original and corrected values), whether the respondent could complete the survey in one sitting, and many other issues of interest to Web survey methodologists. We analyze these paradata, in combination with knowledge about the survey questions (e.g., which items are more complex or sensitive on face value), with an orientation toward interpreting the paradata as measures of respondent burden (e.g., length of the time to complete questions), cognitive difficulty (e.g., changes to answers) and other psychological states that can lead to measurement error, item nonresponse error, or unit nonresponse error (in the form of break-offs). The relationship of post-hoc Web survey paradata analysis to traditional lab-based usability testing at the Census Bureau will be discussed. Implications for mode-specific Total Survey Error components that can be linked to the interaction of respondents' psychological states and Web survey design will be emphasized.

Choice of Content Presentation Mode in Web-Based Survey Administration

Larry Osborn, Knowledge Networks (losborn@knowledge networks.com); Wendy Mansfield, Knowledge Networks (wmansfield@knowledge networks.com); Carl Ramirez, Government Accountability Organization (RamirezC@gao.gov); Jill Lacey, Government Accountability Organization (LaceyJ@gao.gov); Virginia Chanley, Government Accountability Organization (ChanleyV@gao.gov); Gregory Wilmoth, Government Accountability Organization (WilmothG@gao.gov); Ana Ivelisse Aviles, Government Accountability Organization (AvilesA@gao.gov)

Survey research often involves information provision to respondents in questionnaires. The information provision allows for measurement of research subjects' reactions, attitudes, and opinions regarding the supplied information. Web-based survey administration, since it supports multimedia, allows for multiple modes to provide such information, including video, audio, or on-screen text. An experiment testing the optimal mode(s) of information provision would be useful for researchers. Recently, KnowledgePanel® provided data for a nationally representative study on peoples' opinions and attitudes around global climate change and geoengineering technologies, conducted for researchers from the U.S. Government Accountability Office (GAO). The Internet-based survey was conducted with 1,006 U.S. residents aged 18 and older and included the presentation of content outlining two examples of geoengineering technologies. Respondents were initially offered the choice of reviewing this content in one of three modes—video/audio, audio-only, or on-screen text. Respondents were subsequently allowed to review the content in one or both of the remaining alternative modes, if desired. Of the total respondents, 175 chose to review the content in a second presentation mode and approximately 50 respondents chose to review in all three modes. Our paper will examine three key questions. First, are there presentation modes that appear to be more or less effective (as measured by item nonresponse, satisficing behavior, etc.) based on a statistical analysis of combined initial and subsequent content reviews by respondents? Second, are there respondent characteristics related to single-mode choice versus multiple mode choices? Third, are there respondent characteristics related to initial mode preference and/or switching to other modes? Based on these data, implications and recommendations for single and multiple information presentation modes will be discussed. We expect the study results to be useful for researchers who design online survey experiments and discrete choice, referendum, and willingness to pay studies.

Providing Clarifying Instructions in a Web Survey
(DC AAPOR Student Paper Award Winner)

Cleo D. Redline, University of Maryland (credline@comcast.net)

Although comprehension is well recognized as a critical component of the survey question response process, much about it remains unknown. Past research has shown that ambiguous or vague concepts can be clarified through the use of definitions, instructions or examples, but respondents do not necessarily attend to these clarifications. The aim of this research was to investigate where and how to present clarifying information so that respondents would recognize it as essential to their answering survey questions correctly. The research questions were: does placing the clarification before a question, in the same font as the question, or incorporating the clarifications into a series of questions improve survey estimates? Eight study questions, modeled after major national surveys, were administered as part of a probability based Web survey (n = 913 completed interviews). The results suggest that respondents attended to the clarifying instructions. There is some evidence to suggest that in line with expectations, respondents anticipate the end of a question and are more likely to ignore clarification placed after a question than before. Respondents answer more quickly when the questions come after the question, suggesting they spend less time processing the clarification in this position, and their answers appear to be little less accurate. The practice of attempting to highlight the question through the use of font is not supported. However, as predicted, there is evidence to suggest that incorporating the clarifications into the questions and asking a series of questions is the most effective approach. Respondents answers appear to be the most accurate when this approach is employed.

Which Web Survey Respondents Are Most Likely to Click for Clarification?

Tania Coiner, New School for Social Research (tania.coiner@earthlink.net); Michael F. Schober, New School for Social Research (schober@newschool.edu); Frederick G. Conrad, University of Michigan (fconrad@isr.umich.edu)

During survey interviews, respondents sometimes interpret question concepts differently than they are intended. For example, a respondent who is asked “How many people live in your home?” may include a child living away at college even though the survey designers’ concept excludes household members whose primary residence is elsewhere. Some survey respondents seem to be aware of such conceptual misalignments and seek clarification when it is available, but others do not, even when they need it, and instead answer confidently and (for survey purposes) incorrectly. What distinguishes respondents who seek clarification from those who do not? This study demonstrates that cognitive and personality factors seem to predict clarification-seeking in a web survey where respondents could click for definitions of key question concepts. In a laboratory experiment, 60 respondents answered behavioral questions on the basis of fictional scenarios, some of which were designed to create potential comprehension difficulty, and for which response accuracy was measurable. Respondents also completed the Symbol Search and Digit Symbol Coding subtests of the Wechsler Adult Intelligence Scale and the Dickman Impulsivity Index (DII, 1990). Results showed that respondents with faster processing scores requested clarification more often and produced more accurate answers; they also took longer to answer (despite their faster processing), presumably because they were more thoughtful. Respondents who scored high on dysfunctional impulsivity requested definitions less frequently, answered less accurately, and answered more quickly. Symbol Search and dysfunctional impulsivity were strong predictors of frequency of requesting clarification, while all three measures strongly predicted response accuracy. Because the DII and Symbol Search measures could potentially be adapted for on-line administration, one can imagine new web interviewing systems that improve response accuracy by diagnosing which respondents are unlikely to request clarification when they need it, and giving them clarification when it is needed but not requested.

Race-of-Virtual-Interviewer Effects

Frederick G. Conrad, University of Michigan (fconrad@isr.umich.edu); Michael F. Schober, New School for Social Research (schober@newschool.edu); Daniel Nielsen, University of Michigan (cyphe@isr.umich.edu)

In developing self-administered interviewing systems that go beyond text, survey designers are faced with choices about the virtual “interviewer’s” (VI’s) attributes. In ACASI and IVR, for example, designers must select interviewers’ voices. Any visual representation of an interviewer (e.g., a photograph, a video) requires choices about the interviewer’s appearance. In earlier work we have demonstrated that such choices can affect answers; for example, we have shown that respondents report different numbers of lifetime sexual partners depending on the extent of a virtual interviewer’s facial movement, much as occurs in comparing answers between human and ACASI administration. Here we investigate whether a VI’s appearance--in particular the attributes that might lead respondents to think of the VI as belonging to a particular racial group or a particular gender--affects respondents’ answers to race- and gender- related questions. In a web survey of 1735 respondents (half Black and half White, half female and half male), respondents answered (clicking or typing) questions about race and gender issues asked by one of 16 animated interviewers (Black, White, female, male). We observed race-of-VI effects on several race-related questions, but no gender-of-VI effects on gender-related questions. For example, more respondents reported strong (rather than not strong) opposition to preferences in hiring when the question was posed by a VI identifiable as White. When respondents were asked (post-interview) to choose a VI for a hypothetical future interview, black respondents were particularly likely to choose a black VI. The fact that non-human “interviewers” can elicit effects akin to those found with human interviewers raises questions about the mechanisms underlying race- and gender- of interviewer effects more generally.

Sharpening Your Instrument: Strategies to Improve Survey Questions

Hitting the Target: How Number of Items, Response Categories, and Sample Size Affect Aggregate Estimates.

Randall K. Thomas, ICF International (randall.k.thomas@gmail.com); John Bremer, Compete Inc. (teamwiess@me.com)

Increasing the number of items to measure a specific concept has been shown to improve the stability of measurement according to the Spearman-Brown formula. Thomas (1999) showed that the number of items and response categories can be jointly described as a response space and that increasing the gradations within the response space improves response stability, reaching an asymptote of reliability. In the field of sampling statistics, a major concern has been how increasing sample size improves precision of estimates of population parameters, following from the Law of Large Numbers (Bernoulli, 1713). While some research has investigated the effects of sample size on inter-sample reliability, little research has investigated the intersection of individual measurement stability and sample measurement stability. In this paper, we extend the response space concept to include items, responses, and individuals to understand how they can affect stability of aggregate estimates (e.g. means, proportions, and variance). We summarize three separate studies that randomly assigned number of response categories and had multiple items assessing the same concepts that allowed us to compare the additive and interactive effects of the number of items, number of response categories, and number of individuals in their effects on aggregate measurement stability and validity.

Matrix Questionnaire Design to Reduce Measurement Error

Andy Peytchev, RTI International (andrey@umich.edu); Emilia Peytcheva, RTI International (epeytcheva@rti.org)

The full set of questions may be essential to fulfilling the survey objectives, yet long survey instruments can be taxing to respondents. As the respondent continues to answer survey questions for an extended period of time, they may not provide the same level of consideration to all survey questions – providing answers with greater measurement error towards the end of the survey. There is little empirical evidence to support the tradeoff between survey length and measurement error, leaving survey practitioners to rely on expectations to support improvement of survey estimates using a shorter survey. Furthermore, empirical research is in dire need of methods to address the challenge of collecting the full set of variables while minimizing the survey length.

One potential solution is the use of split or matrix questionnaire design in which the survey is modularized and respondents are randomly assigned to receive a different subset of the survey modules. In this way, no questions are asked so late in the survey that the measurement error properties of the data are compromised. The full sets of variables are then multiply-imputed for all respondents. This method provides a tradeoff between measurement error in the long survey and increased variances of estimates in the matrix design. We pose two research questions:

1. Whether there is greater measurement error when questions are asked late in the survey, and
2. Whether a matrix questionnaire design can provide estimates with lower total error (MSE).

We use data from a web survey experiment in which a set of questions were asked either early or late in the instrument. We found strong evidence for greater measurement error when the questions were asked

late in the survey. This paper will present an evaluation of the use of a matrix questionnaire design instead of the full survey.

Text Box Formatting in a State Omnibus Survey: Content, Response and Perceptions

Mindy Anderson-Knott, University of Nebraska-Lincoln, SSP (mandersonknott2@unl.edu); Amanda Richardson, Bureau of Sociological Research, University of Nebraska-Lincoln (arichardson2@unl.edu); Ashley Frear Cooper, Bureau of Sociological Research, University of Nebraska-Lincoln (afrear2@unl.edu); Julia McQuillan, Bureau of Sociological Research, University of Nebraska-Lincoln (jmcquillan2@unl.edu); Stacia Jorgensen, Bureau of Sociological Research, University of Nebraska-Lincoln (shalada2@unl.edu)

Previous research with college students demonstrates that more respondents write responses in boxes for open-ended comments when the box contains no lines instead of lines. Substantive responses, however, do not differ by type of box (lines versus no lines) (Dillman & Christian 2004). Using a state-wide omnibus survey of over 2,000 Nebraskans, we expand on studies of college student populations by examining differences in lined versus no-lined formatting in a general population survey. Our evaluation of these effects is centered on the open-ended question concerning current public sentiment in the state ("What is the most important issue currently facing the state of Nebraska?"). In addition to formatting differences, we also provide analyses of demographic subgroups to assess if the effect is consistent across groups. We hypothesize that we will find no differences between key respondent characteristics such as age, education, gender, and rural/urban. If the box type differentially influences response by subgroups, then the messages reflected in the open-ended responses could be biased. Findings will help to guide quality data collection efforts by ensuring that survey formatting does not inadvertently influence the length of response or the substantive content of the response.

Optimizing the Design of a Question Intended to Measure Expected Starting Salary

Curtiss Cobb, Stanford University (clc2003@stanford.edu); Jon A Krosnick, Stanford University (krosnick@stanford.edu); Brent Bannon, Facebook (bbannon@facebook.com)

Past research suggests that when answers can be either numeric or categorical, open-ended questions yield more accurate measurements than closed-ended questions. However, many researchers prefer to ask closed questions to measure income on the assumption that an open question might be perceived to be intrusive and might ask for more precision than respondents can provide. This paper reports an experimental comparison of question formats in terms of item non-response, unit non-response, and the distribution of answers. The optimal question format began by asking an open question; respondents who declined to answer the open question were then asked a closed question. In comparison to an open question alone or a closed question alone, the combination approach led to the lowest unit and item non-response rates and acquired data with maximum refinement without distorting the distribution of answers. This study therefore suggests that the best measurement approach for collecting unbounded numeric answers might be the use of an open/closed pair of questions.

Bounding Estimates: Effects of Question Clarifications on Self-reported Media Exposure

Ronald Z. Szoc, ICF International (rszoc@icfi.com); John F. Kunz, ICF International (jkunz@icfi.com); Samantha L. Schwartz, ICF International (sschwartz@icfi.com); Randall K. Thomas, ICF International (randall.k.thomas@gmail.com)

Questions concerning the number of hours respondents spend doing a variety of activities within specific time frames are common within surveys. Respondent estimates are believed to be somewhat accurate in indicating how long they have spent doing such activities, especially for shorter intervals (e.g. 'a typical

week'). Many surveys attempt to measure how much people spend in media-related activities – watching TV, listening to the radio, reading newspapers/magazines, or using the Internet. Since most people do not readily know the number of hours in a week, we were interested in how the use of question clarifications could affect estimates – does providing people with boundaries with the length of time spent in terms of 'waking' hours per week or in terms of total hours per week cause a lower and more realistic estimate of the number of hours engaged in such activities? In a web-based survey, respondents were randomly assigned to one of 3 conditions: 1. Control ("About how many hours in a typical week do you spend doing the following?" with no additional respondent instructions); 2. Waking Hours (with "The average number of hours people are awake in a typical week is about 112 hours." as the respondent instruction); and 3. Total Hours (with "The total number of hours in a week is 168 hours." as the respondent instruction). Contrary to our expectations, we found that providing bounding clarifications actually increased the estimated hours for each item. We further explored how self-reported media exposure was related to a number of items measuring consumerism and brand loyalty and discuss the impact that clarifications have on respondent accuracy.

Something Borrowed, Something New: Exploring Post-Participation Incentives.

Evaluating the Impact of "Music Downloads" as Instantly Delivered Contingent Incentives.

Kelly L. Bristol, The Nielsen Company (kelly.bristol@nielsen.com); E. Nicole Bensky, The Nielsen Company (nicole.besky@nielsen.com); Lukasz Chmura, The Nielsen Company (Lukasz.Chmura@nielsen.com); Dinaz Kachhi, The Nielsen Company (Dinaz.Kachhi@nielsen.com); Michael W. Link, The Nielsen Company (Michael.Link@nielsen.com)

In the past, extensive research has examined variation in incentives and its impact on the response rate by comparing the effects of cash versus non-monetary incentives, contingent versus non-contingent incentives, and combinations of the above. In order to maximize the output of the invested dollar amount, downloadable music was tested as a means for delivering contingent, non-monetary incentives instantly and determining its impact on the response rate. Compared to traditional contingent incentives delivered via mail, instant delivery of a contingent incentive may have greater impact and utility. Additionally, the number of songs received (5 vs. 10 songs) might impact the perceived value of the incentive and motivation to co-operate. The findings will help to study the impact of cash vs. non-monetary incentives and its appeal among a cross section of demographics.

Effect of a Post-Paid Incentive in a Patient Experience of Care Survey

Julie A Brown, RAND Corporation (julieb@rand.org)

RAND and Kaiser Permanente Southern California field-tested a survey on consumer exposure to and use of health information technology (HIT) as part of the Consumer Assessment of Health Providers and Systems project (CAHPS®). We assessed respondents' experience with their doctor and use of several types of HIT including: appointment setting via a website or e-mail, e-mail communication with doctor's office, requesting prescription refills via a website or e-mail, HIT use by doctors and its perceived helpfulness, consumer use of doctor's office website, ease of using website, accuracy and timeliness of patient information or results posted on website. We used the Internet as the primary mode of data collection and mail as the secondary mode. We included a test of a post-paid incentive against no incentive. The sample of 3,600 patients was randomly divided into the experimental group, which was offered a \$5 incentive (cash for mail and choice of cash or Target e-certificate for Internet mode) for completing the survey, and the control (no incentive) group. The response rate was 8% higher for the experimental group (57% vs. 49%) and two-thirds of the Internet respondents in the experimental group

chose cash. The proportion of surveys completed via the Internet did not differ significantly between the experimental and control groups (73% vs. 72%). There were no significant differences in the age, gender, education, health status, reports of care experience, or ratings of care experience between respondents in the experimental and control groups. The incentive group had more Asian respondents than the control group (8% vs. 4%). The incentive increased the response rate (and reduced bias associated with nonresponse). Participants in the experimental group do not differ significantly from the control group. Limited data are available on nonresponders and analysis of that data to identify differences between responders and non-responders is pending.

An Experiment Evaluating Different Monetary Gift Amounts Given to Respondents in the Medical Expenditure Panel Survey (MEPS)

Wendy Hicks, Westat (WendyHicks@Westat.com); Ralph DiGaetano, Westat (RalphDiGaetano@Westat.com); Sue Connor, Westat (SueConnor@Westat.com); Pat Ward, Westat (PatWard@Westat.com); Huseyin Goksel, Westat (HuseyinGoksel@Westat.com); Virender Kumar, Westat (VirenderKumar@Westat.com); Trena Ezzati-Rice, Agency for Healthcare Research and Quality (Trena.Ezzati-Rice@ahrq.hhs.gov); Doris Lefkowitz, Agency for Healthcare Research and Quality (Doris.Lefkowitz@ahrq.hhs.gov)

The Medical Expenditure Panel Survey (MEPS) is a national study of health care use and expenses, sponsored by the Agency for Healthcare Research and Quality (AHRQ) and the Centers for Disease Control & Prevention (CDC).

The MEPS household component enrolls a new panel of 8,000-10,000 households each year. Each household is interviewed 5 times over a period covering two years. From the overlapping panels, MEPS produces annual estimates of health status, health care use and expenditures, and health insurance.

To encourage participation, respondents receive a monetary gift for their participation. Since 2007, the payment amount has been \$30 per interview. However, in the past several years MEPS has experienced growing difficulty in maintaining response rates across the multiple rounds of data collection.

For the 2008 MEPS panel, AHRQ introduced a respondent gift experiment which assigned every household in the panel to a \$30, \$50 or \$70 payment group. Results from the first round of data collection indicated that both the \$50 and \$70 payment groups attained statistically higher response rates as compared to the \$30 group ($p < 0.01$), but the response rates between the \$50 and \$70 groups did not differ. However, the composite response rate across all five rounds of data collection for the \$70 incentive group was significantly higher than both the \$30 and \$50 incentive group by 12 and 4 percentage points, respectively ($p = 0.01$). In terms of costs, the \$50 payment amount paid for itself, while the \$70 payment resulted in an estimated net increase in data collection costs relative to the current payment.

This paper presents results from the full data collection cycle for the 2008 MEPS panel, including all 5 rounds of data collection. Results focus on response rates, level-of-effort measures and selected survey estimates and item missing rates as indicators of data quality.

Refusal Conversion Incentives and Participation in a Longitudinal Study of Older Adults

Michael Colicchia, NORC at the University of Chicago (colicchia-michael@norc.org); Meredith Czaplewski, NORC at the University of Chicago (czaplewski-meredith@norc.org); Angie Jaszczak, NORC at the University of Chicago (jaszczak-angie@norc.org)

The National Social Life, Health, and Aging Project, a longitudinal, multi-modal study of approximately 3000 older adults born between 1920 and 1947, examines the interaction between aging, relationships, and health outcomes. In 2005–2006, Wave 1 respondents were initially offered an incentive of \$100. Near the end of Wave 1 increased incentive payments, ranging from \$200 to \$500, were offered to pending cases. Two hundred twenty one respondents received an incentive greater than \$100 in Wave 1 yielding a final response rate of 75.5 percent.

The Wave 2 sample includes returning respondents, current romantic partners of returning respondents, and non-hostile individuals who did participate in Wave 1. All respondents, including the 221 respondents who received an increased incentive in Wave 1, were initially offered the \$100 incentive. We are planning to increase incentives during the final weeks of Wave 2.

While existing literature demonstrates the positive effects of monetary incentives on participation in longitudinal surveys in general, there is limited literature on the effects of refusal conversion incentives on participation in subsequent waves, and near none that focus on the older adult population. In this paper, we will examine incentives across waves and analyze respondent characteristics to determine the relationship between these factors and participation. Particularly, we are interested in respondents who “held out” for an increased incentive in Wave 1 while accepting the baseline \$100 in Wave 2 and, conversely, respondents who participated in Wave 1 for \$100 but did not participate in Wave 2 until offered an increased amount. We will also examine whether the presence of a selected Partner (and thus a combined \$200 incentive for the household) may have contributed to survey participation. Finally, we will determine whether Wave 1 non-responders who participated in Wave 2 differ in the incentives accepted.

When Do Incentives Not Work?

Inger Marie Christensen, SFI Survey (imc@sfi.dk); Maja Mortensen, SFI Survey (mmo@sfi.dk)

A case study from Denmark. This experiment was carried out during data collection of the PIAAC pilot in Denmark in April until July 2010. 2675 respondents were drawn at random to participate in the PIAAC, the Programme for the International Assessment for Adult Competencies. The respondents were divided into 4 groups where 3 groups were offered 3 different incentives in the letter of introduction in order to boost the response rate.

The groups:

- i) Control group: The respondents in this group were offered no incentives to participate in the survey
- ii) Group 1: The respondents were offered a voucher with the value of 20 us dollars after completing the interview
- iii) Group 2: The respondents were offered to participate in a lottery (winner chance 1 pct.) The price is 2000 us dollars after completing the interview
- iv) Group 3: The respondents were offered to choose between incentive one and two.

1516 respondents completed the interviews.

The presentation will answer the following questions:

Question 1: What is the most popular incentive in group 3: A small cash reward or participation in a lottery with a larger reward?

Question 2: Which group has a significant higher response rate than the control group? Are there significant differences among the 3 groups?

Question3: Do the 3 different incentives have a different impact on respondents depending on different geographical and demographic characteristics?

Surveys in the Islamic World: Methods & Findings.

Are Opinion Surveys Undercounting Muslims Around the World?

Conrad Hackett, Pew Research Center's Forum on Religion & Public Life

**(chackett@pewforum.org); Brian Grim, Pew Research Center's Forum on Religion & Public Life
(bgrim@pewforum.org)**

There is growing interest in surveying Muslims in countries where Muslims are a minority and in conducting surveys in Muslim majority countries. However, we investigate whether nationally representative surveys with fewer than 2,500 respondents generally under sample Muslims and produce underestimates of the size of Muslim populations, even after the application of recommended survey weights. We analyze this using the World Religion Database (www.worldreligiondatabase.org), in which staff at the Pew Research Center's Forum on Religion & Public Life, IIASA's Age and Cohort Change Project, and Boston University's International Religious Demography Project have recorded the share of Muslims reported in censuses and Demographic and Health Surveys around the world as well as cross-national opinion surveys such as the World Values Survey, Gallup World Poll, Pew Global Attitudes Project, and International Social Survey Program. Compared with census data, we find that opinion surveys tend to count fewer Muslims, sometimes when Muslims are a small minority and sometimes when they are a sizeable majority in a country. We evaluate the extent to which this may stem from overestimations by censuses, which have their own set of methodological problems ranging from question wording (e.g., only providing forced-choice options) to methodology (e.g., curbside interviews). Alternatively, this may stem from underestimations in opinion surveys due to problems surveying Muslim women, the omission of disproportionately large Muslim youth populations from the survey universe, the misclassification of Muslims, the challenge of contacting Muslims in countries where they are a minority, variation in measurement methods, or variation in context, form, and response rate across data collection methods. We describe differences between census and survey data for key characteristics of Muslim adults to assess how representative Muslim respondents in opinion surveys are of the population of Muslim adults in a country. We conclude with recommendations for survey researchers.

Afghan Women's Perspective on Negotiating with the Taliban

Stephen R. Hornbeck, D3 Systems, Inc (shornbec@vt.edu)

In 2010, Afghan president Hamid Karzai began a campaign to open negotiations for reconciliation with the Taliban. These negotiations included allowing former Taliban officers and officials to run for public office and also offered amnesty for Taliban combatants in exchange for an agreement to end fighting. These negotiations have been supported in a large part by the U.S. military. However, it is still remains unclear how the Afghan population as a whole views the opening of negotiations to the Taliban.

This paper will address the Afghan public perception of the government of Afghanistan opening negotiations to the Taliban. As the reintroduction of the Taliban into government could lead to a loss of rights of women and destroy the gains women have made since the fall of the Taliban, this paper will focus the differing perceptions between Afghan men and women towards these negotiations. The paper will also present data that demonstrates how the geographic location, ethnicity of Afghan women, and the security situation of Afghan women shape their views towards opening negotiations.

This paper will also speak to Afghan's views on the need for preconditions of negotiations. Of the 75% of the population that favors negotiations with the Taliban, three quarters feel that there must be a settled peace with the Taliban before these negotiations can occur. Again, gender plays a prominent role in this view as a higher percentage of women (82%) require this precondition than do men (70%).

The results presented in this paper are from a recent nationwide poll in Afghanistan conducted by ACSOR Surveys and sponsored by D3 Systems, Inc. The nationwide Afghan survey consists of interviews with 2,056 randomly selected Afghans. The respondents were interviewed in person by native Afghans from May 12 to May 24, 2010. With a 95% confidence interval, results from this survey can.

Public Attitudes in Yemen

Gary Langer, Langer Research Associates (glanger@langerresearch.com)

There is growing interest in surveying Muslims in countries where Muslims are a minority and in conducting surveys in Muslim majority countries. However, we investigate whether nationally representative surveys with fewer than 2,500 respondents generally under sample Muslims and produce underestimates of the size of Muslim populations, even after the application of recommended survey weights. We analyze this using the World Religion Database (www.worldreligiondatabase.org), in which staff at the Pew Research Center's Forum on Religion & Public Life, IIASA's Age and Cohort Change Project, and Boston University's International Religious Demography Project have recorded the share of Muslims reported in censuses and Demographic and Health Surveys around the world as well as cross-national opinion surveys such as the World Values Survey, Gallup World Poll, Pew Global Attitudes Project, and International Social Survey Program. Compared with census data, we find that opinion surveys tend to count fewer Muslims, sometimes when Muslims are a small minority and sometimes when they are a sizeable majority in a country. We evaluate the extent to which this may stem from overestimations by censuses, which have their own set of methodological problems ranging from question wording (e.g., only providing forced-choice options) to methodology (e.g., curbside interviews). Alternatively, this may stem from underestimations in opinion surveys due to problems surveying Muslim women, the omission of disproportionately large Muslim youth populations from the survey universe, the misclassification of Muslims, the challenge of contacting Muslims in countries where they are a minority, variation in measurement methods, or variation in context, form, and response rate across data collection methods. We describe differences between census and survey data for key characteristics of Muslim adults to assess how representative Muslim respondents in opinion surveys are of the population of Muslim adults in a country. We conclude with recommendations for survey researchers.

The Importance of Ethnicity in Afghan Public Opinion

John Richardson, D3 Systems (john.richardson@d3systems.com)

Society in Afghanistan is dominated by two ethnicities, Pashtuns and Tajiks, who make up the majority of Afghans. However, there are significant numbers of Hazaras and Uzbeks living in the country in addition to several other minorities. The demographic composition of the country is varied with some of its 34 provinces containing almost exclusively one ethnicity while others are very multi-cultural.

This paper is an analysis of the role ethnicity plays in Afghan public opinion. Opinions in Afghanistan vary greatly by geography. That is to be expected in a country with an active insurgency that is stronger in some locations than in others. However, how much difference in these opinions is also related to the ethnicity of the respondents?

These questions are explored in a variety of subjects including perceptions of security, governance, and development projects in a survey conducted by D3 Systems. D3 Systems has conducted nationwide, face-to-face public opinion surveys in Afghanistan every month since September 2005. D3's Afghan nationwide surveys consist of interviews with randomly selected Afghan adults aged 18+ throughout Afghanistan. Respondents were interviewed in person by native Afghans. Interviewers were gender-matched with respondents. D3's survey is n=2056 with an associated margin of error of +/- 2% at the 95% confidence level.

One of the difficulties in examining sub-populations within a national survey is that the sample sizes for minorities will be small with large margins of error. Using publicly available data from the Asia Foundation, ABC News, and the U.S. Army Corps of Engineers' Measuring Progress in Conflict Environments (MPICE) this study shows how it is possible to mitigate the uncertainties for these minority populations to increase confidence in their comparison with large sub-populations.

Measuring Progress to Peace in Afghanistan; Applying the US Institute of Peace MPICE Framework

Karl G Feld, D3 Systems, Inc. (Karl.Feld@d3systems.com); Brianna Augenreich, D3 Systems (brianna.augenreich@d3systems.com); Veronica A. Gardner, D3 Systems (veronica.gardner@d3systems.com); Anne Pessala, D3 Systems (anne.pessala@d3systems.com); John Richardson, D3 Systems (john.richardson@d3systems.com); William Hayes, D3 Systems (will.hayes@d3systems.com); Eric Chappell, D3 Systems (eric.chappell@d3systems.com)

The *Measuring Progress in Post-Conflict Environments* (MPICE) evaluation framework was created by the US Institute of Peace as a metric system to be applied to any pre- or post-conflict environment. It is designed to understand what drivers of social conflict must change in a country to improve chances for stability and reconstructions. These levers of change are based on empirical data and a tested model. They are intended to inform policy-makers with data to be used in the face of competing policy opinions and interests.

D3 Systems has completed the first year applying the MPICE framework of survey, qualitative and secondary data collection and analysis in Afghanistan. The project is funded by the US Army Corps of Engineers in support of the International Security Assistance Force (ISAF) mission in Afghanistan. This paper discusses the empirically grounded findings of the framework and their implications for the policy debate on future activity in Afghanistan.

Particular emphasis will be placed on research which measures creating a safe and secure environment, promoting political moderation and stable democracy, grounding the rule of law and a sustainable economy and promoting social well-being. Hundreds of data components feed into MPICE Afghanistan annually. MPICE uses ninety-six social, economic, political and cultural metric selections for Afghanistan.

The measurement process includes an annual, nationwide survey n=10,000 of interviews with randomly selected Afghan adults aged 18+ throughout Afghanistan. Native Afghan interviewers are gender-matched with respondents. The associated margin of sampling error ranges from +/- 2 to 3% at the regional level (+/- 1% nationwide) with a 95% confidence level. The survey data is combined with secondary quantitative data from over 50 sources and qualitative work from 200+ populations and thousands of interviews annually.

Understanding Public Perceptions on Global Warming

The Persistence of American Public Opinion on Climate Policy

Bo MacInnis, Stanford University (bo@macinnis.org); Jon Krosnick, Stanford University (krosnick@stanford.edu)

With evidences from fifty-years of trends in Americans' policy preferences, Page and Shapiro (1992) postulate: aggregate public opinion on policy is stable over time and collective opinion responds to important events in predictable ways; a stronger form of the former is "parallel publics" whereby various subgroups in the population, who might differ in their levels of policy preferences, follow a similar path of the changes in their preferences. In this study, we assess Page and Shapiro's stability proposition by reviewing Americans public opinion on climate policies to examine: Has Americans' support for climate policy been stable over the past decade? And have the subpopulations demonstrated a similar trend in their support for climate policies? The contributions of the present investigation are two-fold. First, global warming has been an important societal issue that commands collective actions to mitigate its adverse consequences, and yet there is no systematic study on Americans' preferences for climate policies. This study fills the void to be the first to test Page and Shapiro's propositions on the persistence of public opinion on global warming. Second, our study will aid to shift public focus towards climate policies. Recent controversy has centered on whether Americans believe global warming has been happening. While it is instructive to understand the newly emerged patterns in public beliefs on the existence of climate change, we postulate that the collective public preference for ameliorative policy would be a more valid and direct measure on how much Americans support the enactment of legislation aimed at ameliorating climate changes, as we presume that policy-makers' willingness to take legislative actions on climate change would be most receptive to the public's willingness to support these legislative actions.

All Climate is Local: Measuring the Effect of Weather Events on Perceptions of Global Warming.

Christopher Paul Borick, Muhlenberg College Institute of Public Opinion

(cborick@muhlenberg.edu); Barry G. Rabe, University of Michigan (brabe@umich.edu)

During much of the first decade of the 21st Century, Americans increasingly indicated that they believed temperatures on Earth were rising. This growth in belief was particularly pronounced between 2007 and 2008 when numerous polls converged in showing that Americans were more likely than ever before to express belief in global warming. But this trend appears to have shifted markedly in more recent years. Numerous surveys between 2008 and 2010 have shown a declining percentage of Americans expressing belief that temperatures on Earth have been increasing. While the consistency of results from national level surveys suggests that Americans are indeed less likely to believe in global warming than in recent years, the reasons behind the declining results is less clear. One factor that has been suggested for the shift in opinion is that weather phenomenon such as severe winter storms and periods of below average temperatures have acted to weaken underlying confidence that the planet is experiencing global warming. In this study we attempt to measure the effect of meteorological events on perception of climate change among residents of the United States.

Results from the 2008, 2009 and 2010 National Surveys on American Public Opinion on Climate Change (NSAPOCC) are examined in order to shed light on the relationship between weather events and belief in global warming. These annual telephone surveys include a series of questions that measure the relative effect of various factors on an individual's position on the existence of global warming. The results suggest that declining levels of belief in global warming during recent years corresponded with individual perceptions of weather events such as cold waves, winter storms and relatively inactive hurricane seasons in the Atlantic Ocean.

Measuring Americans' Issue Priorities: A New Version of the Most Important Problem Question Reveals More Concern About Global Warming and the Environment.

David Scott Yeager, Stanford University (dyeager@stanford.edu); Jon Krosnick, Stanford University (krosnick@stanford.edu)

For decades, numerous surveys have asked Americans the "Most Important Problem" (MIP) question: "What do you think is the most important problem facing this country today?" Global warming and the environment have rarely been cited by more than a tiny number of respondents in these surveys in recent years, which might seem to suggest that these have not been the most important issues to Americans. This paper explores the possibility that an additional method of assessing the public's priorities might support a different conclusion. Three experiments embedded in national surveys (two done via the Internet, the other done by telephone) show that when asked the traditional MIP question, respondents rarely mentioned global warming or the environment, but when other respondents were asked to identify the most serious problem that will face the world in the future if nothing is done to stop it, global warming and the environment were the most frequently mentioned problems. Furthermore, a large majority of Americans indicated that they wanted the federal government to devote substantial effort to combating problems that the world will face in the future if nothing is done to stop them. Thus, future surveys might include both versions of the MIP question to more fully document Americans' priorities.

Declared vs. Real Purchasing Behavior Affected by Producers' Action against the Environment
Jiri Remr, Charles University in Prague, Faculty for Social Sciences (remr@centrum.cz)

Some of the current research data show the resolution of many individuals to boycott or refrain from purchasing the products from companies that pollute the environment or that are unfair to their employees or that simply behave in a way that gives rise to the societal conflict. At the same time, another data indicate that the verbal reports about the change in purchasing intentions could be biased because people consider certain answers as socially desirable within the given context.

Primary goal of the presented research is to estimate the extent of social desirability response bias. Useful indication of immensity of the problem with pretended behavior provided many years ago LaPier in his research with Chinese couple. Our current attempt is to show to what extent the social desirability response bias can distort the self-reported behavioral data. Is it just a minor effect that may not bother us or is it actually affecting our data in a substantive way? Our analysis goes even further: actual data enable to show the profile of the respondents whose propensity to give biased answers is the most striking. What are the socio-demographical characteristics of those who pretend a desirable behavior? What are the differences in probability to pretend among different subgroups based on a relation to the given subject (environmental issue)? Finally, we come up with some remarks on how to diagnose the effect in survey researches and how to improve the quality of its results.

The extent of above described effect is shown on data coming from nationwide (Czech Republic) representative survey: N>1.500 respondents, quota sampling, face-to-face interviews; field works were conducted during October 2010.

The Stability Of American Public Opinion On Global Warming: Towards Explaining The Existence Beliefs Trends.

Jon Krosnick, Stanford University (krosnick@stanford.edu); Jon Krosnick, Stanford University (krosnick@stanford.edu)

There has been an observation and agreement among scholars that Americans' judgment about global warming showed a noticeable drop between 2008 and 2009. What has remained the focal point of disagreement among public opinion pollsters is whether the observed decline is temporary or permanent. A source of discord is the lack of understanding of what factors might have caused the decline. The present study fills the void by formulating a set of hypotheses drawing upon the existing theories and evidences on how publics form their beliefs about global warming and testing these hypotheses using the data from our 2006-2010 surveys. Generally, people form their perceptions through personal experiences and indirect learning from informants with the perceived credibility of and consensus among informants regulating the impact of informants on public opinions. Public judgment on global warming have been found to be associated with political party affiliation and economic conditions with a general finding of a lower acknowledgement among Republicans and in economic adversaries. The existing theories and evidences suggest several common suspected forces for which the recent decline might be attributable to. However, we have ascertained that these common suspects, such as declining public trust in scientists, declining public willingness to pay for remedial policies, rising influence of Republicans, or increasing visibility of skeptics, do not contribute to the ups and downs in the public opinion of the existence of global warming, but the unusually cool temperature of 2008 appears to. The results of this study are to substantiate the argument that the observed 2008-2009 decline in global warming existence belief is a result of 2008 breaking the warming trend of world temperature, and is to provide evidence to cast doubt to some purported notions that Americans had permanently shifted their perceptions of and preferences and support for global warming.

Thursday, May 12, 2011

3:00 p.m. - 4:00 p.m.

Poster Session 1

Public Opinion on Gun Control Revisited: Collective Consensus or Unbridgeable Ideological Divide?

Bryan C Parkhurst, University of Cincinnati (parkhubc@gmail.com)

Though support for gun control has been fairly high over several decades, polling data in recent years has shown a noteworthy decline in public support for stricter gun control. At the same time researchers have only begun to get a handle on what drives public opinion on gun control. Using data from the NORC GSS, Celinska (2007) has developed a particularly useful measure of respondents' ideology of individualism vs. collectivism for predicting both gun ownership and opinions on gun control. Her analysis clearly demonstrated that respondents falling toward the individualism end of the spectrum were significantly more likely to own firearms and to oppose their regulation by the government, whereas those inclined toward collectivism were noticeably more likely not to own firearms and to support gun control. Her analysis, which was based on GSS data from 1984 to 1998, may now be somewhat dated, however, in light of recent drops in public support for gun control, growth of gun purchases in reaction to the election of President Obama, and the recent Supreme Court decision on the applicability of the 2nd Amendment to local and state gun control laws. In this presentation, the author revisits American public opinion on gun control in this altered political climate and reexamines Celinska's conclusions using updated NORC GSS data from the past decade. The findings suggest that consensus on this enduring issue remains illusory, despite contrary polling data reported by the Gallup Organization, among others.

The Impact of Survey Design Modifications on Health Insurance Coverage Estimates in a National Longitudinal Health Care Survey

Steven B Cohen, Agency for Healthcare Research and Quality (scohen@ahrq.gov); Trena Ezzati-Rice, AHRQ (tezzatir@ahrq.gov); Marc Zodet, AHRQ (mzodet@ahrq.gov)

National health insurance coverage estimates for the overall population and specific population subgroups are critical to policymakers and others concerned with access to medical care and the cost and sources of payment for that care. The Medical Expenditure Panel Survey (MEPS) is one of the core health care surveys in the nation that serves as a primary source for these essential insurance coverage estimates. The survey is designed to provide annual national estimates of the health care use, medical expenditures, sources of payment and insurance coverage for the population. In 2007, the survey experienced two dominant survey design modifications: a new sample design attributable to the sample redesign of the National Health Interview Survey, and an upgrade to the CAPI platform for the survey instrument, moving from a DOS to a Windows-based environment. This study examines the impact of these survey design modifications on the national estimates of insurance coverage.

The overlapping panel design of the MEPS and its longitudinal features are particularly well suited to assess the impact of survey redesign modifications on estimates. Since two independent nationally representative samples are pooled to produce calendar year estimates, one has the capacity to compare estimates based on the “original survey design” in contrast to those derived from the “survey redesign.” This paper examines the correlates of nonresponse incorporated in the estimation techniques and adjustment methods employed in the survey, and the measures utilized for post-stratification overall and by panel. Particular attention is given to assessing the level of convergence in coverage estimates based on the alternative designs as well as the alignment of model based analyses that discern which factors are associated with health insurance classifications. The paper concludes with a discussion of strategies under consideration that may yield additional improvements in the accuracy for these critical policy relevant survey estimates.

New Techniques for Weight Trimming

Benjamin Phillips, Abt SRBI (phillipsb@srbi.com); Seth Brohinsky, Abt SRBI (s.brohinsky@srbi.com); Graham Wright, Brandeis University (gwwri@brandeis.edu); Chintan Turakhia, Abt SRBI (chintan@srbi.com)

Whether by design or as the result of adjustments, survey weights commonly exhibit large variation in size. The result of a wide range in variation of survey weights is the inflation of sampling error, leading to decreases in the precision of survey estimates. A variety of heuristics have been developed to counteract the impact of extremely high or low weights by truncating weights above or below a certain point. These techniques, however, are neither systematic nor theoretically grounded. This is of particular concern because large weights are not arbitrary but are used to avoid underrepresentation of elements of the population with low probabilities of selection or high probabilities of nonresponse. Arbitrary trimming may reduce variance at the price of unacceptable increases in biases.

This paper will investigate the use of systematic procedures to minimize sampling error induced by wide variation in weights. We use nonlinear programming and other optimization techniques to find the transformation that minimizes mean square error (MSE) of estimates of multiple characteristics of interest. MSE is a metric that balances losses in precision and bias. We investigate four transformations to reduce the influence of extreme weights, three of which differ fundamentally from present approaches. The first systematizes present techniques for weight trimming by finding the truncation point that minimizes MSE. The remaining techniques reduce variation while retaining the ordinal relationship between weights.

through compression. The second uses a linear compression function above a threshold, where both the threshold and the degree of compression are chosen to minimize MSE. The third applies a nonlinear transformation (raising the weight to a power greater than zero and less than one), where the power is chosen to minimize MSE. The fourth applies a nonlinear transformation of the weight above a threshold, where the threshold and power are chosen to minimize MSE.

Accounting for Diffuse Diversity: Accurately Polling Asian Voters in the City of Los Angeles.
Nathan Aderhold, Loyola Marymount University, Leavey Center for the Study of LA
(nate.aderhold@gmail.com)

While the City of Los Angeles is widely known for its racial and ethnic diversity, accurately accounting for that diversity in polling practices can be troublesome. To achieve a representative sample of exit poll voters, in 2005 Loyola Marymount University (LMU) researchers developed the 'racially stratified homogenous precinct approach'. While the method has increased overall data reliability for the city, a sampling dilemma emerged: in LMU's 2008 General Presidential Election Exit Poll, the Asian voter demographic was underrepresented. After evaluating the polling places selected in 2008, it was discovered that many precincts categorized as 'Asian', and thought to be centered in Asian neighborhoods, instead were located outside those areas.

In an effort to enhance precision in precinct selection for LMU's 2010 Gubernatorial Election Exit Poll, a new methodology was devised to accurately locate polling places in highly-populated Asian neighborhoods. Various interactive mapping programs, such as SimplyMap and Los Angeles Times' Mapping L.A., were used to produce a visual representation of demographic data from the 2000 U.S. Census and 2009 American Community Survey. To increase the number of potential Asian respondents a four-step process was employed: 1) Neighborhoods with the highest Asian populations—based on Mapping L.A.—were identified; 2) Using SimplyMap, exact census block data was layered over the neighborhoods for a more precise look at the Asian demographic; 3) Precinct outlines in these designated neighborhoods were layered over the neighborhood boundaries; and 4) Polling places were selected based on the centrality within the neighborhood.

The implementation of these more calculated measures in the precinct selection process for the 2010 exit poll helped diminish the sampling dilemma, achieving a more representative sample overall. Using these mapping programs in a more systematic approach can increase data-set reliability and create a truer representation of diverse populations in cities across the United States.

Perception about Characteristics of Scientists, Results of the Survey of Youth Perception of Science and Technology in Bogotá.
Edgar Mauricio Bueno Castellanos, Colombian Observatory of Science and Technology
(embuenoc@gmail.com)

The Colombian Observatory of Science and Technology -OCyT- developed, in 2009, a survey about the perception of Science and Technology in students of the last two years of high school in Bogotá, Colombia. A set of questions inquired about the perception of these students of the characteristics of a scientist. Answers were analyzed using multivariate data analysis tools. Three classes of students were identified: students that think of a scientist as a person with special intellectual capabilities, students that see scientists as isolated people and, a third class, composed by students that think of scientist as ordinary people. Given that the sampling design -stratified according to the nature of school (official or private)- should not be ignored at the estimation stage, the analysis was carried out taking into account the selection probabilities assigned to schools. In addition, because of the high nonresponse in private

schools -data collection was allowed in 13 out of 31 private schools (42% of schools, 53% of expected students)-, it was decided to impute values corresponding to item nonresponse using the methodology of the nearest neighbor and use calibration methods for unit nonresponse, reducing the nonresponse bias.

Deliberative Polling and Small Sample Sizes: Robustness of the Deliberative Polling Model
Sean J Westwood, Stanford University (seanjw@stanford.edu); Alice Siu, Stanford University (asiu@stanford.edu); James S Fishkin, Stanford University (jfishkin@stanford.edu)

The Deliberative Polling model of public consultation inhabits a space between traditional polling and social science experimentation. In traditional polling, sample sizes for acceptable results are usually assigned a floor of 300 participants, while psychological experiments are often considered successful with samples of fewer than 20 per cell. It is less clear what the lower limit for acceptable generalizability is for the Deliberative Polling model. This paper investigates the issue of sample sizes in Deliberative Polls with data gathered from two polls that had lower than expected turnouts (Power2010 and La Plata 2009). Power analyses using effect sizes from previous polls show that samples of 19-35 are sufficient for paired t-tests for most issues. Although larger samples are used to assert representativeness of the general population along socio-demographic and attitudinal measures, the phenomenon of opinion refinement is sufficiently supported with fairly small samples. For initial sampling, La Plata used a geographically stratified random sample for initial face-to-face interviews (n=1,476), and Power2010 used a representative online panel (n=300). The participants that arrived for deliberation in both polls were remarkably representative of the overall population in terms of demographics and attitudes, with minimal statistically significant differences between participants and non-participants. While larger samples are always desirable, we provide compelling evidence that Deliberative Polls, as quasi-experiments, are tenable measures of informed public opinion with surprisingly small samples.

Improving Questionnaire Design of Establishment Surveys for Field Data Collection
Lorelle Vanno, The Nielsen Company (lorelle.vanno@nielsen.com); Jennie Lai, The Nielsen Company (jennie.lai@nielsen.com); Teresa Jin, The Nielsen Company (teresa.jin@nielsen.com); Michael Link, The Nielsen Company (michael.link@nielsen.com)

In the recent years, survey researchers around the world are starting to focus their attention on survey methodology of establishment surveys for businesses, farms and institutions. The effective design of the questionnaire becomes a major consideration for field interviewers to collect data accurately and efficiently. The Nielsen Company collects data for establishment surveys from retailers in 95 countries worldwide. We will demonstrate how improvements of the questionnaire design (specifically on format, layout, instructions, question order, etc.) can maximize data quality and efficiency of field data collection for establishment surveys. We will first discuss the limited research on questionnaire design and data collection methods for establishment surveys; then highlight past challenges of field interviewers for Nielsen's Retail Establishment Survey and share lessons learned of administering the improved questionnaire. These qualitative insights will allow survey researchers of establishment surveys to adapt their improvement of the questionnaire design based on our findings.

Evaluating Strategies for Reducing Non-Response Among Under-Represented Groups
Jason Minser, Abt SRBI (j.minser@srbi.com); Courtney Kennedy, Abt SRBI (c.kennedy@srbi.com)

The 2009-2010 Cincinnati Travel Household Survey was the first household travel survey (HTS) to collect data exclusively through GPS technology. Traditionally, HTS designs featured hard copy travel diaries that were relatively burdensome for respondents. While the switch to GPS technology is expected to reduce measurement error, there are persistent concerns about differential response rates across key

subgroups (e.g., low income households, zero vehicle households, 4+ person households) and the potential for non-response bias. The Cincinnati survey featured numerous strategies for improving response rates, including advanced letters, incentive programs, and multiple reminder calls. In this analysis, we evaluate the performance of each of these strategies in improving response rates for groups that tend to be under-represented in HTS designs. We hypothesize that a combination of tactics works best in reducing non-response and increasing representativeness among under-represented groups. The cost-benefit of each strategy will be evaluated to determine the most effective strategy to incorporate in a HTS context.

Enumerating Parents in a Household Survey: Challenges with Nontraditional Households, Same-Sex Parents, and Spanish-Speakers.

Cameron McPhee, American Institutes for Research (cmcphee@air.org); Sarah K Grady, American Institutes for Research (sgrady@air.org)

This poster will use data from 58 cognitive interviews to identify issues related to collecting information in a household survey about parents of a sampled child, with special attention to same-sex parents, other nontraditional household structures, and Spanish-speakers. Most research conducted on household enumeration has used roster data; this analysis is unique because it deals with capturing household adults' relationships to the sampled child.

The National Household Education Survey (NHES), undergoing a redesign from an RDD telephone survey to an ABS mail survey, has conducted methodological experiments, including questionnaire redesign. Sections of the questionnaire that were designed to collect information about children's parents (called "Mother/Female Guardian" and "Father/Male Guardian" sections) were adapted to "Parent 1" and "Parent 2" sections in order to accommodate a diversity of family structures.

As part of the upcoming 2010 NHES Field Test (n= 60,000), a random sample of respondents will receive the revised questionnaires with the "Parent 1/ Parent 2" sections. In the questionnaire redesign process, cognitive interviews were conducted with 58 parents, including 6 same-sex parents, 11 parents in other, nontraditional households, and 9 Spanish-speaking parents. Lessons were learned about measuring familial relationships as they relate to a sampled child.

Descriptive statistics about nontraditional households, same-sex couple households, and Spanish-speakers from the RDD 2007 NHES will help illustrate dimensions of the challenges in enumerating parents. Recommendations for collecting information on parent relationships for a diversity of household structures will be presented.

A Comparison of Collecting In-School Parent Permission Forms

Greg Mosorjak, RTI International (mosorjak@rti.org)

Purpose: To examine methods that maximize parental permission rates for student participation in school-based education studies while meeting Federal guidelines for maintaining human research rights for study participants.

Significance: Gaining parental permission for student participation in school-based studies can be a formidable task, whether obtaining active permission for more sensitive, high-risk data collection or "opt-out" parent permission for less sensitive, more benign data collection. In this presentation, we will share lessons learned from large-scale studies that utilized both types of parent permission and insured that the permission forms were properly completed and available to researchers throughout the entire data collection period.

Procedures: Using information gathered from three recently completed national educational studies, we examined three methods of distributing, tracking, and verifying parent permission forms by data collection staff – active, opt-out, and a combination of the two..

Principle Findings: Using "active" consent procedures yielded a lower participation rate because forms had to be returned and clearly marked and signed indicating if student had permission to participate. Using the Protection of Human Rights in Research, we are discussing when one type of consent is preferred over another, when passive consent is allowed, and how to identify a legitimately signed and completed consent form.

Conclusions: There is a noticeable difference in return rates of parent permission forms based on the type of permission sought, and the distribution and return method. While using the “opt-out” procedures will likely have a higher participation rate, most require “active” parent permission.

Testing the Effects of Eliminating the Mandatory Long Form for the Canadian Census of Population

Fred Hazelton, Statistics Canada (hazefre@yahoo.ca)

In summer 2010 the Canadian Census of Population was the subject of much media attention. The government's decision to eliminate the mandatory long form caused a national discussion about the role of the Census. Not only did the changes affect the public's perception of the census but it also affected the logistics associated with the paper questionnaire. The Questionnaire Design Resource Centre was asked to examine the impact of the changes with respect to two factors. 1) The public's response to the media attention and; 2) the ability of respondents to complete the new format of the short form. This poster will show the flow of events from the government's decision to testing the new form. Several versions of the form were tested including various placements of instructions, symbols, and contrasting colors. Testing was completed under a rigorous time constraint and with the pressure of potential increased non-response. The poster will demonstrate how subtle differences in format can have great affects on both measurement error and public perception.

Identifying Efficiencies in Completing a Complex Telephone Interview

Julie A. Pacer, Abt SRBI (pacer@srbi.com); Suzanne Bogue, Abt SRBI (bogue@srbi.com); Cathy Joseph, Abt Associates (Cathy_Joseph@abtassoc.com); Sarah Minden, Brigham & Women's Hospital, Harvard Medical School (sarahminden@gmail.com)

Surveys of populations with a chronic condition can be lengthy, complex, and costly. For these special populations, it is important to keep respondent burden low to achieve acceptable response rates within a reasonable field effort. Ideally, respondents would complete an interview in one telephone call. However respondents with a chronic condition may require multiple interview sessions, which drives up field costs. It can be challenging to manage this type of sample both considerately and efficiently. Since 2000, Abt Associates has collaborated with the National Multiple Sclerosis Society on the Sonya Slifka Longitudinal Multiple Sclerosis Study. This study tracks the characteristics and implications of MS and its treatments over time. It utilizes a nationally representative sample of approximately 4,600 people, over two cohorts, at various stages of MS. At the outset of the study respondents were interviewed every six months, however in 2009 it was redesigned into an annual interview. This annual interview is complex. It takes approximately 60 minutes to complete and collects information on disease status, activities of daily living, quality of life measures, general health, and detailed utilization information on health care visits and medications obtained in the past year. Some questions request consultation of medical records, prescriptions, and receipts. Respondents are encouraged to prepare their information in advance of the interview.

This research seeks to explore how respondents who completed this complex interview in one interview session differ from respondents who completed the interview over multiple interview sessions. Both respondent and survey characteristics will be considered. We will review lessons learned to streamline respondent participation and retention in both the survey and study. Results of this analysis may be of interest to researchers who interview respondents with a chronic condition or field complex telephone interviews, particularly those which collect detailed information.

Using a Probability-Based Online Panel to Survey American Jews

Graham Wright, Cohen Center for Modern Jewish Studies - Brandeis University (gwwri@brandeis.edu); Benjamin Phillips, Abt SRBI Inc. (b.phillips@srbi.com); Josua Tobias, Cohen Center for Modern Jewish Studies - Brandeis University (jtobias@brandeis.edu); Jordan Peugh, Knowledge Networks (jpeugh@knowledge networks.com); Karen Semans, Knowledge Networks (ksemans@knowledge networks.com)

Americans Jews represent less than 2% of all adults in the U.S. This low prevalence makes creating a probability-based sample of American Jews cost-prohibitive. We explore the use of a probability-based online panel to reach a representative national sample of respondents who are Jewish by religion or background, the latter being identified by screening panel members with no religious affiliation.

In 2010, we conducted three surveys of Jewish respondents using Knowledge Networks' (KN) KnowledgePanel®. This online panel of approximately 50,000 U.S. residents aged 13 and older is built through random-digit dialing and address-based sampling and includes participants who would not otherwise have online access. The panel had three distinct advantages. The large panel size and prior information on religious affiliation allowed a larger sample than would be possible in a sample specifically drawn for this purpose. Panel cooperation rates were high and demographic information available on panel nonrespondents allowed for adjustment for nonresponse. Issue saliency bias was reduced by the generic nature of panel recruitment.

We examine the coverage properties of the KnowledgePanel® sample with respect to demographic and religious factors in comparison to surveys of American Jews from nonprobability online panels, traditional mail panels, RDD designs, and meta-analyses of surveys with items on religion. The panel appeared to be representative with respect to sex, region, and education. However, young adults aged 29 and below were under-represented. The panel initially appeared to be biased with respect to Jewish religious denominations. However, estimates of the denominational composition of American Jews were strongly influenced by question wording and response options as well as rapid denomination change among American Jews.

We also conduct sensitivity analyses of the impact of weighting assumptions on outcomes of interests. In general, estimates were quite robust to weighting assumptions, strengthening confidence in the sample.

How was Dinner? Results from the Bento Box Questionnaire Testing for the 2012 Census of Agriculture

Jaki S. McCarthy, USDA Nat'l Agricultural Stat. Service (jaki_mccarthy@nass.usda.gov)

In preparation for the 2012 Census of Agriculture, multiple methods of testing were used to improve the questionnaire and data collection procedures. As in a Japanese bento box meal, these 5 methods were used to provide a balanced and complementary testing "meal". The "bento" testing approach for the Census was outlined in a previous paper, and this paper summarizes the wholesome results of using this approach. Review of feedback from field staff, calls to the toll free number, the amount of item editing and imputation, cognitive interviews, pretesting and large scale field testing were all combined in the redesign of the questionnaire and data collection procedures. Several examples will be presented. How the results have changed the data collection and questionnaire for the 2012 Census of Agriculture will be discussed. The paper will conclude with a discussion of how the combination of testing methods was clearly a more satisfying meal than the single testing dishes alone.

The Reunion Crashers: Locating and Tracking Longitudinal Study Respondents After a 30-year Hiatus.

Leslie Scott, American Institutes for Research (lscott@air.org); Susan Lapham, American Institutes for Research (slapham@air.org); Sandra Eyster, American Institutes for Research (seyster@air.org); Sabine Horner, American Institutes for Research (shorner@air.org)

This presentation evaluates the success of tracing activities used to find current contact information for participants in a longitudinal study who were last contacted more than 30 years ago. Locating study participants after a 30-year hiatus is crucial to the success of a 50-year follow-up study to examine the relationship between cognitive and non-cognitive skills and life course outcomes. Project TALENT is a large, nationally representative longitudinal study of men and women who were in high school in 1960. The American Institutes for Research (AIR) collected extensive information in four half days of testing on the characteristics and cognitive abilities of approximately 440,000 high school students. Follow-up surveys were conducted 1, 5, and 11 years after high school graduation. Between 1960 and 1978 students were sent an annual newsletter with a request for contact information updates. However, 1978 was the last contact made with the participants. Since 2009, we have employed multiple methods to locate individuals who are now in their 60s. In 2009 and 2010 AIR contracted with LexisNexis, obtaining current contact information for 77% of participants not identified as deceased. In 2010 we sent representatives or materials to 50-year high school reunions and obtained over 10,000 contact records. We present results of our tracing activities, including record matching from LexisNexis and reunion activities and using the reunion results to assess the validity of the LexisNexis records search to evaluate the effectiveness of each of the different tracing methods. Preliminary results show a high level of correspondence between addresses obtained through the two modes, showing that that LexisNexis is a valid source of contact information for this longitudinal study, and that tracing through reunions yields additional information not obtained through LexisNexis.

The Activeness of the Issue Publics in the New Media Environment: Online Political Expression as a Mediator Between Issue Public Membership and Political Participation.

Hsuan-Ting Chen, University of Texas at Austin, School of Journalism (hsuanting.chen@gmail.com); Sun Ho Jeong, University of Texas at Austin, School of Journalism (sunho.jeong1@gmail.com); Homero Gil de Zúñiga, University of Texas at Austin, School of Journalism (hgz@austin.utexas.edu)

The concept of issue publics, depicting the way in which individuals relate to issues they esteem important while paying little attention to others, was raised to argue against citizens' decreasing competence. This concept has gained scholars' attention since the new media environment, particularly in the high-choice media, as it provides individuals with a wealth of information and diversified contents from decentralized media outlets, allows individuals to select information by weighting competing preferences among different content. Scholars have also indicated that members of issue publics are motivated to use high-choice media, such as the Internet that allows them to selectively expose themselves to relevant information. However, to fully understand citizens' competence, the relationship between issue public membership and political participation has not been gathered. Nor has research to date inquired about specific activities of the issue publics on the Internet.

This study aims to understand the activeness of issue publics by exploring issue publics' online activities, especially political expression that might affect political participation. This study assumes that members of the issue publics would have a higher level of political participation because they are more cognitively and behaviorally involved in certain issues and more likely to obtain relevant mobilizing information. In addition, the political expression as an active online behavior that stimulates their information elaboration would serve as a mediator between issue public membership and political participation.

Through the use of nationally representative survey data (N = 232), the analysis of multiple mediation model (Preacher & Hayes, 2008) suggested that the relationship between issue publics and offline political participation is mediated through political expression online and on social network sites. In addition, political expression on social network sites has a stronger mediating effect on the relationship between issue publics and offline political participation than online political expression does.

Using "Sufficiently Knowledgeable Adult" As an Alternative Respondent Selection Rule for Child Health

Donna Eisenhower, NYC Department of Health and Mental Hygiene (deisenho@health.nyc.gov); Stephen Immerwahr, NYC DOHMH (simmerwa@health.nyc.gov); Tara Merry, Abt SRBI (t.merry@srbi.com); Andy Weiss, Abt SRBI (A.Weiss@srbi.com)

National telephone surveys of child health have long used "most knowledgeable adult" (MKA) as the respondent selection rule. By definition, data from the MKA should be the most accurate proxy data, and data on the MKA can be singularly attached to a child. We evaluated an alternative respondent selection rule -- "sufficiently knowledgeable adult" (SKA) -- as a less costly means for collecting proxy data on child health.

[Methods] Data come from the New York City Department of Health and Mental Hygiene's 2009 Child Health Survey (children under 13), a citywide telephone survey in 3002 households (2499 with multiple adults). Adult respondents were asked if they were sufficiently knowledgeable, or could answer questions, about the selected child's health or health care. Another adult was interviewed when this person answered no. In multiple adult households, the responding adult was asked if other adults in the household were equally knowledgeable, less knowledgeable, or more knowledgeable about the selected child's health and health care. Data quality was measured by comparing means and level of item non-response of those identifying themselves in each of the knowledge categories. Cost data were measured by the number of call-backs and compared to surveys using MKAs.

[Results] Among respondents in households with multiple adults, 61% said they were equally knowledgeable to other adults (no MKA); 13% said another adult was more knowledgeable (the other is an MKA); and 26% said the other adult (s) were less knowledgeable about the child's health (they are MKA). There were no significant differences in data quality among the groups, but fewer callbacks were required for SKAs in this survey than in surveys using MKAs.

[Conclusions] SKA selection reduces costs and potentially non-response, and may also be more appropriate as child rearing activities are increasingly shared more equally between couples and within households.

Meta-Analysis of Cognitive Interview Findings and Recommendations for Establishment Surveys from 2005-2009.

Kristen Louise Cibelli, U.S. Census Bureau/University of Maryland (kristen.louise.cibelli@census.gov)

Cognitive interviewing has become a well-accepted method for pre-testing survey instruments for both household and establishment surveys. Most studies examining lessons learned from cognitive interviews are survey-specific, focusing on the results from testing one particular survey instrument. However, studies consolidating response error results across surveys are not common, particularly among establishment surveys where cognitive research is just beginning to gain critical mass. This poster reports on the results of a meta-analysis of Findings and Recommendations from cognitive testing conducted by

Economic Program staff at the U.S. Census Bureau between 2005 and 2009. Results from the meta-analysis allow for the examination of overarching patterns in the types of measurement errors identified across a range of establishment surveys that varied in topic, target population, frequency and other establishment survey characteristics. The meta-analysis classifies the nature of findings based on a set of codes related to instructions, the question, a combination of the question and instructions, the data available to the respondent, formatting, burden and the respondent. We coded a total of 548 findings reported for 18 cognitive studies, including findings that were more generally associated with the instruments along with those associated to specific questions. We further examine the source of ambiguous concept, the most common type of error identified in the meta-analysis, to learn more about the context in which these errors arose. We look at the nature of the survey, business sector, type of question, and individual respondent and business attributes associated with the ambiguous concept for a subset of five surveys.

Charting the Vote: Maintaining Diversity and Measuring Exit Poll Reliability

Brianne Gilbert, Loyola Marymount University - Leavey Center for the Study of LA

(brianne.gilbert@lmu.edu); Melissa Woehrstein, Loyola Marymount University - Leavey Center for the Study of LA (melissa.woehrstein@lmu.edu)

After the discrepancies in exit poll results in the 2000 and 2004 Presidential elections, controversies surrounded exit polling methodology, specifically that of poor sampling techniques and inaccurate results. To address these methodological issues, in 2005 a team of researchers at Loyola Marymount University (LMU) implemented the use of an innovative sampling technique known as the “racially stratified homogenous precinct approach” to accurately sample voters in the city of Los Angeles.

Using the 2005 methodological approach as the foundation for a longitudinal analysis, the LMU research team has since implemented the racially stratified homogenous precinct approach in its 2008 Presidential Election and 2010 Gubernatorial Election Exit Polls. Research shows that the city of Los Angeles is racially segregated, thus voters are more likely to be voting in racially homogenous precincts. The key component of the stratified approach is the selection of racially and ethnically homogenous precincts throughout the city. This sampling method was not only validated, but it produced more accurate results than other exit polls in Los Angeles.

Although the stratified approach has been used in other exit polls, analysis is needed to determine the reliability of this methodology in subsequent elections. To complete this longitudinal analysis, the LMU research team will compare the official Statement of Vote for each election to the 2008 and 2010 LMU exit poll results. Specifically, the LMU research team will combine statistical analysis and visual comparisons of both precinct and citywide data to determine if the methodology is as accurate as it was in the 2005 inaugural poll.

It is the intent of the LMU research team that if this stratified approach maintains reliability it may be applied to other exit polls of diverse regions throughout the United States thereby avoiding the inaccuracies found in 2000 and 2004.

Social and Economic Characteristics of Mexican Migrants.

Paul Francisco Valdes Cervantes, Parámetro Investigación (paul.valdes@gmail.com); Jesús Irineo Carreño Rodríguez, Parámetro Investigación (leonardodejesus@gmail.com); Igor Vivero Avila, Universidad Autónoma del Estado de México (ivivero@yahoo.com)

The survey will be conducted in the state bus terminals ejectors migrants to the United States: Zacatecas, Michoacán, Tamaulipas, Coahuila, Puebla, Oaxaca, Veracruz and Mexico State in December of 2010.

800 face to face surveys point to Mexican immigrants flow returning from the United States of America for the holiday period of Christmas and New Year (December 2010).

The study provides information to understand the characteristics of the legal or illegal migrants residing in the United States of America.

The migrant population returning to their places of origin has emotional ties that make it usually goes back to where much of the income obtained as a migrant, and this gives a producer role within their community.

Get information to understand the determinants of legal and illegal migration, the sector in which they work, destination of expenditure in the place of residence, access to the financial system in the United States and Mexico, the frequency of sending money to Mexico. Other important features of the study is to understand the social situation in which encuesta while residing in the United States, you know who lives, if you have relatives in the United States, the number of relatives in the United States, the frequency and the motive of his visit to Mexico.

It also has indicators of whether a medical service in the United States, if you belong to any civil partnership in the United States, if at any time has been discriminated against or if someone has infringed their rights as human beings in the United States and Mexico.

The fact that conduct interviews in Mexico can help raise the level of response.

Technological Advances to Reduce Survey Error

Chris Ellis, RTI International (ellis@rti.org); Nathan Sikes, RTI International (sikes@rti.org); Adam Sage, RTI International (asage@rti.org); Joe Eyerman, RTI International (eyerman@rti.org); Brian Burke, RTI International (bjb@rti.org)

Reduction of survey error is an enduring and common concern among researchers. Recent procedural and technological advances in the industry have improved our ability to reduce survey error. Examples include providing GPS coordinates in address-based samples to reduce coverage error and computerized capture and transfer of household rosters and sample member demographics to reduce sampling error. However, some uses of new technology (e.g., the introduction of a separate handheld computer for household screening) may reduce error in one area while increasing it in another. e.g., the linking of multiple computer devices is sometimes problematic. Additionally, it is often challenging to determine which emerging technology will (or will not) assimilate effectively into a research organization's suite of services.

This poster will present initial test results and a framework for discussion across three primary subtopics:

- First, we will discuss how hardware innovation(s) can be harnessed to support the multi-faceted work of the field interviewer, allowing for the doorstep screening, household interview, retention and transmission of data, while ensuring affordable data collection to clients.
- Second, noting the increasing availability of global positioning system (GPS) technology, we will discuss how GPS can be used to increase the efficiency and monitor the efficiency and legitimacy of field interviewer work.
- Third, we will discuss what programming solutions can be brought to bear on interviewer coding as a source of survey error. In particular, can computerized analyses run against submitted data quickly and efficiently detects error in interviewer coding, transcription of open-ended responses, and/or overuse of open-ended response options?

Based on actual system testing and development, we will summarize our initial findings, the immediately available benefits of such technological advancements, and the next steps for development and/or implementation.

Understanding the Tea Party

Jocelyn Kiley, Pew Research Center for the People & the Press (jkiley@pewresearch.org); Alec Tyson, Pew Research Center for the People & the Press (atyson@pewresearch.org); Scott Keeter, Pew Research Center (skeeter@pewresearch.org)

Over the last two years, the Tea Party has emerged as an important force in U.S. politics. This paper investigates the correlates of active support of the Tea Party movement. We examine the demographic makeup and particular political attitudes of Tea Party supporters, as well as how they compare both to the public as a whole and to those who actively disagree with the Tea Party. The paper also utilizes the Pew Research Center's 2011 political typology (which uses cluster analysis to sort the public into homogeneous groups based on fundamental values and beliefs), as a lens through which to view Tea Party supporters.

The paper draws upon several Pew Research Center surveys conducted in 2010 and 2011 to describe the demographic characteristics of Tea Party supporters and how they compare to the general public. In addition to demographics, attitudes on politics and policies will be analyzed and compared with the public with a particular focus on the signature issues of the movement: health care, the federal budget deficit, and government spending.

Using the 2011 political typology, an ongoing Pew Research effort, the paper examines supporters of the movement beyond party identification and current policies; what common social and political values unite Tea Party supporters? What are the strongest predictors of support for the movement?

Which Technologies Do Respondents Use in Online Surveys – An International Comparison?

Lars Kaczmirek, GESIS (lars.kaczmirek@gesis.org); Dorothee Behr, GESIS - Leibniz Institute for the Social Sciences (lars.kaczmirek@gesis.org); Wolfgang Bandilla, GESIS - Leibniz Institute for the Social Sciences (wolfgang.bandilla@gesis.org)

Design decisions for Web surveys are restricted by the assumptions about the technologies respondents have available. Measurement problems might occur when fully labeled scales are displayed on small computer screens or when respondents participate via cell phones and other mobile devices such as Netbooks, iPhone, Ipad, or Blackberry. In these cases, the required equidistance of scale points could be violated. Other technologies whose availability are relevant in this context are Flash technology and the respondents' connection speed, that are key indicators for successful video presentations, and JavaScript which is widely used in automatic data validation procedures. JavaScript is also necessary for all interactive question types such as automatic tally questions or visual analog scales. In the process of designing a survey, the availability of these technologies is then highly relevant for the technical pretest. As pretesting is restricted to the most common combinations of technology, such as specific browsers, mobile devices, and connection speed, it is important to know which combinations really are the most common in the target group.

This study provides exactly this data on available technologies for countries with different Internet penetration rates, namely Canada, Denmark, Germany, Hungary, Spain, and the United States (N=480 per country, quotation on age, gender and education). Data was collected automatically, similarly to the collection of paradata, in January 2011 while respondents participated in an Internet survey. The participants were sampled from online access panels. The results provide information about the

availability of technology in different demographic groups: How do respondents access online surveys (connection speed, browser, mobile devices)? What technology can survey researchers safely design for (screen size and used window size, Flash, JavaScript)? The study shows that most surveys can use a wide range of design choices, but also that specific groups of respondents need a conservative approach.

Online Survey Software: A Comparison of the Features of Common Packages.

Alexandra Cooper, Social Science Research Institute, Duke University (cooper@duke.edu);

Patrick Miller, Duke Initiative on Survey Methodology (patrick.miller@duke.edu)

Online survey tools provide a cost-effective mechanism for creating and delivering surveys and collecting and accessing the resulting data; they are widely used. Any number of such packages exists, and it is not always easy for the researcher to know what interface will best serve his or her need. To address this issue, we have prepared an overview comparing the features of common packages. We examine several packages that are available for researchers to purchase (including Checkbox, Qualtrics, and Viewsflash, all of which are provided at no cost to faculty and students by our home institution, Duke University) and several that are available at no cost online (at least for simple, short surveys) including PollDaddy, SurveyMonkey, and Zoomerang. Our comparison examines a number of features, including (a) ease of entry – how simple is it to add features defining text, such as italics, and what features does the interface offer to import from other sources, such as a Word document; (b) functionality – how well does the system incorporate commonly used survey mechanisms, such as skip patterns, branching, piping, and randomization; (c) appearance – how easy is it to customize the look and feel of a survey instrument, from background and foreground color to table size and complexity and inclusion of graphics; (d) administration – what alternatives are available for sharing the survey instrument with research subjects, tracking responses, and sending reminders; (e) data storage – where does the package store responses, and how securely; and (f) data analysis – how versatile are the self-contained tools each package offers for examining results, and how easy does the package make it to download the data. While we focus on the needs of our primary constituents – faculty and students – our comparison is of broad relevance to survey researchers generally.

Age-Related Differences in Reported Computer and Internet Usage Based on Question Type: ‘A Great Deal’ of Variability.

Jennifer C Romano Bergstrom, U.S. Census Bureau (jennifer.romano@census.gov); Erica L Olmsted-Hawala, U.S. Census Bureau (erica.l.olmsted.hawala@census.gov); Wendy A Rogers, Georgia Institute of Technology (wendy@gatech.com); Jon A Krosnick, Stanford University (krosnick@stanford.edu)

What does it mean to have a great deal of experience using computers? The answer people provide depends on how they interpret the qualitative descriptor. Age-related difference in interpretation can substantively affect outcome measures. We developed an extensive questionnaire to measure participants' computer and Internet experience. The questionnaire used a Likert-type scale to measure comfort using a computer and the Internet to do various tasks. Respondents were also asked how many times they did specific things during the last week, such as use the Internet to find information for a hobby, for social networking, to play games, to read news, etc. In a sample of 100 respondents (older adults aged 65-75; middle-age adults aged 40-50; young adults aged 18-28), we found that although all groups reported having “a great deal” of experience using computers and the Internet on the Likert-type scale items, the actual reported experience for activities differed by age group. Older adults reported using the computer and Internet to complete fewer types of tasks than middle-age and young adults, and they reported using the Internet fewer times during the last week compared to middle-age and young adults. Thus, “a great deal” of experience meant different things to the different age groups. Surveys

designed to gather data on user experience with computers and the Internet must include questions about specific activities and amount of time spent doing such activities. In this paper, we show that merely asking participants to select a value (e.g., none, a little, a moderate amount, a lot, a great deal) for how much experience they have had using the Internet is not a valid measure. Moreover, the effects of different interpretations may lead to skewed assessments of age-related differences in usage patterns.

Nonresponse Bias in the Survey of Youth Perception of Science and Technology in Bogotá
Edgar Mauricio Bueno Castellanos, Colombian Observatory of Science and Technology
(embuenoc@gmail.com)

The Colombian Observatory of Science and Technology -OCyT- developed, in 2009, a survey about the perception of Science and Technology in students of the last two years of high school in Bogotá, Colombia. The survey sampling design was stratified according to the nature of school (official or private). During the data collection stage, two main sources of nonresponse were detected. The first one, as a consequence of the important difference in the response probability according to the nature of school: the survey was implemented in the 16 official schools included in the original sample (100%), while only 13 out of 31 private schools (42%) allowed to collect information. The second source corresponds to students who belong to schools in which access was allowed, but did not assist during the days when survey was applied. Estimates, initially, were obtained modifying the original sample sizes by those observed. Subsequently, it was decided to obtain new estimates taking into account the nonresponse effect; to achieve this goal, the values corresponding to item nonresponse were imputed using the methodology of the nearest neighbor and the calibration method was used for unit nonresponse. The results obtained for both cases don't show visible differences, especially when estimating a ratio; even though, some great differences were observed when estimating totals. Although it is impossible to determine with certainty which methodology was more precise, results obtained using the combination imputation-calibration are more reliable because it considers auxiliary information, which is not present in the first estimation methodology.

Pursuing an Education While Facing Serious Physical or Cognitive Difficulties: Examining Educational Attainment for Young Adults with Self-Reported Disabilities
Angelina N. KewalRamani, American Institutes for Research (akewalramani@air.org)

In 2009, approximately 19.5 million people between the ages of 16 and 64 reported having a physical or cognitive disability. The traditional systems of power that individuals with disabilities often encounter in school and society can severely limit their opportunities for achievement. In addition, students with physical or cognitive disabilities may experience taunting or bullying in school, as well as other adversities that interfere with learning. Youth with disabilities who overcome these obstacles and complete high school are more likely than their dropout counterparts to participate in positive behaviors, such as pursuing a postsecondary degree. However, as the number of young adults with disabilities continues to grow, tracking measures of educational progress such as completing high school remains a formidable challenge. Historically, the ability to directly compare the high school completion rates of students with disabilities with the rates of students without disabilities has been limited by the different data sources and definitions used by the Office of Special Education Programs (OSEP) and the National Center for Education Statistics (NCES). Data sources traditionally used for estimating high school completion lacked valid questions about disabilities that would have allowed for meaningful comparisons. However, in 2008 the American Community Survey (ACS) revised its disability items based on negative feedback received from data users. The revised questions capture respondent difficulty in vision, hearing, cognitive, ambulatory, self-care, and independent living. This paper will evaluate the ACS 2009 disability items with a focus on the educational attainment of young adults with disabilities. More specifically, the research will

compare overall high school completion rates of 18- to 24-year-olds with and without disabilities. Analyses will be conducted by disability type, gender, and race/ethnicity. The research will reveal whether the ACS disability items provide reliable estimates of the educational attainment of young adults with self-reported disabilities.

The Depths of American Public Opinion on Climate Change: Harmonizing the Results of Diverse Surveys.

Jon Krosnick, Stanford University (krosnick@stanford.edu); Bo MacInnis, Stanford University (bo@macinnis.org)

Many survey organizations have been tracking public opinion on climate change. Different surveys employ different questions, which tap into public perceptions of related but different constructs that are centered on global warming, and consequently, the results differ across diverse surveys. However, the cross-survey difference in public opinion on climate change has contributed to disagreement among scholars and pundits and thus propelled recent disputes on whether Americans' acknowledgment of global warming has declined. In this study, we use global warming measures in our surveys to harmonize several global warming measures that are tracked by other survey organizations, namely, Pew Research Center, Gallup Organization, CNN/ORC, and NBC News/Wall Street Journal. To enable such cross-surveys harmonization, in two of our most recent national surveys conducted in June, 2010 and November, 2010, we added the questions by the aforementioned organizations, with the exactly same question wording and presentations of response choices. Our analysis indicates that results from these diverse surveys share the common depths of public opinion on global warming and can be successfully harmonized using the judgments built from our surveys. The harmonization exercise allows us to bridge the gaps among multiple public opinion measures on global warming, to advance our collective understanding how Americans perceive global warming, and to bring harmony among survey organizations to the discord in this area of public opinion.

Sampling Transgender Communities to Assess Healthcare Access and Develop Public Health Programs.

Judith Bradford, The Fenway Institute (jbradford@fenwayhealth.org)

Transgender is an umbrella term used to describe persons whose gender identity or expression does not conform to the sex they were born as or assigned to at birth. To address concerns about the prevalence of HIV cases among this population and healthcare system readiness to provide culturally competent care, the Virginia Department of Health, its HIV Prevention Community Planning Committee, and a university-based research team conducted a statewide study to assess HIV-related behavioral risks, determine the health-related needs of transgender Virginians and evaluate their access to transgender-sensitive healthcare. A Statewide Transgender Task Force was formed to ensure that transgender persons provided guidance and were engaged throughout all phases of the study. We conducted focus groups with both gender vectors - male-to-female (MtF) and female-to-male (FtM) and with racial/ethnic diversity, and used results to refine a quantitative survey based on previous reports. To offset Task Force concerns about the willingness of population subgroups to participate, we provided both internet and paper-pencil versions of the questionnaire. Of 367 respondents, 61% responded online and 39% completed paper forms. Characteristics of the subsamples differed by geographic region, race/ethnicity, type of community (urban, suburban, rural), living alone or with others, home ownership, health insurance, involvement with other transgender people, and disclosure of gender nonconformity to primary care provider. Discrimination in healthcare, education and housing was reported by 41% percent of the sample. We used multivariable logistic regression to examine the prevalence and types of transgender-related discrimination. Variables independently associated with increased odds of discrimination were:

being unemployed (aOR=1.83; 95% CI=1.03-3.24; p=0.04); being “out” to a regular primary care provider (aOR=2.80; 95% CI=1.39-5.64; p=0.004); and being more connected to the transgender community (i.e., having contact with a greater number of other transgender people in the past 6 months; aOR=1.23; 95% CI=1.03-1.47; p=0.03).

Sample Re-use: An Application in Canadian Media Ratings.

Kimberley Lafleur, BBM Canada (klafleur@bbm.ca)

Sample re-use is defined as using the same respondents for two or more surveys. Benefits of sample re-use include: ability to meet difficult sample targets, reduced enumeration costs, more stable data to accurately measure trends, increased return rates, and the potential for more detailed analyses of audience flow.

The drawback to sample re-use is the possibility of introducing respondent bias. Respondents who are willing to participate in more than one survey may be more engaged and therefore may have different behavior than the general population.

The Sample re-use test was undertaken for BBM's radio diary service. The radio diary service consists of two surveys per broadcast year, Fall and Spring. Since not all markets are included in both surveys, the interval between surveys for those respondents participating in sample re-use may be up to one year.

Two different approaches were used for this test:

1. Re-enumeration: All households that returned at least one diary in Spring 2008 were re-enumerated for Spring 2009.
2. Direct Mail: Diaries were mailed to all households that returned at least one diary in Spring 2008 based on the Spring 2008 enumeration files.

Overall tuning patterns for the test sample were similar to production for both the re-enumeration approach and the direct mail approach. Some differences in tuning were found for specific age groups. Return rates for the test groups were higher than production for both approaches. Cost per return was reduced for both approaches. The qualitative profile of test respondents differed from production for household size, hours worked and income.

A simulation was run incorporating sample re-use with the production dataset at 3 different levels (10%, 20% and 30%). Although there were some differences found between the sample re-use data and production data, these differences did not impact overall tuning levels at the 20% sample re-use level.

The Social Aspect of the Digital Divide

Edward Paul Johnson, Opinionology, formally Western Wats (ejohnson@opinionology.com)

When most people focus on the digital divide they think of those without the technology requisite to becoming part of the sampling frame. However, there is another dimension to the digital divide that is not commonly discussed. How many people have internet and a phone, but will not participate in an internet survey while they will participate in the phone survey? This non-response error could be just as devastating as coverage error to a well run research design (Groves, 1989). In the years that Western Wats and Opinion Outpost has conducted phone recruit to web panels and phone recruit to web surveys, we have encountered lower response rates than to regular RDD phone surveys. We hypothesize that there is a significant aspect to the digital divide that is social rather than technological in nature.

To test our hypothesis we conducted a RDD phone survey of the individuals in the United States with both internet access and a phone. These respondents were called and asked to complete a web survey. Those that agreed to participate were sent an email invitation to take a 10 minute survey about attitudes towards online surveys, shopping, and social media sites. Respondents who did not agree take the web version were given a shortened 5 minute. Those that agree to take the web version and did not take after several reminders were contacted again on the phone and given the 5 minute version of the survey. The sample contained 9,879 valid residential numbers with a contact rate of 38% and a cooperation rate of 15% which resulted in 581 interviews. Data differences in attitudes and behaviors are highlighted in this presentation.

Predicting Mobility in Special Populations: Lessons from the Making Connections Survey

Kate E Bachtell, NORC at the University of Chicago (bachtell-kate@norc.org); Michael Latterner, NORC at the University of Chicago (latterner-michael@norc.org)

In 2010, Bachtell et al. reported on the challenges of collecting retrospective addresses. While this research focused on the efforts to attain higher quality address data from respondents, a precursor to such efforts and the subject of our current research is the difficulty of tracking and eventually contacting mobile respondents. Accurate predictions of mobility between collection periods are necessary in planning longitudinal surveys, as mobility influences survey costs.

Research involving diverse and special populations brings challenges for producing evidence-based estimates of mobility. Data derived from five sites studied as part of the Making Connections Survey, an evaluation of low-income communities in Ten U.S. Cities, reveal differences in residential mobility between sites. In this paper we first explore the geographic nature of these differences and examine factors which may influence them, including household and neighborhood characteristics. The second portion of our paper presents five Making Connections sites as case studies of mobility among special populations. We will compare data from the Making Connections survey to data collected from populations with similar geographic and demographic characteristics through the Current Population Survey to explain differences in mobility across sites, and to consider the varied “lessons” offered by the two sources. We will test several logistic regression models that have been shown to predict residential mobility, highlight intervening factors, and assess the fit of these models to the Making Connections data as a means to inform future methods of predicting mobility.

Do Global and Detailed Questions Achieve Equivalent Results? A Case Study from the Consumer Expenditure Survey

Jeanette Duddy Davis, Bureau of Labor Statistics (Davis_Jeanette@bls.gov)

The Consumer Expenditure Survey (CE) collects data on household expenditures over the course of a year. The interview is long and can be burdensome. Asking about some expenditure categories as broad global questions has been proposed as a way to shorten the interview. This paper describes the results of a random-digit dialing field experiment designed to determine whether expenditures reported in response to global questions would be comparable to those reported through CE’s current set of detailed items. Respondents were randomly split into two panels and asked about the same topic (clothing). One panel was asked about expenditures with three global questions while the other was asked about expenditures with eight global questions. Both panels were then asked the same set of detailed CE expenditure questions. Overall the three-question global approach yielded results that were comparable to the current set of detailed questions, while responses to the eight-question series were consistently and significantly higher than the CE detailed items. The paper describes differences in the global questions, presents the results of the investigation by sub-topic, and shows that the time savings is relatively small.

The Search for a Definition of Socioeconomic Status for a Survey of High School Students
Tommy Holder, RTI International (tholder@rti.org)

The High School Longitudinal Study of 2009 (HSL:09), conducted for the National Center for Education Statistics, tracks more than 21,000 U.S. 9th graders through their high school years and beyond. The main focus of HSL:09 is to understand students' academic goal-setting and decision-making regarding postsecondary educational destinations and career choices. The survey emphasizes especially the STEM areas (science, technology, engineering and mathematics).

Research has shown socioeconomic status (SES) to be an important correlate of achievement and other educational outcomes. The very definition of SES has the potential to significantly affect the interpretation of data. Although student SES is typically calculated with three components (family income, parents' occupation(s) and parents' education level), alternative constituents are sometimes used, and the SES construct can be made sensitive to other potentially relevant factors such as family size and geographic location. (For example, it may be asked whether a family of seven should have the same SES value as a family of three holding family income, parental occupation and education constant).

Given the importance of SES to researchers, several definitions of the construct were gathered from the literature for a comparative evaluation. Included in the examination are variants of the three-component SES definition cited above. This paper details the correlational analyses conducted to evaluate the candidate definitions and the SES formula adopted for HSL:09. Our research suggests that controlling for additional correlates produces a more effective SES measure in comparison to an overall definition used by many surveys and one that better meets the criteria of construct validity.

The Political Implications of a Drive to Digitize: An Examination of Audience Engagement with Newspapers.

Ann E. Williams, Georgia State University (annwilliams@gsu.edu)

With the emergence and rise of digital technologies, changes in the economy of the press are in turn changing the shape and nature of public communication, civic engagement, and political involvement. This paper employs public opinion data from a randomly selected, representative national sample (N ~ 2200) to approach questions about how newspapers' move from print to digital form has shaped individuals' interaction with news texts in ways that ultimately influence the participatory momentum of the electorate. Drawing from multiple theories of public opinion, the paper offers a conceptual model of audience engagement that operates from the perspective of individual members of the public rather than from the position of the industry. The analyses suggest that the political value newspapers have long brought to an informed and participatory electorate are not uniform across offline and online modes of news reading. While online news reading provides opportunities for information exchange and networking that can work to enhance the political value of news texts, traditional reading patterns that correlate directly with certain forms of participatory outcomes are not as strong for many who receive news primarily from online news counterparts. Unique measures that assess breadth and depth of audience engagement also show that individual user preferences and practices contribute significantly to the political impact of online news reading. The social and normative implications of the findings are discussed. In a larger social context, the outcomes of this study suggest that the economic value of newspapers' 'drive to digitize' also brings cost to the public. Continued study in this area is needed as the political value that the press carries to news audiences is increasingly mediated.

Thursday, May 12, 2011

3:00 p.m. – 4:00 p.m.

Demonstration Session 1

Demonstration 1A.

(IQ)2: A Cutting-Edge, Online Hybrid Solution for Qualitative/Quantitative Insights From Your Target Audience Located Anywhere in the World.

Michael Remolona, Blackstone Group (michael.remolona@bgglobal.com); Mike Burmester, Blackstone Group (mike.burmester@bgglobal.com)

Blackstone Group has developed an Internet-based data collection and interviewing methodology, called (IQ)2, that is ideal for gathering opinions and exploring insights from hard-to-reach, diverse and geographically disparate special populations.

The methodology utilizes a wide range of interactive tools to capture individual insights that are at turns deep and qualitative and unbiased and quantitative, as well the unique insights that emerge from group discussion. This careful integration of quantitative, qualitative, individual and group exercises overcomes common research drawbacks such as groupthink, normative social influence and backroom or observer bias. Because participants complete many exercises and activities concurrently, researchers are able to gather large amounts of information in a small amount of time.

The tools utilized in the (IQ)2 methodology include: an online meeting platform supporting phone and VoIP audio/verbal communication as well as web-camera technology; a robust, web-based quantitative survey software package; a text-based chat room that allows both public and private communication (especially good for sensitive topics); a screen-sharing module; a whiteboard that can be overlaid on a wide range of visual and audio stimuli; and a separate, hidden channel through which observers and facilitators can communicate. This versatile set of tools can be adapted to study any number of topics and is compatible with the majority of computer operating systems and only requires a high-speed Internet connection for participation. The technical capabilities of participants are determined in advance as part of the screening, sampling and recruiting process.

In addition to being an online option for executing in-person focus groups, (IQ)2's flexible platform is an effective alternative for conducting other types of qualitative research, including in-depth interviews, dyads and triads.

(IQ)2 has been tested extensively by Blackstone Group during the past eight months in various contexts and has proven to be a dynamic and valuable methodology for public policy research.

Demonstration 1B.

Completing Web Surveys on Cell-enabled iPads.

James Dayton, ICF International (jdayton@icfi.com); Heather Driscoll, ICF International (hdriscoll@icfi.com)

Intercept field data collectors working in outdoor environments using electronic devices face a number of challenges, such as conducting interviews in bright sunlight, wind, and/or rain. Some in-field data collection even requires exposure to dirt, grime, and natural specimens. These interviewers need data collection devices that are lightweight, can be held in one hand or suspended from the neck, are protected from the elements, have outstanding battery lives, are easy to read in various lighting levels, and have large screens able to display an easy-to-read font and graphics.

Researchers at ICF sought a data collection option that was both innovative and well-suited for conducting interviews in challenging environments without many of the cost and data security implications that can accompany more traditional CAPI implementations. Our solution was web-based surveys on cell-enabled tablets (specifically 3G iPads).

We will demonstrate a cell phone-enabled iPad set up to intercept anglers working from docks and beaches and on boats. The demonstration will review how visual aids will assist anglers and interviewers with the identification of various fish species, the use of the upcoming camera feature to capture a visual record of species that can later be confirmed by ICF fish biologists, and how the GPS function can be used to help assure interview quality and track members of a large field force. We will also demonstrate how networked scheduling applications can be used to direct interviewers to specific locations at specific times, as well as track progress toward a variety of specific interview quotas. Finally, we will demonstrate if the iPad can be configured to mimic ease of use and other functional advantages of the current clipboard paper and pencil system through the use of flexible comment fields and the ability to toggle across multiple ongoing surveys.

Demonstration 1C.

Facilitating Deliberative Consultation: A system for Automated Deliberative Polling Analysis.
Sean Westwood, Stanford University (seanjw@stanford.edu)

While issues of sample size, sample selection and survey design dominate the methodological considerations for traditional public opinion research, recent movements toward deliberative opinion refinement (AmericaSpeaks, Deliberative Polling and Citizen Juries) present new methodological and logistical challenges. This paper addresses the methodological and analytical complexity for Deliberative Polls and offers a free and open source tool automate much of the analysis needed for Deliberative Polling. Deliberative Polls combine a traditional opinion poll conducted through various means (face-to-face, telephone or online) with a follow-up deliberative event. Standard tabulations of the initial poll begin the process, but to claim generalizability of the small deliberative sub-sample drawn from the initial poll's sample, it is necessary to compare the deliberative sample to the initial sample/census data. Labeled as 'representativeness' this analysis compares both the deliberative and non-deliberative groups along socio-demographic measures and across all attitudinal measures. At the conclusion of the deliberative event refined opinion is captured and compared at the individual-level to the opinions gathered in the initial poll through paired t-tests for each question. Further, measures of satisfaction and knowledge gain are also generated. Each Deliberative Poll varies in language(s), number of questions, number of time points and in various other ways. Considering the wide variance in poll structure and the need to produce consistently formatted data for both lay and educated audiences I created a software tool in R and PHP that fully automates the process of generating all statistical analysis and frequencies for Deliberative Polls (representativeness, opinion change, knowledge gain and satisfaction). The software integrates with OCR scanning software and produces HTML and Word formatted output with plots for each question. The

software is free and provides results in minutes that would otherwise take several individuals days of manual labor in STATA or SPSS.

Demonstration 1D.

A Stata Program for Respondent Driven Sampling
Matthias Schonlau, RAND Corp (matt@rand.org)

Respondent-driven sampling (RDS) is a sampling technique typically employed for hard-to-reach populations (for example, homeless people, people with AIDS). Briefly, initial seed respondents recruit additional respondents from their network of friends. The recruiting process repeats iteratively, thereby forming long referral chains. It is crucial to obtain estimates of respondents' network sizes (for example, the number of friends with the characteristic of interest). RDS shares some similarities with snowball sampling, but the theoretical foundation for inference using RDS samples is much stronger. I will give a brief overview of this technique and demonstrate my implementation of RDS in Stata.

Thursday, May 12, 2011

4:00 p.m. - 5:30 p.m.

Concurrent Session B

A Generation of Societal Change

Measuring Change with Changing Measures
Peter V Marsden, Harvard University (pvm@wjh.harvard.edu)

Repeated surveys such as the General Social Survey (GSS) track social trends over extended periods of time. They stress the exact literal replication of question wordings over time, holding the measurement instrument constant in order to discern change. Questions in repeated surveys do change occasionally, however, as new issues emerge, previous issues become more salient, or near-consensus emerges on existing items.

Such changes in items pose challenges to analysts who seek to assess change in a domain over periods in it is measured using different items. Over-time comparisons of simple additive scales based on those items administered at the different points in time may give misleading indications of trends, especially when questions having skewed marginal distributions are replaced by new items on which respondents are more evenly divided.

This paper uses item response theory methods to assess over-time change in a latent trait measured using different indicators over periods of time, for the case in which the indicators of the trait measured at one period overlap partially with those available for other periods. Illustrations of the approach are drawn from the GSS, for trends in nontraditional gender-role orientations and in verbal ability.

Trends in ANES Survey Data: 1988-2008. Vincent L. Hutchings, University of Michigan (vincenth@umich.edu); Simon Jackman, Stanford University (jackman@stanford.edu); Gary Segura, Stanford University (segura@stanford.edu)

This paper provides trend data on American National Election Study (ANES) questions from 1988 through 2008. Particular emphasis will be paid to questions about salient social groups, ideological orientations, and economic policy items. Implications will be drawn for the upcoming 2012 ANES survey instrument.

The Millennials Come of Age.

Rich Morin, Pew Research Center (rmorin@pewresearch.org)

The Millennial Generation encompasses roughly 50 million men and women aged 18 to 29. As a generation, Millennials are more ethnically and racially diverse than older adults. They're less religious and less likely to have served in the military but more likely to have a tattoo. They also are on track to become the most educated generation in American history. The Great Recession slowed their entry into careers and first jobs, but they are more upbeat than older adults about their own economic futures as well as about the overall state of the nation. They are less skeptical than their elders of government and believe government should do more to solve problems. Politically they are the most likely of any generation to identify themselves as liberals and Democrats. Millennials also are the first "always connected" generation: More than eight-in-ten say they sleep with a cell phone by their bed and almost two-thirds acknowledge they sometimes text while they drive. This profile of the Millennial Generation emerges from a Pew Research Center telephone survey conducted Jan. 14 to 27, 2010 on landline and cell phones with a nationally representative sample of 2,020 adults, including an oversample of 830 respondents 18 to 29.

Stability and Change in U.S. Society Since 1980: A Review of Gallup Trends.

Jeffrey M. Jones, Gallup (jeff_jones@gallup.com); Lydia Saad, Gallup (lydia_saad@gallup.com)

The paper examines changes in Americans' attitudes and behaviors over the last 30 years of Gallup polling on key trends on moral issues, social policy and religion, both overall and by age group.

Cross-Cultural Survey Methods: I.

Influence of Health Care Attributes and Language Preference on Disparities in Willingness to Engage in Cancer Prevention among Hispanic and White Americans.

Jocelyn Landau, University of Pennsylvania (Jocelyn.Landau@uphs.upenn.edu); Katrina Armstrong, University of Pennsylvania (karmstro@mail.med.upenn.edu)

Racial and ethnic disparities have been documented across all domains of health care, but the etiologies of these disparities remain poorly understood (Jacobs et al., 2005). Language preference and proficiency have been associated with health-related behaviors, disease prevalence, and receipt of health care services among Hispanics, but lack of sufficient individual-level population-based data has limited our ability to understand the extent of language-associated disparities in health (DuBard & Gizlice, 2008). National surveys offer an important way to examine health disparities and language barriers across racial and ethnic groups.

This study employs conjoint analysis to assess the effects of health care attributes and language preference on willingness to engage in cancer prevention measures among Hispanics and Whites. Two

conjoint modules were designed, each containing three attributes that were selected from health policy discourse during the time of data collection. Data are from a random digit dial (RDD) telephone survey of a nationally representative sample of American adults (N=982). All participants were given the option of completing the survey in English or Spanish. Hispanics were oversampled to allow for more robust comparisons across ethnic groups.

The findings reveal that varying health care attributes influence willingness to engage in prevention measures to reduce cancer risk. Differences across ethnic groups were observed, but language preference proved to be a stronger predictor of these differences. Language preference directly affected overall willingness to take a chemoprevention drug and moderated the effect of different health attributes on willingness to take a predictive genetic test or chemoprevention drug, even after controlling for race/ethnicity, income, education, gender, health insurance status, and health care system distrust. Results suggest that the attributes associated with health care delivery and treatment have differential effects on medical uptake across ethnic groups, and language preference is an important consideration in these effects.

A Nation of Immigrants: Innumeracy and Distortion Surrounding Estimation of Hispanic/Latino Group Size.

Jonathan D. Stringfield, The University of Illinois at Chicago (jstrin1@uic.edu); Ryan A. Sporer, The University of Illinois at Chicago (rspore2@uic.edu)

Studies on survey respondent's estimation of minority group sizes have been fairly consistent in the finding that respondents tend to overestimate the size of minority groups. The studies of this phenomenon, often labeled as "innumeracy" for cognitive response functions or "distortion" for mechanisms relating to threat or prejudice, have not explicitly focused upon Hispanic/Latino populations in a U.S. context. Given the dramatic growth of Hispanic/Latino populations in the U.S. in recent decades, coupled with media and political attention focused upon immigration reform, respondent estimation of the Hispanic/Latino population seems to be both a superlative and as of yet largely understudied mechanism for understanding U.S. distortion/innumeracy. Further, this phenomenon is uniquely informative to the perceptual impetus behind such measures as Senate Bill 1070 in Arizona (the "Arizona Immigration Law").

The recently completed 2010 Chicago Area Study (CAS) combines respondent-level estimations of Hispanic/Latino group size, in addition to numerous attitudinal and cognitive measures framed in reference to Hispanic/Latino populations on diverse topics such as "threat," prejudice, political participation, educational/knowledge measures, and political media exposure. Thus, results from the 2010 CAS provide an opportunity to at least partially disentangle the effects of survey-response cognitive measures (innumeracy) and attitudinal measures (distortion) related to minority group estimation heretofore only partially disassembled in extant literature. While the local-nature of this study limits national generalizability, focusing upon immigration issues in local communities provides an opportunity to isolate issues of innumeracy and distortion in areas where immigration issues are particularly salient to the everyday lives of respondents. Results from this paper will add to the collective knowledge of how innumeracy/distortion operates in minority group estimation in a U.S. context, and inform general understandings of how attitudes relating to immigration may inflate perceptions of Hispanic/Latino group size (which has become increasingly manifest in political rhetoric and anti-immigration legislation).

Cross-Cultural Measurement of Self-Rated Health.

Sunghee Lee, Univ. of Michigan (sungheel@isr.umich.edu)

With the recent growth of immigrant populations, U.S. has seen a sharp increase in ethnic and linguistic minorities. The largest immigrant group is Hispanics, and more than 60 percent of the linguistic minorities is Spanish speakers, counting at 34.5 million. Given such trends, there has been an effort to include linguistic minorities, who were traditionally considered as part of nonrespondents, in the population-based data collection especially by the federal government. It has become a standard practice for numerous government-administered and sponsored surveys that were formerly administered only in English to add Spanish an interview language. While this budding effort of including these populations is likely to decrease nonresponse bias, this practice introduces a room for errors stemming from inequivalent measurement properties of survey instruments.

This paper focuses on the measurement properties of self-rated health (SRH), one of the most widely used survey items around the world. Data for the study come from three sources, the California Health Interview Surveys (CHIS), the English Longitudinal Study of Ageing (ELSA), the Survey of Health and Ageing and Retirement in Europe (SHARE). These surveys all include SRH. CHIS and SHARE are conducted in multiple languages. When combined together, these three provide a multi-lingual and multi-cultural survey environment. CHIS 2007 and 2009 as well as the first waves of ELSA and SHARE include an experiment that randomized the order of the SRH item. We will first examine whether the population-level age-gender adjusted health status is comparable across survey languages and countries. We then will analyze the order experiment data to examine whether there is an order effect and, if so, how the order effect behaves with respect to the interview language.

Methodological Challenges in Implementing Comparative Ethnographic Evaluation Research across Race/Ethnic Populations in the 2010 Census: An Asian Case Study.
Laurie Schwede, Census Bureau (laurel.k.schwede@census.gov)

The Census Bureau is conducting the 2010 Census evaluation, "Comparative Ethnographic Studies of Enumeration Methods and Coverage in Race/Ethnic Groups." The evaluation aims to identify reasons why miscounts of some minority subpopulations persist across decennial censuses and to suggest improvements. Expert ethnographers conducted a coordinated set of small-scale systematic Census field observation/debriefing studies of live Census Coverage Measurement (CCM) survey interviews to address the same research questions with the same methods at the same time in 9 U.S. race/ethnic sites. In most sites, the design called for researchers to try to observe most interviews in households of a particular race/ethnic group.

This is the first paper on this evaluation. This paper presents a case study of one site to identify methodological and substantive issues encountered while attempting to apply that overall design of maximizing observed interviews with one small, targeted hard-to-reach subpopulation: dispersed Asian, primarily Chinese, households in the CCM sample in a California metropolitan area. Data come from observations/debriefings in 2010 of 40 live CCM interviews, 17 of them with Asians and the rest with other race/ethnic households. This paper identifies and discusses factors affecting the enumeration of Asian and other households in this site and how they contribute to the overall evaluation. Factors for Asians include: finding and engaging enough Asian households, language and cultural barriers, mistrust of outsiders, lack of community liaisons and bilingual interviewers, problems with a language identification card, question wording, and on-the-fly written and oral translations. Creative solutions field personnel developed to overcome issues are identified. Factors affecting enumeration of other race/ethnic households include mobility, mistrust of outsiders, and question wording, particularly in using standardized race and Hispanic origin questions to classify race/ethnic identity. Implications of these

results for the overall evaluation and for planning and implementing other comparative subpopulation studies will be identified.

Enhancements to Address Based Sampling.

A Data Collection Model for Address Based Samples.

Jill M. Montaquila, Westat (jillmontaquila@westat.com); J. Michael Brick, Westat (mikebrick@westat.com)

The availability of address based sampling (ABS) frames based on the U.S. Postal Service Computerized Delivery Sequence (CDS) files has brought on a revolution in sampling and data collection for both general population and subpopulation surveys. These sampling frames are able to provide very good national coverage of households in the U.S. With the appending of telephone numbers to the addresses, ABS frames offer survey organizations a choice of modes for data collection.

In this paper, we briefly discuss considerations when using the ABS frames that are associated with each of the modes—in-person, mail, and telephone. We describe a mail-based data collection model that has been used for both general population and subpopulation surveys. The paper includes discussion of results from two surveys—one general population survey and one subpopulation survey. Both surveys were conducted in two phases, beginning with a mail household screener followed by an extended survey administered by mail or phone. The mail screener was used to collect information about household members that was necessary for determining eligibility and sampling. Upon receipt of the completed screener, we sampled one eligible household member for an extended survey. In one of the studies, the extended was administered by telephone; in the other study, mail was the primary data collection mode for the extended survey.

This paper pulls together our experiences and results from these two data collection efforts. We present the mail screener response rates, discuss the ability to obtain telephone numbers from the household through the screener, and present both the mail-based and telephone-based extended survey response rates. The paper also includes discussion of the results of methodological experiments that were embedded in these two studies and the implications of these results for a general ABS data collection model.

Creating an Improvement Universe for an Address-Based Sampling Frame.

Timothy Kennel, U.S. Census Bureau (tkennel@survey.umd.edu)

It is well documented that the coverage of many Address Based Sampling Frames is poorer in rural areas than urban areas. Our evaluation of the U.S. Census Bureau's address based sampling frame confirms many of the findings in the literature. Area frame listing is one of the most popular coverage improvement methods. However, due to the high cost of listing, it is most efficient to list only segments with the most coverage errors. To improve the coverage of the U.S. Census Bureau's Master Address File in rural areas, we are planning to perform a listing operation in several states with the highest potential for coverage bias. Within those states, we plan to define a coverage improvement universe of blocks where the frame is most in need of updating. We describe the process of determining which states have the highest potential for coverage bias and the process of defining which blocks in those states to put into an improvement universe.

Supplementing Address-Based Sampling Frames with Physical Addresses of Housing Units with Unlocatable Mailing Addresses.

Bonnie E. Shook-Sa, RTI International (bshooksa@rti.org); Douglas B Currivan, RTI International (dcurrivan@rti.org)

When utilizing an address-based sampling (ABS) frame for an in-person household survey, households with unlocatable mailing addresses are a primary source of undercoverage. Unlocatable mailing addresses such as post office boxes, rural route boxes, highway contract boxes, and simplified addresses cannot be linked to housing units on the ground. For this reason, approximately two million residential unlocatable mailing addresses are typically excluded from the national ABS frame for in-person surveys. Most of the undercoverage occurs in rural areas where unlocatable addresses are more prevalent. Low ABS coverage in rural areas often leads researchers to resort to a hybrid frame that supplements ABS coverage with costly field enumeration (FE) methods in areas without adequate ABS coverage.

Databases of listed addresses derived primarily from white pages contain physical addresses for a significant number of households with unlocatable mailing addresses. We explore the feasibility of white page supplementation using data from the 2009 National Survey on Drug Use and Health (NSDUH). The NSDUH provides national, state, and substate data on substance use and mental health in the civilian, non-institutionalized population age 12 and older. Data are collected on a quarterly basis each year, with approximately 140,000 household screenings and 67,500 interviews completed annually.

We estimate the gains in coverage provided by white page supplementation and evaluate the feasibility of combining white page lists with the locatable mailing addresses on the ABS frame. We also investigate the decreased reliance on FE that results from white page supplementation in a hybrid frame approach. We estimate that white page supplementation could add as much as 7.6 percent coverage to the household population in rural segments. This would result in cost savings by reducing the number of segments that need to be field enumerated.

The Impacts of Frame Construction Methods on Survey Estimates: Traditional Listing vs. The USPS DSF.

Katie Dekker, NORC at the University of Chicago (dekker-katie@norc.org); Ned English, NORC at the University of Chicago (english-ned@norc.org); Colm O'Muircheartaigh, NORC at the University of Chicago (colm@norc.uchicago.edu)

This paper closely examines the types of addresses that tend to be included or excluded from address frames. Address databases, derived from the USPS delivery sequence file (or DSF), can serve as a frame for face-to-face or multi-mode address-based samples. The DSF has been demonstrated to have excellent coverage properties in urban and suburban areas, but may be less useful for sampling frame generation in rural areas with non city-style delivery. At question is when is it appropriate to employ the DSF vs. create traditional listings; there is ambiguity surrounding which environments may be best suited to each. Survey research organizations would generally prefer to employ geocoded addresses from the USPS DSF whenever possible, due to cost and efficiency advantages. The use of the DSF would not be indicated if it would potentially result in the undercoverage of target households, however. We work with the Residential Energy Consumption Survey (RECS), which made a-priori decisions at the segment level as to whether DSF coverage was likely sufficient or whether it were necessary to conduct traditional listing. Following data collection, we conducted a new listing in segments that were near the threshold between the DSF vs. traditional listing, in order to quantify any potential undercoverage. Additionally, we closely examined the addresses that were added in this post data collection effort. We check to see if any addresses added during listing were present on the DSF but may have been mis-geocoded, or were missing from the DSF altogether. We also determine if added housing units tended to be in multi-unit buildings, near segment boundaries, or had other observable characteristics.

Enhancing Address Based Sampling with Contingent Incentives and Sample Indicators.
Lukasz Chmura, The Nielsen Company (Lukasz.Chmura@nielsen.com); Tracie Yancey, The Nielsen Company (Tracie.Yancey@nielsen.com)

In an effort to enhance address based sampling, The Nielsen Company conducted a test meant to explore the following: (1) the use of large contingent (“pay for performance”) incentives in combination with non-contingent (“pay up front”) incentives (2) geo-coding to identify black race households at the sample selection phase and (3) use of vendor-supplied sample indicators.

By moving to the use of larger, contingent incentives (combined with an initial modest non-contingent incentive) the focus is on rewarding households that perform the requested task. This shift allowed us to offer responding homes a larger incentive than under the current incentive plan, which uses non-contingent incentives only. The use of larger contingent incentives was hypothesized to have a more positive impact on response rates as compared to the current incentive methodology, and provides many cost efficiencies.

The sample indicators, either provided by the vendor or added using geo-coding, were used at the pre-recruitment and mailing stage. Households that confirmed they were in the hard to reach demos and those believed to be in the hard to reach demos were given higher incentives. Sample indicators were hypothesized to have a positive impact on both response rates and sample quality.

The test showed that the use of contingent incentives resulted in significantly lower response rates. The impact on demographic distributions was mixed. However, the use of geo-coding to identify black race households at the sample selection phase had a positive impact on both response rates and sample quality.

Factors Influencing Survey Participation.

Response Rates and Interviewer Personality.

Jacob Sohlberg, Stony Brook University (jacobsohlberg@gmail.com); Leonie Huddy, Stony Brook University (lhuddy@notes.cc.sunysb.edu); Kathleen Donovan, Stony Brook University (katiedonovan12@gmail.com); David C Martin, Stony Brook University (dmartin66@gmail.com)

Interviewers are a crucial ingredient in the survey research process. In an increasingly difficult environment dominated by declining response rates and increasingly expensive interviewer-mediated surveys, it is important to understand the factors that boost response rates even slightly. We thus examine the degree to which interviewer personality characteristics influence variations in survey cooperation and response rates. Past research has examined the effects of personality on job performance and it finds that those who are diligent and outgoing tend to be more effective (Barrick and Mount 1991; Hurtz and Donovan 2000). But to our knowledge few studies have investigated the effects of interviewer personality on telephone survey response rates. We draw on a burgeoning literature in psychology and political science on the big five factors of personality (e.g., McCrae and Costa 1987; Gerber 2010; Mondak 2010). The five personality factors include conscientiousness, extraversion, openness to experience, agreeableness, and neuroticism. Specifically, we expect interviewers high on conscientiousness, extraversion, openness to experience, and agreeableness to obtain higher response rates. Those who are low on neuroticism should also obtain higher response rates. We test these predictions with data from a local-area RDD telephone survey conducted at a university survey research

center in the fall of 2010. Roughly 30 interviewers conducted the interviews. In multivariate analyses we test the effects of interviewer personality characteristics on refusal, cooperation, and callback rates on the first and subsequent call attempts while controlling for the effects of other interviewer characteristics such as experience, concerns about gaining respondent cooperation, and other demographic characteristics.

The Influence of Personality Traits and Motives for Joining on Participation Behavior in Online Panels.

Florian Keusch, Vienna University of Economics and Business, Austria (florian.keusch@wu.ac.at)

Due to the dropping penetration rate of landline telephones and a general decline in the willingness to participate in telephone surveys, data collection using online methods – especially online access panels – are becoming more popular all over the world. Though widely adopted among marketing researchers critics still fear that the new sampling methodology leads to biased results produced by a breed of survey-savvy volunteer respondents who are solely interested in monetary incentives and therefore cannot be compared to the general population. This study seeks to give in-depth insight into the personality of online panelists by analyzing their motives for joining the pool as well as their personality traits (Big Five, materialism).

In a survey among 1,729 members of an Austrian online access panel participants were asked to answer standardized scales measuring materialism (Richins 1987) and the Big Five personality traits (Rammstedt & John 2005). In addition they had to name the reasons for joining the online panel. The data was then analyzed for its influence on participation behavior in the pool during the last 13 month. The results show that money is a relevant motive but not the dominating reason for becoming an online panel member. It is also found that psychographic characteristics have rather little influence on participation behavior within the panel.

Motivated Underreporting In Screening Interviews.

Frauke Kreuter, Joint Program in Survey Methodology (fkreuter@survey.umd.edu); Stephanie Eckman, Institute for Employment Research (steph.eckman@gmail.com); Roger Tourangeau, University of Michigan (rtourangeau@survey.umd.edu)

Many national surveys include screening interviews intended to identify members of the eligible population or members of subgroups slated for oversampling. Underreporting of members of these groups drives up survey costs and may introduce bias into the estimates. There is evidence that members of the target populations are often underreported in screening interviews. One of the best documented instances of such a screening shortfall occurred in the National Longitudinal Survey of Youth, 1997 Cohort, with a coverage ratio of only about 70 percent for the targeted age group. All other age groups have coverage rates above 90 percent. These screening shortfalls could reflect respondent motivation to screen out rather than refuse (and be subject to conversion attempts). They could also reflect interviewer motivation to screen households out, since interviewers are often graded based on their nonresponse rates but not based on their eligibility rates.

This presentation reports results from an experiment in a telephone survey; the experiment varied two factors designed to affect the motivations of the household informant and one factor designed to affect interviewer motivation. We varied (1) the amount of information about the target population in the advanced letters; (2) the format of the screening questions (full household rosters versus screening questions that explicitly ask about the target population); and (3) the interviewer payment scheme (bonuses for completed interviews versus bonuses for completed screeners). Our design allows us to estimate main effects of each of these factors as well as their interactions. We expect differences in

eligibility rates across the different payment groups if the interviewers contribute to the motivated underreporting effects seen in the literature.

Multivariate Models of Intended 2010 Census Participation.

Gina K Walejko, Science and Technology Policy Institute (gkwalejko@gmail.com); Peter V. Miller, Northwestern University (p-miller@northwestern.edu); Nancy Bates, U.S. Census Bureau (nancy.a.bates@census.gov)

The Census Continuous Tracking Survey (CCTS) measured knowledge, attitudes, intention to participate, and reported participation in the 2010 Census from early December, 2009 to late April, 2010. The study was designed to provide real time information to Census managers about public views of the Census and likely participation. This paper takes a retrospective look at the entire tracking data set (n>24,000) to identify how attitudinal and demographic factors influenced intended and reported participation in the Census during the months prior to, and the weeks following Census Day, April 1, 2010. Multivariate models examine the relative predictive power of demographic and knowledge/attitudinal predictors combined, as well as the influence of individual variables, particularly awareness of Census campaign messages.

Interviewer Effects & Performance Measurement.

Gender of Interviewer Effects on In-Person Surveys on Potentially Sensitive Topics.

Joseph W Lenski, Edison Research (JLenski@edisonresearch.com)

Edison Research has conducted three national surveys of undergraduate college students for the Associated Press and mtvU. These three surveys included an experimental design to measure the effect of the gender of the interviewer on the results with approximately one-half of the completed questionnaires at each location handed out by male interviewers and one-half handed out by female interviewers. The three surveys used the same methodology of in-person self-administered interviews conducted at a random sample of 40 four-year universities around the country. The first survey was conducted in February and March of 2008, the second survey was conducted in April and May of 2009, and the third survey was conducted September 2010. The sample sizes were 2,253 interviews for the 2008 survey, 2,240 interviews for the 2009 survey, and 2,207 interviews for the 2010 survey. The results of the surveys are analyzed to determine the impact of gender of interviewer on the demographics of the samples obtained and for the results of specific questions on potentially sensitive topics. Our conclusions are that college students are more likely to express worries about finances, safety on campus, mental health and emotional distress when handed a questionnaire by a woman than when handed a questionnaire by a man. The analysis will also include response rates by gender of interviewer crossed by gender of potential respondent.

Interviewer Gender Effects on Survey Responses to Marriage-Related Questions.

Mingnan Liu, Purdue University (mingnanliu@gmail.com); Kevin Stainback, Purdue University (kstainba@purdue.edu)

This research draws from social desirability perspectives to examine the effect of interviewer's gender on interviewee's responses to survey questions related to marriage utilizing the 2006 China General Social Survey (CGSS). In particular, this article examines two specific research questions. First, do respondents provide different responses to male and female interviewers? Second, do the effects of interviewer's

gender differ for male and female respondents? The specific items examined include, (a) if married men are happier than unmarried men, (b) if married women are happier than unmarried women, (c) if getting married is better than being single, and (d) if staying married is better than getting divorced. The results indicate that interviewer's gender significantly influences responses to marriage related questions, with female interviewers receiving more pro-marriage responses than male interviewers. Further analyses reveal that female respondents provided significantly different responses to the marital happiness questions to female than male interviewers, with female interviewers receiving more pro-marriage responses. Male respondent's, however, were unaffected by interviewer's gender on these questions. We conclude with implications for future research.

The PAIP Score: A Propensity-Adjusted Interviewer Performance Indicator.

Brady T West, Institute for Social Research (bwest@umich.edu)

Evaluating interviewers based on their response rates is complicated in most surveys. By random chance, interviewers may call cases that are more or less difficult to interview. In addition, interviewer response rates can only imperfectly be computed because of the contributions of other interviewers' prior contacts with those cases (calling a case after a contact from an expert interviewer may pose different difficulties than calling a case after a contact from an inexperienced interviewer). This paper proposes and evaluates an interviewer performance indicator that attempts to repair these weaknesses. The proposed indicator is computed using a three-step algorithm. First, for each active case, available paradata are used to estimate the propensity that the next contact with the case will generate an interview. Second, if the interviewer assigned the case obtains a successful interview on the next contact, the interviewer receives a score of 1 minus the estimated response propensity for the contact; a non-successful contact by the interviewer results in a score of 0 minus the estimated response propensity for the contact. Finally, for each interviewer, the contact-level scores are averaged over all contacts, resulting in a propensity-adjusted interviewer performance (PAIP) score. Addressing an important drawback of previous interviewer performance measures discussed in the literature, this performance indicator gives large credit to the interviewer who obtains success on very difficult cases, and only a small penalty given failure with such cases. The indicator gives only small credit to success on very easy cases and larger penalties given failure with easy cases. This paper illustrates computation of the PAIP score using two different surveys (one face-to-face, one telephone), and assesses the validity of the indicator as a new metric for evaluating the performance of interviewers.

Multifactor Approach to Assessing the Quality of a Face-to-Face Interview.

Michael W Link, The Nielsen Company (Michael.Link@Nielsen.com); Pambos Charalambous, The Nielsen Company (Pambos.Charalambous@Nielsen.com); Jennie Lai, The Nielsen Company (Jennie.Lai@Nielsen.com); TraShawna Boals, The Nielsen Company (Trashawna.Boals@Nielsen.com)

Face-to-face interviewing is often characterized as producing the highest quality survey data when compared to other modes, yet rigorous assessments are rare and often limited to computer-assisted recording approaches. Interviews with non-traditional populations, such as establishment owners where recall may bias attempts to collect specific factual information, often make assessment even more difficult. Here we provide a multifactor conceptual framework for assessing the quality of face-to-face interviews in an ongoing panel of store owners. The model includes attributes of the environment and context in which the interview occurs (level of privacy, number of interruptions), characteristics of the interviewer (experience, number of previous interviews with the same store), store owner attributes (knowledge level of store products, interest/focus during interview), interviewer-store owner interaction (language barrier, length of time store has been in the panel), and interview characteristics (length,

availability of supporting documentation). Interviewers completed a brief 20 question assessment after interviews with more than 2,500 store owners in August 2010, which included their assessment of the quality of the interview. Additional empirical information about the store and survey responses are also used to develop a model to predict first the interview's assessment of quality and then a more objective measure of quality based on how closely store responses mirror those of owners from similar stores. The model shows that factors at all levels, particularly those related to the shop owner and experience of the interview are key for understanding the quality of the data collected. The results have implications for interviewer training, face-to-face interview procedures and how researchers assess the quality of complex survey efforts.

Increasing the Efficiency of CATI and CAPI Monitoring Operations at Statistics Canada.

Caroline Pelletier, Statistics Canada (caroline.pelletier@statcan.gc.ca)

Computer-assisted interviewing operations require interviewers to have good interviewing skills. For example, they must simultaneously maintain a good rapport with the respondent, clarify the meaning of a question if necessary, ask probing questions when the responses are ambiguous, categorize the responses correctly, and enter the data accurately in a live production environment. Monitoring provides an opportunity to evaluate the interviewers' skills. Formal monitoring is already in place for evaluating the performance of interviewers conducting computer-assisted telephone and personal interviews (CATI and CAPI). They provide an objective evaluation of the interviewer's performance at systematic points in time and help identify potential problems with the questionnaire. CATI monitoring involves unobtrusively listening to the interaction between the interviewer and the respondent while CAPI monitoring involves digitally and unobtrusively recording portions of interviews and then listening to the recordings.

Several activities were undertaken recently at Statistics Canada to increase the efficiency of CATI and CAPI monitoring operations. These activities will be covered in this presentation.

- The evaluation criteria which used to differ among social, agricultural, and business surveys were harmonized.
- The definition and frequency of monitoring sessions were adjusted to better respond to operational needs.
- Priorities for monitoring interviewers were automatically derived. Instead of having each site identify the interviewers to be monitored during the day, each site is provided with a priority list. The list is generated objectively based on statistical rules and parameters available in the monthly work schedule. It avoids over and under-sampling of interviewers and meets the expected number of monitoring sessions per interviewer.
- The information pertaining to the monitoring is now automatically filled online before reports are generated.
- The content of reports was improved to provide relevant, timely, and constructive feedback to interviewers, supervisors, monitors, and managers.

Media Exposure & IT Influence on Opinion and Behavior.

The Internet and Media/Social Time Displacement Since 1995.

John P. Robinson, University of Maryland (ROBINSON@SOCY.UMD.EDU)

A major concern about new technologies, like the Internet and other IT, is how much they might displace time spent on existing media and social interaction. Three highly-publicized early studies of the initial impact of IT indicated it was reducing time on both social life and mass media use. However, a number of

high-quality national surveys since then – from the Pew Center, the General Social Survey (GSS), the Survey of Public Participation in the Arts (SPPA) and the American Time Use Survey (ATUS) -- have not replicated these results. Indeed, they find some support for Internet and other IT use sometimes being associated with increased social life and media use (especially reading).

The present article examines whether Internet users spend less time involved in arts events and other leisure activities. It uses data from two separate high-quality national surveys about the public's arts and leisure participation (the GSS2110 and the Census-Bureau's SPPA2008) to examine whether users of IT participate less in various arts other leisure activities. Both show that, even though the two studies used different IT questions and examined different arts activities over different years, progressively higher arts participation was found among Internet users, generally by heavier Internet users after education and other predictors of arts activity are controlled. The SPPA data also show that users and heavier IT users are also more active in a variety of other free time activities like movies, volunteering and sports. IT use here then becomes a way of extending or enhancing live attendance rather than displacing it. However, the results also raise questions about whether these correlations result from a "response set" of respondents overestimating their activity participation. However, they still provide minimal evidence of Internet's ability to reduce time on other leisure activities.

The Knowledge Gap in Generation X.

**Jon D Miller, International Center for the Advancement of Scientific Literacy
(jondmiller@umich.edu)**

The knowledge gap hypothesis is now 40 years old and has become a standard component of communication theory. In other work, Miller, Garcia, Fernandez, and Salmon (forthcoming) have examined changes in the knowledge gap over the last 20 years using a series of cross-sectional national studies. This paper uses the 23-year record of the Longitudinal Study of American Youth (LSAY) to examine the development of the knowledge gap in a cohort of young adults (now age 35 to 39) that represent Generation X. The LSAY began with a national probability sample of 7th and 10th grade public school students in 1987 and approximately 4,000 of the 5,000 eligible participants provide annual data to the LSAY on a variety of variables.

To examine the development of the knowledge gap in Generation X, this analysis examines three outcome variables: (1) a measure of civic scientific literacy that reflects the ability of each respondent to read and understand science in popular media, (2) a measure of understanding of influenza taken in the midst of the 2009 epidemic, and (3) a measure of understanding of climate change as a public policy issue. These variables reflect Tichenor, Donohue, and Olien's (1970) original proposition that "education is a powerful correlate of acquisition of knowledge about public affairs and science from the mass media" (TDO, 1970, p. 160). A set of structural equation models are used to estimate the relative influence of gender, education, religious attitudes, and various kinds of media consumption to each of the three outcome variables.

The findings show that (1) the knowledge gap is alive and well in Generation X, (2) formal education and Internet use combine to produce larger differences that were found four decades ago in TDO's original work.

Influences From the Mass Media, the Internet and Social Media on the Time Trend of Public Opinion About the Toyota Brand.

**David Fan, University of Minnesota (dfan@umn.edu); David Geddes, Evolve24
(david.geddes@evolve24.com); Felix Flory, Evolve24 (felix.flory@evolve24.com)**

Methods have been developed for predicting the impact of persuasive information on the time trend of public opinion and have been used to predict the time trend of the Index of Consumer Sentiment from mass media coverage with R-squared of 0.83 for 1978-2000 (Fan and Cook, Journal of Mathematical Sociology, 2003; 27:1-23).

At the time of those studies, electronic databases like LexisNexis included the majority of news content in the press. For a topic like Consumer Sentiment, most people were likely to follow the press in the same way that the public forms its opinion about the weather from weather forecasts.

In more recent years, however, there has been an explosion of electronic information available to the public. Therefore, it is important to test the impact on opinion due to these new information sources.

For such tests, we will to predict the time trend of public opinion for a sample designed to represent the general population. These dependent variable data will be the BrandIndex of YouGovPolimetrix, Inc. computed as the average of six questions about Toyota on: general impressions, quality, value, satisfaction, recommendation to friends, and corporate reputation. The data are collected daily and can be aggregated for any appropriate time interval.

In an extension of prior methods, our predictors will be a wide range of persuasive communications including social media, blogs, and online news. The data will come from Evolve24 and will be scored by computer for ideas likely to persuade the public about improving or declining BrandIndex values for Toyota.

A crucial test will be the extent to which these systematically available data can represent the effects of advertising which is much more difficult to capture. Such studies can be extended to politics where candidates with massive investments in advertising nevertheless lost.

Response to Surveys of High-Profile Topics: The Effects of Media Coverage and Public Engagement on Response to the National 2009 H1N1 Flu Survey.

Nicholas D Davis, NORC at the University of Chicago (davis-nicholas@norc.org); James A Singleton, Centers for Disease Control and Prevention (xzs8@cdc.gov); Lina Balluz, Centers for Disease Control and Prevention (lib7@cdc.gov)

The National 2009 H1N1 Flu Survey (NHFS) was conducted by NORC at the University of Chicago on behalf of the Centers for Disease Control and Prevention from October 2009-June 2010. The purpose of the survey was to monitor national and state-level influenza A (H1N1) monovalent and trivalent seasonal vaccination coverage rates during the 2009-2010 season. This survey was conducted during a period of considerable media attention to influenza. The goal of this paper is to identify the impact of that attention – as well as corresponding shifts in public interest in influenza – on the CASRO response rate and its components (resolution, screener completion, and interview completion rates). A comparison of NHFS response rates over time to changes in the same rates for the Behavioral Risk Factor Surveillance System (BRFSS), another health-related but not influenza-specific survey, is examined to help quantify this effect. By separating long-term trends in response from shorter-term changes and comparing to aggregate indices of media coverage and public interest in the “flu,” the paper aims to determine whether there is evidence that changes in media attention have either long-term or short-term effects on response. We attempt to determine the effects of changes in response with regard to the potential for non-response bias in survey estimates, identifying practical implications for conducting future pandemic influenza or other surveys during public health or other emergencies.

Information or Affirmation?

Nicholas Brigham Schmuhl, University of Wisconsin-Madison (nbschmuhl@wisc.edu); Albert C. Gunther, University of Wisconsin-Madison (agunther@wisc.edu); Heather Akin, University of Wisconsin-Madison (heathereakin@gmail.com)

Recently, mass communication scholars, media analysts, and members of the public have expressed concerns about the partisan “echo chamber” surrounding discourse of political and social issues. In a media landscape defined by an explosion of online sources and hundreds of cable television channels, polarized audiences are increasingly able to seek and find agreeable content, perhaps blunting the information function of the mass media. We may find ourselves in a new era of minimal effects in which the media consumer – assisted by new social media-inspired selection algorithms – utterly controls his or her own exposure to information.

To test these ideas, this study employs a web-based experiment and a sample of highly involved partisans on the issue of childhood vaccinations and their purported connection to autism spectrum disorders. After measuring preexisting attitude extremity and group identification, we manipulate participants’ initial exposure to either congenial or counter-attitudinal news articles on the topic. We then achieve a behavioral measure of subsequent media selectivity, allowing participants five minutes to choose and read from a randomized sample of related articles from the actual online media environment. Tracking both sequence of choices and reading time per selection, we investigate whether partisan audiences are likely to choose agreeable media content, whether initial exposure affects subsequent selections, and the extent to which media selection is influenced by factors such as group identification, affective predisposition, and perceived reach of information. Further, we reexamine traditional cognitive dissonance theories, seeking a more nuanced mechanism of selective exposure, perhaps mediated by participants’ preferred affective states.

Early analysis suggests a robust pattern of motivated selective exposure. These response patterns hold major implications regarding the mechanisms that may be driving an increasingly polarized issue-public. Our findings extend concepts of selective exposure from partisan politics into the issue-oriented arena.

Proliferation of Polls (Good & Bad) & the Way Media Uses Them with a Less-Than-Critical Eye.

Proliferation of Polls (Good and Bad) and the Way Media Uses Them with Less than a Critical Eye.
Susan Pinkus, S.H. Pinkus Research Associates (spinkus@sbcglobal.net); Mark DiCamillo, The Field Poll (markd@field.com); Phil Trounstone, CalBuzz (phil@trounstone.com); Whit Ayres, Ayres, McHenry & Associates, Inc. (whit@ayresmchenry.com); Mark Blumenthal, Huffington Post (mark@huffingtonpost.com); Peter V. Miller, Northwestern University (P-Miller@northwestern.edu)

There has been a proliferation of polls -- ones that don't disclose methodology publicly or even when asked and the media that is uncritical of these polls that they are writing about. There have been several articles plus an open letter submitted by 10+ leading partisan pollsters (Dems and Reps) who have openly criticized these polls (that unfortunately are lumped together with good, sound methodological polls). I would like to get together a panel of pollsters/media to talk about this issue which is extremely important to the polling industry. This would dovetail nicely with someone from AAPOR council to talk about the transparency initiative. For example Phil Trounstone at CalBuzz.com wrote an article, "Beware Murphy, Rasmussen and Other B.S. Artists" (about the Meg Whitman campaign and the spin on robopolls

during the CA gov's campaign) and Mark DiCamillo who wrote a piece in CalBuzz on "Why increased robopolling in California is Troubling."

Racism and Sexism in Contemporary Society.

What Black and White Americans Believe about Racial Progress, and How it Matters for Affirmative Action.

Matthew DeBell, Stanford University (debell@stanford.edu)

Support for policies intended to promote racial equality or to help disadvantaged groups is conditional upon beliefs about racial progress – or at least, it is alleged to be. Studies have concluded that support for policies promoting racial equality depends on the belief that racial progress has already occurred (e.g., Brodish et al., 2008). But this literature has used flawed measures and conceptualization of opinion about progress and has made generalizations about the population without representative sample data to warrant such generalizations.

The present study uses new measures of beliefs about racial progress from a nationally representative sample (ANES 2008-2009 Panel Study). The paper presents improved measures of opinion about progress and the multifaceted rationale for preferring the new measures to the old ones. Next, it describes the differences between blacks and whites in perceptions of progress toward equality when these improved measures are used. Finally, it presents analysis of the role of such opinions in shaping attitudes toward affirmative action policies intended to promote equality. Contrary to previously published findings, this study demonstrates that support for equality policies (affirmative action) is not conditional on beliefs in racial progress. Instead, policy preferences depend strongly on beliefs about something related to progress but distinct from it: how much current racial conditions differ from the ideal. Implications for psychological theory founded on now-falsified claims about the role of racial progress are discussed.

Discriminatory Attitudes in Europe: As Hard To Get At As in the USA?

Mario Ignatov, University of Cincinnati (m@rioignatov.com)

European integration is presently based on four basic freedoms: free movement of goods, capital, services and persons. The lack of EU-harmonized fiscal and labor force integration proved to be a major drawback in fighting the ongoing world economic recession. The lack of EU-harmonized immigration policy runs similar risks, given the continent's bitter experience of ethnic conflict in the recent past. This research project comes to grips with the methodological problem of discovering very little variance in the Eurobarometer-selected outcome measures explained by various predictors. Some of this may well be due to the design of the survey questions being unable to meet certain validity criteria, or more likely to the substantive matter of discrimination being a highly sensitive issue, subject to social desirability concerns, thus colliding with expressing politically correct views. The current study has drawn upon previous research in social identity literature and developed a model for assessing individual background characteristics that theoretically predict discrimination against immigrants and ethnic minorities. Specifically, the author tested six hypotheses, derived from an existing theory about who the "average" European ethnic discriminator might be. The tests were based on secondary analysis of Eurobarometer survey data, using multiple and logistic regression techniques. The results confirmed four of the six hypotheses. The strongest predictors of individual discriminatory attitudes were having acquaintances from various minority groups, political ideology, education and community size. Nevertheless, the model is not nearly robust enough to explain significant amounts of variance, which means that the Eurobarometer will need to find new ways of measuring discriminatory attitudes by developing better survey questions that come to grips with the social sensitivity of expressing discriminatory attitudes

towards ethnic minorities, much like the problem of measuring racial attitudes in the USA, as Howard Schuman and others have amply demonstrated.

Racial Aversion, Obama's Election and Governing a Divided Country.

Monika L McDermott, Fordham University (mmcdermott@fordham.edu); Cornell Belcher, brilliant corners Research and Strategy (mmcdermott@fordham.edu)

Many pundits and academics have declared that Barack Obama's victory in the 2008 election demonstrated that race was not a factor in voters' decisions. In contrast, this paper argues that race actually did matter to voters' reactions to Obama, and to both their vote choice and presidential evaluations after the election. Specifically, racial aversion – an index of negative racial attitudes measured in repeated battleground surveys over the course of the election and post-election – was a factor in voters' preferences. Analysis of political surveys demonstrates that for the most racially averse voters' racial attitudes were key in their voting decisions, but for the marginally racially averse these attitudes were not activated during the election. Since Obama took office, however, previously dormant racial attitudes have begun to play a larger role in individual judgments, resulting in negative attitudes toward Obama and his policies.

Exploring Women's Status and Rights in Iraq.

Kevin Tillmann, D3 Systems (kevin.tillmann@d3systems.com)

The status of Iraqi women in daily practice has remained largely hidden from international attention since the U.S.-led invasion of Iraq in 2003. Even with extensive coverage of political instability, increasing violence, and significant turmoil in the region, questions still surround what happens in the everyday lives of women in the country.

There have been some documented instances of women running for political office and other jobs in the public sector, prompting a public perception that improved security has brought improved freedom for Iraqi women. Previous research has shown that the majority of Iraqi women serve a domestic role in society, with very few in high positions of employment (D3 Systems, 2007). Most women indicate dissatisfaction with this situation.

This paper publicizes change in the opinions of women about their daily lives from March 2007 to November 2010 in Iraq. The analysis tests the publicly held hypothesis that Iraqi women today enjoy higher life satisfaction and freedom than their female peers in March 2007. It establishes a definitive measurement of happiness across Iraq's national female population and provides analysis of daily life within various attitudinal and demographic segments. It especially focuses on issues of domestic and social gender conflict; personal finances and purchases, occupation status, personal safety, and access to health care and education.

The D3 Systems studies were conducted face-to-face by native Iraqi and Kurdish interviewers nationwide in Iraq in March 2007 and November 2010, respectively. Total n=1,100 in each survey with a sampling margin of error +/- 3%. The surveys were national randomly sampled studies of the general population throughout the country.

Predicting Biased Behavior With Implicit Attitudes: Results From A Voting Experiment.

Cecilia Hyunjung Mo, Stanford University (mo_cecilia@gsb.stanford.edu); Wendy Gross, Stanford University (wtgross@stanford.edu); Tabitha Bonilla, Stanford University (tabitha@stanford.edu)

Although explicit racism has nearly vanished, and explicit sexism has decreased, many scholars worry that voters may still exhibit implicit sexist or racist voting behaviors. In other words, voters decide "as if" they were sexist or racist even though they may genuinely believe themselves to be unbiased. Yet, scholars differ on how to measure these implicit attitudes and behaviors. In September 2008, we conducted an online experiment to examine the relationship between (1) sexist voting decisions and (2) prominent measures of implicit like the Implicit Attitude Test (IAT) and Affect Misattribution Procedure (AMP). Through this 2008 experiment, we test the predictive validity of the two most prominent measures of implicit sexist attitudes. Implications for research design and analysis of implicit attitudes measures are discussed, including the American National Elections Study's decision to include the AMP over the IAT.

Friday, May 13, 2011

8:00 a.m. - 9:30 a.m.

Concurrent Session A

Assessing and Adjusting for Unit Nonresponse Bias.

Analyzing Nonresponse and Nonresponse Bias in the Occupational Employment Statistics Survey Using Regression Trees.

Polly Phipps, Bureau of Labor Statistics (pphipps_p@bls.gov); Daniell Toth, Bureau of Labor Statistics (toth_d@bls.gov)

To analyze nonresponse and nonresponse bias for the Bureau of Labor Statistics Occupational Employment Statistics survey (OES), we build a regression tree model. The tree model identifies employment size, industry sector, multi-establishment status, and metropolitan area size as having a significant impact on an establishment's propensity to respond, with important interaction effects. Among small establishments, organizational complexity drives the response rate, while for establishment with large employment, the industry and area population size are important. Using administrative wage data linked to the sample, we investigate the possibility that the establishment characteristics are related to wages; wage estimates are a major OES outcome variable. Fitting the same regression tree model to the wage data, we find that categories with the lowest response rates have an above average wage per employee, while those with the highest response rates have below average wages. Differences in average wages between respondents and nonrespondents persist across the categories, giving cause for concern that the nonresponse bias could be non-ignorable. We discuss the advantages of using regression trees to investigate nonresponse, and outline possible remedies to deal with potential nonresponse bias, including: focused contact, nonresponse follow-up, and responsive design; adoption of different types of data collection methods; and use of administrative data as auxiliary data in nonresponse adjustments.

Identifying Sources of Non-Response Bias: A Mixed-Mode Health Survey of a Low-Income, Culturally Diverse Population.

Donna McAlpine, School of Public Health, University of Minnesota (mcalp004@umn.edu)

Lower response rates in surveys has resulted in increased efforts to encourage participation, including creative incentives, mixed-mode designs, refusal conversations and more aggressive attempts to reach

potential respondents. While such efforts appear to increase response rates, little is known about whether they reduce response bias.

Understanding response bias requires information about the entire sampling frame. Even when this information is available, researchers have commonly grouped the sample into completes and non-completes, which may mask differences between types of non-respondents. In contrast, this study examines total non-response bias, and bias associated with specific reasons for non-participation.

Data come from a mixed-mode survey (mail and telephone) of enrollees in a public health care program (RR 44.3%). Information for 10,455 eligible participants was available from enrollment applications (demographics) and claims data (utilization and diagnoses). Dispositions were categorized as completes (mail or phone), refusals, unable to locate, no answers, technical barrier, unable to participate and other.

The phone survey increased the participation of respondents from minority communities; but, overall, non-respondents were disproportionately African American, Asian or American Indian. Respondents to the phone survey were healthier than those to the mail survey; thus, the final sample of completes was similar to non-completes on these measures. The final sample underrepresented those who used more intensive care such as the emergency room.

The most common reason for non-completion was being unable to locate the respondent (56%), compared to only 8% refusing. Reason for non-response was associated with bias. Being unable to locate respondents increased the under-representation of minority groups and those who used more intensive services, but non-response due to 'no answers' reduced such biases.

Reducing response bias, not increasing response rates, should be our focus. Given limited resources, researchers should be thoughtful about which types of non-responders to target to reduce bias.

Assessment of Bias When Field Operations are Curtailed in a Mixed Mode Telephone and Face to Face Survey of Persons with Disabilities.

Eric Grau, Mathematica Policy Research, Inc. (egrau@mathematica-mpr.com)

In-person interviews are generally seen as necessary to adequately survey a population of persons with disabilities since sample members' disabilities may preclude them from completing an interview by telephone or mail. The National Beneficiary Survey (NBS), a mixed mode survey of persons with physical and/or mental impairments, includes both Computer-Assisted Telephone Interviews (CATI) and Computer-Assisted Personal Interviews (CAPI). The purpose of this analysis is to assess how estimates are affected, if field operations, needed to obtain CAPI data, were scaled back or eliminated in the NBS. We focus our attention on a subpopulation of beneficiaries of Supplemental Security Income or Social Security Disability Insurance who participated in the Ticket to Work program at least once between January 1, 2005 and October 2, 2005. We use administrative paradata with process details about specific sample members that records each locating and interview attempt. We attempt to reconstruct disposition codes as if (1) there had been no field operations and (2) field operations were limited to 60 days. We then compare attributes of respondents and nonrespondents in the original design with those defined as respondents or nonrespondents with less field effort. We look at comparisons of age, race, gender, and key variables that are available for (almost) all members of the target population. In addition, we compare selected survey estimates to assess whether the differences in responses from CAPI and CATI as observed in Sloan, Wright, and Barrett (2006) result in differences in the estimates when CAPI efforts are removed or scaled back.

Evaluating the Impact of Interviewer Observed Auxiliary Information in Nonresponse Adjustments.
Jeffrey M Gonzalez, U.S. Bureau of Labor Statistics (gonzalez.jeffrey@bls.gov)

Nonresponse weighting methods encompass a variety of postsurvey statistical adjustments to compensate for unit nonresponse. They are primarily viewed as a tool for reducing bias due to nonresponse but can potentially increase the variance of the adjusted estimate. Effective auxiliary information for use in the nonresponse adjustment should be highly predictive of response and the key survey outcomes (Little and Vartivarian, 2005). When information of this type is used, it has the potential to reduce nonresponse bias without increasing the variance estimates.

Obtaining information on the full sample that satisfies those criteria can be challenging, for sampling frames often do not include rich auxiliary information that is related to key survey variables. A potential way to circumvent that problem is to have interviewers make and record observations about each sample unit when locating, contacting, and soliciting participation; thus, information would be available for the respondents as well as the nonrespondents. Such interviewer observations have been collected as part of the U.S. Consumer Expenditure Quarterly Interview Survey (CE) since 2005. Interviewers are asked to note the sample unit tenure status and any concerns regarding participation in the survey (e.g., distrust of government) for all sample units regardless of cooperation.

Using the CE data this research will investigate the value of incorporating interviewer observed auxiliary information into a nonresponse adjustment for the CE. We first examine the strength of the bivariate correlations between each of tenure and a perceived concerns index (a variable summarizing the extent to which a sample unit expressed any concerns about completing the interview) with response propensity and key survey outcomes (e.g., expenditures). We then estimate response propensities and examine differences in point estimates when this information is and is not included in the nonresponse adjustment. Finally, we evaluate the resulting impact on the variances of expenditure estimates.

Using Nonresponse Propensity Scores to Improve Data Collection Methods and Reduce Nonresponse Bias.

Morgan S Earp, USDA National Agricultural Statistics Service (Morgan_Earp@nass.usda.gov)

The USDA's National Agricultural Statistics Service (NASS) conducts the quarterly Crops/Stocks Survey. The Crops/Stocks Survey collects detailed data on crop acreage, yields, production, and quantities stored. These data are used to set national and state level estimates of acres planted, harvested, production, and on-farm grain stocks. Nonresponse error is a huge concern, especially when setting estimates at the state level. Using administrative and previously reported data, we created nonresponse propensity scores. These propensity scores were provided to state offices along with sample masters, and states were asked to document any refusal/noncontact avoidance methods used. We analyzed six quarters of data to predict likely nonrespondents and establish baselines for assessing the effectiveness of state level treatments.

This paper discusses the use of classification trees to create nonresponse propensity scores and assesses the utility of proactively employing them in data collection to reduce nonresponse rates and nonresponse bias. This study aims to assess and compare state level refusal/noncontact avoidance techniques over time, specifically those targeting establishments that have the greatest effect on state and national estimates.

This research will enable NASS to determine whether prior knowledge of nonresponse likelihood can be used to improve data collection methods and ultimately reduce nonresponse bias.

Cross Cultural Survey Methods: II.

The Color of Culture.

Peter Mohler, University of Mannheim Germany (peter.mohler@uni-mannheim.de)

Culture is the core cohesion factor of societies. If there is no common culture, there is no stable societal cohesion. Values play a very prominent role in defining a common culture. The question is, whether the American political culture is color-blind in the sense that being proud to be American unites Americans irrespective of the color of their skin.

Habitually American surveys use race as a standard demographic in their analysis, differentiating often between African American, Hispanic, White, and Others. The first three categories are rooted in differences of skin color (even the now "soft" Hispanic one).

It can be shown that the relation between one's skin color (vulgo race) and one's culture as measured by surveys is not a valid measure of a common culture.

Evidence from an experiment measuring value orientations could not find statistically difference between African-American and White students on collectivism-individualism. Really surprising are results from an earlier GSS oversampling of African-American respondents asking for 'feeling American'. It can be shown that standard analyses using Black or African-American as a single demographic indicator fail to describe accurately a highly heterogeneous group: two thirds of African-American respondents report answers that are identical to those of Whites and Hispanic. One third, however, reports more different answers. This one third of respondents creates the statistical significant different results between Whites and African Americans.

Melting these two sub-groups of the African-American respondents into the 'Black' category does great injustice to each of them: the former are, now for decades, reported to be "different" from other American groups, while sharing the same values, and the latter, as their grief and burden is 'softened' by the former and is not as visible as it should be.

Thus the color of culture is there but not as we are used to see it.

Uneasy Coexistence: Current Relations between Aboriginal and Non-Aboriginal People in Canadian Cities.

Sarah Robertson, Environics Research Group (sarah.roberton@environics.ca); Keith Neuman, Environics Research Group (keith.neuman@environics.ca)

Relations between Aboriginal and non-Aboriginal peoples in North America have been troubled for several hundred years. While considerable progress has been made in the past few decades, many issues remain unresolved, and tensions continue to fester where these populations intersect. Aboriginal peoples continue to lag well behind the rest of the population on most social, economic and health indicators, and continue to face racism and discrimination. In Canada, most of the attention has been devoted to addressing the issues facing First Nations peoples living on reserves, and little is known about the more than 50 percent of the Aboriginal population now living in the country's cities.

In 2008-09, the Environics Institute conducted the first-ever national study of urban Aboriginal peoples in Canada, to advance and reframe the national conversation with and among Aboriginal peoples through an innovative and multifaceted research study. This research sought to better understand and document the experiences, aspirations, expectations, and identities of Aboriginal people, exploring new areas of inquiry such as the factors leading Aboriginal peoples toward success, autonomy, cultural confidence and spiritual meaning. Further information about this research is available at www.uaps.ca

The study included major surveys of both Aboriginal and non-Aboriginal Canadians in 10 cities across the country, providing a unique opportunity to better understand the perspectives these two groups have toward each other. This paper will address three aspects of the study:

- a) How Aboriginal peoples are living within the larger non-Aboriginal context of urban life in Canada;
- b) How urban Aboriginal and non-Aboriginal Canadians view each other, and the extent to which their respective perceptions reflect true understanding versus stereotypes and misconceptions; and
- c) How an innovative research methodology was developed to successfully conduct this research with a hard-to-locate population, within a politically charged socio-cultural context.

What makes Minority Populations Hard-to-Reach in a Cross-Cultural Perspective?

Mathew Stange (mathew.stange@huskers.unl.edu)

Sometimes it is difficult to collect data from either subgroups of a survey's target population or the entire target population. Various factors contribute to this. Such populations are often described as "hard-to-reach", even if the actual reaching of them is not always quite the problem. Groups that feel stigmatized or seek to go unnoticed might be hard-to-find and contact or if contacted might be hard to persuade to participate. Others might pose difficulties in terms of interviewing requirements or the questions asked. When populations cannot be interviewed or the interview/questionnaire is burdensome or threatening for respondents, data quality is often impaired. Generalizations made to the hard-to-reach population or to a population that includes one or more hard-to-reach groups can thus be limited. Additionally, if survey instruments are not designed to fit specific needs of a hard-to-reach population measurement error can result.

Little methodological research has examined the characteristics that earn populations the description hard-to-reach in either survey research in general or in cross-cultural and cross-national survey research. This paper reports aspects of a systematic literature review undertaken to identify the range of groups often described as hard-to-reach in cross-cultural survey research, to ascertain what characteristics earn each of these populations this description and, ultimately, to establish similarities and differences across cultural groups. The paper focuses on a number of minority populations, such as documented and undocumented immigrant workers, ethnic minorities, and LGBT individuals. The research is part of a larger project aimed at attempting to address what makes populations hard to reach in terms of study design and implementation.

Design Considerations for a Cross-Cultural Enumeration Survey.

Teresa (Ye) Jin, The Nielsen Company (teresa.jin@nielsen.com); Justin T. Bailey, The Nielsen Company (Justin.Bailey@nielsen.com); Kelly Bristol, The Nielsen Company (Kelly.Bristol@nielsen.com); Michael W. Link, The Nielsen Company (Michael.Link@nielsen.com)

Researchers have acknowledged that understanding cultural differences, including the thoughts and behaviors of respondents in a multicultural context, are essential to create a valid and reliable enumeration survey. To encourage cooperation and optimize data quality, there are a wide variety of factors need to be considered, including survey mode, instrument structure, respondent's culture, question wording, and culturally valid response options.

In order to establish estimates of media usage, The Nielsen Company initiated an enumeration survey in mainland China using a computer-assisted telephone interviewing (CATI) mode. One primary goal of this research was to discover whether the newly designed telephone interview survey, using best practices for survey design, can reach the targeted response rate in an initial pilot study and result in continued improvements for implementation. Pilot data was analyzed to investigate item nonresponse and examine data quality. Meanwhile, qualitative interviews with the call center team were conducted to explore the effectiveness of the phoning scripts, as well as respondents' reaction during interviewing. Further analyses will be focused on discovering whether respondents with different demographic characteristics react to the survey differently.

The findings of this research will shed light on how to overcome methodological, cultural and operational barriers when developing a survey in a cross-cultural context. We will also share some lessons learned during the course of designing this survey for researchers involved in developing cross-culture surveys.

Do Incentives Affect Data Quality?

Effects of Differential Incentives on Response Rates in Four Countries for a Web-based Follow Up Survey: ITC Four-Country Wave 8 Survey.

Kathleen McSpurren, Survey Research Centre, University of Waterloo (kmcspurr@uwaterloo.ca)

On-line survey administration has many appealing features to researchers, most obviously, the potential cost savings compared to telephone or face-to-face modes. In addition, there may be a preference for self-administered web surveys by the public, given the convenience and minimal intrusion.

The paper is an analysis of an embedded experiment in a follow up survey in four English-speaking countries in 2010. The data is part of the International Tobacco Control (ITC) Project, being conducted in 20 countries around the world. The ITC project evaluates the effects of national tobacco policies using longitudinal survey data strategically collected before and after policy changes. The 20 countries included in the project represent half the world's population, and 60% of the world's smokers. The four English-speaking countries are: Australia, Canada, the United Kingdom and the United States.

The Four Country Recontact survey is in its eighth wave. Two experimental cells were included in the Wave 8 recontact survey, randomly assigning participants to a regular, or additional financial incentive, to encourage completion of an on-line version rather than once again participating by telephone-administered questionnaire. Differences in response rates by mode, as well as reported motivations to complete by web, would be analyzed in the paper. Specifically, the research questions are:

- a. Does the amount of incentive directly impact the response by web?
- b. Are there differences by country? Do differences by country exist in web uptake generally?

- c. Similarly, are there gender/age differences by country for uptake, and gender/age differences by incentive amount?
- d. Within the web respondents, further analysis would include stated mode preference for follow up by incentive amount, as well as tendency to answer honestly by web.

Effects of Mode and Incentives on Response Rates, Costs, and Response Quality in a Mixed Mode Survey of Alcohol Use Among Young Adults.

John Stevenson, University of Wisconsin Survey Center (stevens@ssc.wisc.edu); Jennifer Dykema, University of Wisconsin Survey Center (dykema@ssc.wisc.edu); Chad Kniss, University of Wisconsin Survey Center (ckniss@ssc.wisc.edu); Penny Black, University of Wisconsin (pblack@wisc.edu); Paul Moberg, University of Wisconsin (dpmoberg@wisc.edu)

A growing challenge in survey research is completing surveys with younger adults. Increased cell phone use in younger cohorts has made typical RDD telephone surveys increasingly ineffective. Researchers turn to other types of samples to help with coverage and response rate issues, often using self-administered surveys (web or mail). Each of these has their own pros and cons. For example, web surveys tend to be cheaper to conduct, have shorter field periods, provide greater flexibility with questionnaire design, and may be associated with lower levels of item-missing data and higher reports of threatening behaviors. Their limitations include coverage and response rates.

In the current study we evaluate the effects of a sequential, web-mail mixed-mode design on a variety of outcomes. A total of 7,200 adults between the ages of 21 and 35 were randomly sampled from driver's license records in eight counties in Wisconsin and asked to complete a questionnaire about attitudes, beliefs, and behaviors about drinking alcohol. Our experiment explores how different modes and pre-incentives influence data quality. For the mixed-mode experiment, sample members were randomly assigned to web-mail or mail-web treatments. Respondents initially received either a postal letter containing a URL link to the survey (web-mail) or a mailed survey (mail-web) along with a pre-incentive followed by two additional contacts. For those that did not respond, the mode was then switched for two more contacts. Within each mode, respondents were randomly to 1 or 2 dollar pre-incentive condition. Our analysis examines the effects of mode and incentives on response rates, costs, survey reports about drinking behaviors, item-missing data, and nonresponse bias in a sample of young adults assigned to the four experimental groups described above (e.g., web-mail with \$2, web-mail with \$1, mail-web with \$2, and mail-web with \$1). Preliminary analyses indicate web-mail condition fared better than anticipated.

Exploring the Impact of Prepaid Cash Incentives on Multiple Indicators of Data Quality.

Rebecca Medway, Joint Program in Survey Methodology, University of Maryland (rmedway@survey.umd.edu); Roger Tourangeau, Joint Program in Survey Methodology, University of Maryland (rtourangeau@survey.umd.edu); Luciano Viera, Jr., Fors Marsh Group, LLC (lviera@forsmarshgroup.com); Scott Turner, Fors Marsh Group, LLC (sturner@forsmarshgroup.com); Sean Marsh, Fors Marsh Group, LLC (smarsh@forsmarshgroup.com)

As survey response rates continue to decline, incentives are increasingly used as a way to motivate sample members to respond. However, beyond convincing sample members to participate, incentives may also influence the way that they approach the task of completing a survey interview – thereby altering the quality of the resulting estimates. For example, incentives have been found to reduce the prevalence of item nonresponse. However, several other indicators of data quality that are widely used in survey research have not often been analyzed in incentives research. For example, little is known about the effect that incentives have on the prevalence of satisficing behaviors. Such indicators are potentially

more informative indicators of the effect of incentives on data quality because they hold respondents to a higher standard than does item nonresponse.

Using data from incentive experiments that were included in two recent surveys, this presentation will discuss the effect of prepaid cash incentives on these additional indicators of data quality. For example, the prevalence of satisficing behaviors will be compared among those who have received an incentive and those who have not received one.

Impact on Data Quality of Making Incentives Salient in Web Survey Invitations.

Chan Zhang, Institute for Social Research, University of Michigan (chanzh@isr.umich.edu)

Incentives are commonly used to boost survey response rates, although the evidence is still limited as to whether they actually reduce nonresponse bias. Besides, we speculate that incentives might also increase measurement error if they attract respondents who participate only to obtain the incentive and so are unwilling to expend much effort. We explore the impacts of incentives on data quality by manipulating its salience in email invitation letters for a Web survey sent to university staff that asks about their use of information technology for work duties. Two designs of the invitations are compared. In one condition, we make incentives salient by starting the email subject line with the incentive and putting a statement about the incentive at the beginning of the invitation letters, in bold and larger font. In the other condition, importance of the survey is emphasized. Thus, both conditions inform the staff of the incentive and the importance of the survey, but with different emphases. To assess nonresponse error, we compare response rates between the users and non-users of a campus-wide, web-based information management system, where the usage status is known for the sampled staff and related to some of the survey variables. To evaluate measurement error, we ask two questions for which we have administrative records. We also examine the indirect indicators of response quality, such as completion times and length of open-ended responses. We find that making incentives salient increases the response rate, particularly among the non-users of the system, which suggests reductions in nonresponse bias for some survey estimates. The incentive salience also affects the make-up of the respondents on their motivations for taking the survey, which we find are related to their response quality. The findings suggest a potential link between nonresponse and measurement error driven by incentives.

GLBT Persons & Couples: Perceptions, Attitudes, Behaviors.

Individual Religiosity, National Religious Context, and Attitudes toward Homosexuals: An Analysis of 40 Countries.

Peter J. Martini, University of Nevada, Reno (pmartini@unr.edu)

Attitudes toward homosexuality vary drastically across the globe, and form the focus of social conflict within societies as well. Religion, in its various components, is among the most robust predictors of attitudes toward gay men and lesbians. Most research to date uses religiosity as a single concept that measures the intensity of belief or devotion of an individual. This paper argues that religiosity is a multidimensional concept in which different dimensions have distinct effects. In particular, I argue that the importance of religion in an individual's life, and the degree to which an individual relies on religion as a means to solve life's problems, are distinct concepts and both are likely exert separate influences on attitudes toward homosexuals. Going beyond the individual level, this paper investigates contextual influences in assessing the hypothesis that the higher national religiosity, the more likely that the nation's citizens will have negative attitudes toward homosexuals. Using the fifth wave of the World Values Survey (N = 66,755) this paper uses factor analysis to differentiate between dimensions of religiosity and

regression analysis to assess the influence of individual measures of the two proposed components of religiosity, as well as national religious context, on attitudes toward homosexuals in 40 countries, with statistical controls for other potential influences. The factor analysis supports the proposition that the importance of religion in an individual's life and the use of religion to solve life's problems, are conceptually distinct from one another as components of religiosity. Findings from regression analyses support the assertion that these two components of religiosity affect attitudes toward gays and lesbians in different ways. In terms of contextual results, the findings reveal no effect of religious context net of development, providing further support to prior research on the liberalizing influence of national development on individual attitudes.

Same-sex Marriage: An Analysis of Question Wording & Context Effects.

Daniel Cox, Public Religion Research (dcox@publicreligion.org); Robert P. Jones, Public Religion Research Institute (rjones@publicreligion.org)

Getting accurate measures of public opinion on the issue of same-sex marriage has challenged pollsters in both the academy and in advocacy campaigns. Public attitudes on same-sex marriage are shifting rapidly, particularly among certain segments of the public. Moreover, on this issue people are often forced to balance competing values (i.e. equality vs. traditionalism), and support or opposition may be contingent on which value is most active at the time. And finally, there is considerable confusion among the public over the terms of the debate, such as the difference between civil marriage and religious marriage, and the consequences for private religious institutions. The lack of conceptual clarity, competing values and shifting attitudes on this issue may give question order, context and wording outsized influence.

Getting an accurate read on public attitudes about same-sex marriage is further complicated due to the different approaches adopted by pollsters to measure it. Currently, there are two different questions commonly used by pollsters to measure support for same-sex marriage: a binary favor-oppose question and a three-part question that includes a middle, civil unions option. These different question formulations elicit significantly different levels of support for same-sex marriage.

This study will rely on two important data sources: a pre- and post-election panel survey conducted shortly before and immediately after the 2010 election and a series of experiments included in omnibus surveys from November 2010 to March 2011. The panel survey design enables us to test alternative formulations of the same-sex marriage question (the 2-part versus the 3-part) on the same respondent. This will allow us to examine differential levels of support for marriage attributable solely to question wording. The omnibus experiments are designed to test how different contexts contribute to greater opposition or support to same-sex marriage.

Latent vs. Explicit Attitudes Toward Same-Sex Marriage.

David Redlawsk, Rutgers University (redlawsk@rutgers.edu); Caroline Tolbert, University of Iowa (caroline-tolbert@uiowa.edu); Allison Hamilton, University of Iowa (allison-j-hamilton@uiowa.edu)

Attitudes toward same-sex marriage have been shifting over the past few years, and while a slim majority of Americans continues to oppose same-sex marriage, opinion is more supportive of civil unions for gays. Public opinion matters since state public opinion on these issues is a significant and positive predictor of adoption of pro-same-sex relationship policies. Yet our understanding of public opinion on these issues is based almost exclusively by measuring explicit attitudes. We assume that respondents convey their true opinions and are not subject to the influence of social context in the public expression of these attitudes.

Research shows that in a climate of increasing tolerance, the accuracy of explicit questions about sensitive issues, such as race or sexuality, can be called into question. Pressures exist for some people to hide their true attitudes in order to conform with social expectations - social desirability bias.

Such bias has been well documented in racial attitudes, but not well understood in the domain of same-sex marriage. Certainly when expressing a prejudiced attitude carries no social stigma we would not expect bias. But as attitudes have shifted on the status of gay people, it is possible some prejudices have become veiled.

We examine this with a survey-based list experiment on same-sex marriage, comparing implied support to explicit support. We find that across the full sample, both implicit and explicit measures report the same support, suggesting a lack of bias in explicit attitudes. However, we find significant bias in sub-samples, such that liberals overstate support when asked explicitly and conservatives understate it. This implies that apparent ideological polarization in explicit attitudes is attenuated in implicit measures. We explore the extent to which this finding is a result of high self-monitoring, where knowing their own ideology, some respondents express conforming explicit attitudes while holding divergent unexpressed beliefs.

Comparing Health-related Behaviors of Community Health Patients Based on Sexual Orientation and Gender Identity.

Judith Bradford, The Fenway Institute (jbradford@fenwayhealth.org)

Sexual and gender minorities (SGM, including lesbian, gay, bisexual and transgender persons) are thought to experience higher rates of suicidal ideation and perhaps suicide attempts when compared to heterosexuals. Few data sets include sufficient numbers of SGM to permit comparisons with heterosexuals. Fenway Community Health Center (FCH) in Boston, a federally-qualified health center with dual mission to provide open door healthcare, regardless of ability to pay, and to address the health and related needs of sexual minorities, conducted a voluntary survey of patients at intake over a 12 months period. A one-page questionnaire assessed health-related risks and included measures of sexual orientation and gender nonconformity, including “questioning” as a response option. There were approximately 5% refusals and 3013 completions, with similar numbers of SGM and heterosexual respondents. To assess the relationships between sexual and gender identity and suicide variables, we considered three items from the survey: (1) In your lifetime, have you ever thought seriously about killing yourself? (2) In your lifetime, have you ever made a suicide attempt? (3) Select if they had been physically attacked and/or verbally attacked. Logistic regression analyses investigated whether “questioning” individuals had the greatest odds for considering suicide, but non-significant odds for attempting suicide, compared to heterosexual, bisexual, and gay and lesbian individuals. Patients who were “questioning” their sexual orientation reported rates of suicidal ideation greater than heterosexual, gay, lesbian, or bisexual individuals. However, “questioning” individuals did not report attempting suicide at rates greater than heterosexual, gay, lesbian, or bisexual people. This finding is of interest as it may identify some protective factor that is inherent in “questioning” individuals, such as being less likely to experience discrimination based on sexual orientation if their orientation is not evident by their self-report, attitudes, or behaviors.

Documenting & Tracking How Couples Meet.

William C McCreedy, Knowledge Networks (bmccreedy@knowledge networks.com); Michael J Rosenfeld, Stanford University (michael.rosenfeld@stanford.edu); Larry Osborne, Knowledge Networks (losborne@knowledge networks.com)

The National Science Foundation supported a large-sample survey using KnowledgePanel to generate information about how couples meet, with a special emphasis on GLBT couples. This presentation will discuss the findings from the baseline study and propose plans for tracking partners as these relationships change over the coming years. Of particular importance are the different meeting patterns of GLBT and hetero couples. These pattern-differentials are hypothesized to influence the persistence or disintegration of the relationships. This presentation focuses on three aspects of KN's panel that were crucial in making this study work: 1) the ability to efficiently oversample self identified GLB respondents, 2) the ability to follow respondents over time at modest cost, and thereby to study couple dissolution rates and 3) the ability to capture respondents' own stories, in their own words. The purpose of the project was to provide a national perspective on how couples in various social contexts meet and stay together or dissolve their relationships. This research is an initial foray into the longitudinal study of partnering in the contemporary era. This project has relevance for both health and sociological topics and is groundbreaking. The panelists in KnowledgePanel provide self-identification as to their sexual identity and more than 1,000 say they are GLBT. These panelists provided a national perspective on how a rare population partners and it could only be achieved with a large national probability sample that incorporated a pre-screening for sexual identity. The sample can be tracked inexpensively because they are in regular contact with Knowledge Networks while doing surveys for other projects. A crucial, and valuable, component to this project is a qualitative aspect of the work consisting of the 'meeting-up-stories' of the couples, both straight and gay. The nature of these stories becomes one of the predictors of relational permanence or dissolution.

Methods to Improve Online Surveys.

A Meta-Analysis of Experiments Manipulating Progress Indicators in Online Surveys

Mario Callegaro, Google (callegaro@google.com); Yongwei Yang, Gallup inc.

(yongwei_yang@gallup.com); Ana Villar, Stanford University (villar@stanford.edu)

A proportion of respondents who start answer online surveys do not complete the survey. Drop-off rates could be of concern if those who do not finish the survey are systematically different from those who complete the survey. Hence, researchers have tried to implement strategies that reduce the likelihood that a respondent will prematurely abandon the survey process. One of these strategies is the inclusion of progress indicators in the survey. However, there is no consensus in the literature regarding its effect on survey drop-off rates.

In this meta-analysis we analyzed 20+ randomized experiments that compared drop-off rates of an experimental group who completed an online survey where a progress bar was shown, to drop-off rates of a control group to whom the progress bar was not shown. In all the studies drop-offs were defined as any respondent who did not fully complete the survey. Three types of bars were analyzed that differ in terms of their pace throughout the survey: a) linear or constant, b) fast first then slow, and c) slow first then fast.

Random effects analysis was used to compute odds ratios (OR) for each study. Because the dependent variable was drop-off rate, an OR greater than 1 indicates that the progress bar group had a higher drop-off rate while an OR lower than 1 indicates that the progress bar group had a lower drop-off rate.

Preliminary results suggest that, contrary to widespread expectations, the progress indicator does not help reduce drop-off rates. Both the constant and slow-to-fast progress bar are in fact increasing the chances of drop-off in comparison to no progress bar.

To our knowledge this is the first meta-analysis study on the topic. Additional literature search will be performed and we are awaiting from some authors to send us data to add to the study.

Interactive interventions in web surveys can increase response accuracy.

Frederick G. Conrad, University of Michigan (fconrad@isr.umich.edu)

Interactive web questionnaires promise to improve survey measurement relative to more static modes, whether online or paper. By designing questionnaires that react to respondent actions, it may be possible to promote behavior that leads to more accurate responses. We are investigating one type of interactivity, namely giving feedback to respondents (“speeders”) when they answer so fast they cannot realistically have read the question let alone thought about the answer (“You seem to have responded very quickly. Please be sure you have given the question sufficient thought to provide an accurate answer.”). In prior research (reported at AAPOR, 2009) we observed that some speeders answered questions about quantities (e.g., “Overall, how many overnight trips have you taken in the PAST 2 YEARS?”) more slowly when they were prompted once or twice; hardcore speeders never slowed down. We were encouraged by the slowdown for at least some respondents but could not be sure it reflected higher quality data. In the current research (n=2565) we explored the relationship between response time and quality by prompting speeders on simple numeracy items for which we could determine response accuracy (e.g., “If the chance of getting a disease is 10%, how many people out of 100 would be expected to get the disease: 1, 10 or 20?”). Prompting slowed some speeders AND increased their response accuracy, suggesting that in the earlier studies the slowdown also improved data quality. The benefits of prompting seemed to persist beyond the intervention itself as it reduced straightlining on later grid questions. Psychologically, interactive interventions seem to work differently than simple instructions. We randomly asked half of the respondents to commit to answering conscientiously; this slowed their responses but did not increase accuracy or interact with prompting. Overall, the current work illustrates the promise of interactive interventions for improving online measurement.

Qualified Success: Effects of Response Format on Requalification for Recontact Studies

Randall K. Thomas, ICF International (randall.k.thomas@gmail.com)

In attempting to locate respondents who have experienced a particular event, surveys often have two parts – a screening section that qualifies individuals and a follow-up section to gather more detailed information from those qualified. In some studies these two sections can be separated by weeks or months, often requiring rescreening to ensure that respondents meet the qualification criteria. How we re-screen has the potential to significantly affect the proportion of people requalifying for the study. This web-based survey project recontacted screened respondents and asked about their purchases of a number of products within the past 3 years using one of three randomly assigned response formats – a yes-no grid (YNG), a multiple response format (MRF - ‘select all’), or a combination grid (CG) that allowed respondents to indicate self, other, self and other, or no one had purchased the product. For both self and household members, we found that the MRF had the lowest requalification rate while the CG showed the highest requalification. We will further present how response patterns (e.g. satisficing and non-differentiation) and need for cognition affect the effects of response formats.

To Link or Not To Link: Why and When Respondents Give Consent.

Marcel Das, CentERdata, Tilburg University, The Netherlands (das@uvt.nl)

The objective of this study is to identify optimal wording for persuading respondents to consent to linking survey data with administrative records, to remove unfounded fear or distrust, but also to ensure that respondents understand what they are agreeing to. The study adds to the limited knowledge on how best to word consent statements and differs from earlier studies in the sense that survey data from the online LISS panel are used for linking.

The LISS panel collects data on a wide range of topics and for many disciplines in the social sciences. The panel is based on a probability sample of Dutch households drawn from the population register by Statistics Netherlands. Households that could not otherwise participate are provided with a computer and Internet connection.

Since 2008, surveys are conducted on a monthly basis. This has resulted in a huge amount of data. A substantial expansion of research opportunities could be achieved by linking survey data to administrative data from national registers. Statistics Netherlands provides a remote access facility for that purpose. Very strict rules apply for using this facility. But still, even in case of confidentiality assurances, respondents may not want their records to be combined with administrative sources. Linking needs consent.

Our study consists of two parts. First, we conducted a qualitative study to explore reactions to a consent statement among a limited number of subjects. Second, based on the lessons learned from the qualitative study, we conducted an experiment on a small part of the LISS panel, in which short and long(er) consent statements were presented to randomly selected groups, either by e-mail or letter. Consent rates for each group are compared. This presentation reports on results from both the qualitative and quantitative testing of informed consent statements.

Measuring User Satisfaction in the Lab: Questionnaire Mode, Physical Location, and Social Presence Concerns.

Matt Jans, U.S. Census Bureau, Center for Survey Measurement (matthew.e.jans@census.gov); Jennifer C. Romano Bergstrom, U.S. Census Bureau, Center for Survey Measurement (jennifer.romano@census.gov); Kathleen T Ashenfelter, U.S. Census Bureau, Center for Survey Measurement (kathleen.t.ashenfelter@census.gov); Jon Krosnick, Stanford University, Department of Communication (krosnick@stanford.edu)

Usability testing and user satisfaction are cornerstones of electronic survey development. There are no clear recommendations regarding how self-reported user satisfaction questions (e.g., Chin, Diehl, & Norman, 1988) should be implemented. Human-computer interaction research and theory suggest that the method matters, and that overly-positive ratings are likely when they are collected on the same computer that previously ran the software being evaluated by the user (Nass & Moon, 2000). We use a 2x2 between-subjects randomized experiment to evaluate the effect of two components of social presence on user satisfaction ratings: data collection mode (paper-and-pencil v. computerized, where the computer used for satisfaction ratings is sometimes the computer that ran the software that was usability tested), and physical proximity of the respondent to the computer on which the usability test was conducted (i.e., in the same room as the usability test or a different room). We present data from about 200 respondents, across 5 usability tests of websites and web surveys. Our initial analyses do not support past research findings that participants rate satisfaction highest when answering on the computer that was used for the usability test. We find the lowest ratings when the satisfaction questionnaire was completed on paper in the same room as the usability test (social presence on the proximity dimension, but not the mode dimension). Individual satisfaction scale items asking about information arrangement, clarity, and navigation within the website or survey showed this difference. This difference was also found in a simple summative satisfaction scale. Our findings bring into question whether the robust social presence findings of Nass and colleagues and their theoretical implications (e.g., ethopoeia) are of concern in all human-computer interaction settings, and may provide guidance for usability testers' decisions about how and where to collect user satisfaction ratings.

Public Opinion and the Economy.

The Public Opinion Deficit: Global Reactions to the Economic Crisis.

Gideon Skinner, Ipsos MORI Social Research Institute (gideon.skinner@ipsos.com)

The global economic crisis of the last few years has created great uncertainty around the world, for financial markets, business, politicians, society, and – not least – for public opinion. There has even been uncertainty over whether it really is a global crisis: one of the more interesting points to come out of the recent G20 summit is the extent to which what is seen as a ‘global’ crisis in the West is actually viewed in parts of Asia as a ‘North Atlantic’ crisis (and certainly our data points to a wide variation from rapidly developing economies such as China and India down to much more worry in the US, Japan and much of Europe).

This presentation from the Ipsos Social Research Institute explores the trends in public opinion and the economy from before the crisis began to the present day. It compares economic confidence in different countries across the world and how this matches to actual economic performance: the countries that have been worst hit, those that are showing signs of recovery, and those that sailed through relatively unscathed. We look at the implications for politicians at the ballot box, and for policy makers, for whom conflicting public opinion doesn’t make it any easier to balance the need to create growth while dealing with some eye-watering levels of public debt. In particular, a European analysis comparing the size of national debt and acceptance of actions to reduce that debt gives us some indication of perception gaps in public opinion, and where - if anywhere - the ‘debt-deniers’ can be found.

Public Opinion on the Economy in Three Countries.

Trevor Tompson (ttompson@ap.org); Jennifer Agiesta, The Associated Press (jagiesta@ap.org)

The economic recession was truly a global event, though there were significant differences country by country in its impact. In this paper we present the results of surveys conducted by The Associated Press and the GfK Group in three countries -- the United States, Vietnam and Turkey -- that we use as case studies to compare and contrast the differential impact of the recession across countries. The surveys include extensive AP-GfK telephone polls in the United States since 2008, plus individual face-to-face surveys in Vietnam and Turkey in 2010. The survey data reveal differing perspectives, both forward and backward-looking, on the impact of the global recession. Americans quickly turned negative about their feelings about the economy and became downright gloomy about the future. Americans believe now that the life will be worse for future generations than it was for them, a remarkable shift in perspective from a population that is normally eternally optimistic. By contrast, people in the hyper-developing Vietnam say they’ve experienced economic progress in the past five years, and remain bullish about the future direction of their own economy and personal finances, though there are significant concerns about unemployment. And people in Turkey also have tremendous concern about the high rate of unemployment, which is higher than the highest rate hit in the U.S., but have mixed feelings about the future direction of their country’s economy over the medium and long term.

A New View on the Economy? Distinguishing between Important Issues and Important Problems.

Christopher Wlezien, Temple University (wlezien@temple.edu)

To measure the importance of political issues, scholars traditionally have relied on a survey question that asks about the “most important problem” (MIP) facing the nation. Increasingly scholars are relying on a variant that asks about the “most important issue” (MII). While we have learned quite a lot about what MIP

captures over time, we know little about MII. Using newly-compiled data from the UK, this paper examines differences in the two items and their dynamics across 18 issues/problems, including the economy. The results of our analyses reveal that MII responses are strikingly similar to MIP responses. While they may be slightly closer to a valid indicator of issue importance, MII responses mostly reflect variation in individuals' assessments of problem status. An effective measure of issue importance remains elusive.

The "Easterlin Paradox": A Spurious Correlation?

Thomas Petersen (tpetersen@ifd-allensbach.de)

According to the "Easterlin paradox," named after the American economic scientist, and substantiated by a meta analysis of 30 surveys in 19 countries, there may be a clear relationship between income and subjective well-being within each country, but no correlation of this kind on the international level: people from rich countries did not report a higher degree of subjective well-being than people from poorer countries. Furthermore, in countries in which the average income of the population is rising, there is no comparable increase in the subjective well-being of the population at the same time. It has been suspected that the decisive predictor may not be the absolute level of income, but the distance between the personal and the average income.

In 2009, the Allensbach Institute conducted a representative survey on the material orientation of the German population. This survey contained a number of questions about different aspects of material wealth and some indicators of subjective well-being. The results of this study show that like in earlier studies, there is a strong correlation between income and subjective well-being, but this correlation disappears when other variables are added. It seems that a kind of "subjective income" is the decisive variable here: The subjective well-being is dependent from the impression that the personal income is justified, that it is satisfying, and that there is better than a few years ago. If these variables are considered in the analysis, the absolute income level drops out of the calculation.

Survey Context & Socialization Among Language Minorities.

All About Us or Part of a Sub-group: The Effect of Survey Context on Response Among Immigrant Respondents.

Angela Fontes, NORC at the University of Chicago (fontes-angela@norc.org)

The collection of accurate financial data is critical to the understanding of financial security. For immigrant groups, who are frequently underrepresented in surveys and more likely to be at-risk economically, the accuracy of financial data is even more important. Moreover, data on the financial status and public program participation of immigrant families may be used in to inform federal and state policies and influence public opinion. In addition to the standard methodological challenges associated with surveying hard-to-reach populations, potential immigrant respondents may be particularly sensitive to confidentiality concerns due to fear surrounding documentation status or simply general mistrust of government institutions. Immigrant respondents may also feel increased pressure to provide socially desirable answers as indicators of acculturation. Survey context may influence an immigrant respondent's willingness to provide accurate information by establishing the comparison group the respondent will be measured against, as well creating a perception of safety and confidentiality (or lack thereof). This paper examines the effect of survey context by comparing responses on a number of financial indicators between two surveys with different contexts: the 2007 Survey of Income and Program Participation (a national survey that includes a representative sample of immigrants), and the New Immigrant Survey (a survey exclusively examining immigrants). Data from these surveys are compared to determine if there

are differences in the reporting of ownership and allocation to financial assets. Analyses are conducted to determine the existence and extent of differences in the reporting of several financial indicators including income, retirement account (IRA, Keogh) ownership and transaction account (savings, checking) ownership. Finally, a measure of federally funded program participation, the receipt and amount received in food stamps is compared between immigrant respondents in the two surveys.

Hispanic Attitudes Toward Immigration and the Language of the Interview.

Kate Kenski, Dept. of Communication, University of Arizona (kkenski@email.arizona.edu); Marisa A. Enriquez, University of Arizona (enriqueem@email.arizona.edu)

Nonresponse is a concern to survey researchers because nonrespondents may differ systematically from respondents (Groves and Couper, 1998; Lavrakas, 1993). Language barriers are a significant contributor to survey nonresponse (Hu, Link, & Mokdad, 2010). Given that the percentage of households speaking a language other than English has been on the rise (Shin & Bruno, 2003), failing to offer interviews in Spanish is potentially problematic if those who answer questions in Spanish differ from those who answer questions in English. Surveys providing Spanish and English interviewing options have revealed demographic differences between Hispanics interviewed in Spanish compared to those interviewed in English (Stepanikova & Cook, 2004). Using data from the 2008 National Annenberg Election Survey (n=716), the relationships between Spanish vs. English interviewing and attitudes toward three immigration topics are examined. Results indicate that Hispanic respondents gave significantly different responses to questions about providing a path to citizenship for some illegal aliens, building a fence along part of the US border with Mexico, and allowing illegal immigrants to obtain driver's licenses depending on whether the interviews were conducted in Spanish or English. These differences persisted in multivariate models, which controlled for several demographic characteristics and ideology. The implications of these findings are discussed.

Symbols and Layout: Unique Issues in Chinese Translation of Self-Administered Survey Forms.

M. Mandy Sha, RTI International (msha@rti.org); Yuling Pan, U.S. Census Bureau (yuling.pan@census.gov)

Translating a self-administered survey form into Chinese requires considerations in both linguistic and sociocultural appropriateness, as well as issues related to form navigation. Prior studies show that symbols used to mark a response differ between English and Chinese writing practice (Pan et al, 2005) and that the translation of commonly-used Yes/No answer categories does not always make logical sense in Chinese (Pan et al, 2009). However, there is no literature that documents these issues in one source or provides evidence from a large number of in-language cases. This paper aims to investigate and identify issues unique to the Chinese language and writing practice and pinpoint the challenges these issues will present to survey data if unaddressed.

This study uses data from 84 Chinese language cognitive interviews from a project conducted by the U.S. Census Bureau. This project pretests the Chinese translation of the American Community Survey (ACS) Language Assistance Guide (LAG). We found that Chinese language rules and writing practice affect form navigation. The symbol X that is commonly used to denote "yes" in a response means "no" according to Chinese writing practice. Because Chinese characters are not derived from Roman letters, it is not possible to "translate" the instructional function of the ALL CAPS text. Chinese characters can also affect the layout of the form because Chinese is usually more compact than English, requiring fewer words to relay the same meaning. In addition, the Chinese translation of Yes/No answer categories without incorporating the verb used in the question can sound awkward and hard to understand for Chinese respondents. We designed some solutions to mitigate the effect of these issues and will discuss

the feasibility of our recommendations. This paper merits attention of questionnaire designers of self-administered survey forms that provide language assistance to respondents with limited English proficiency.

Factors Contributing to Differences in Reported Party Identification Across Polls in 2010.

Elizabeth Hamel, Kaiser Family Foundation (lizh@kff.org); Bianca DiJulio, Kaiser Family Foundation (biancad@kff.org); Sarah Cho, Kaiser Family Foundation (sarahc@kff.org); Claudia Deane, Kaiser Family Foundation (claudiad@kff.org); Mollyann Brodie, Kaiser Family Foundation (mollyb@kff.org); Larry Hugick, Princeton Survey Research Associates International (larry.hugick@psra.com); Jonathan Best, Princeton Survey Research Associates International (jonathan.best@psra.com)

Political party identification is a fluid concept and the share of people identifying as a Democrat, Republican or something else can change from year-to-year, month-to-month or with political cycles. However, among polls conducted at similar times, the proportions of the public reporting allegiance to a certain party tend to be similar. In 2010, the Kaiser Family Foundation Health Tracking surveys consistently reported a higher percentage of Democrats relative to Republicans than other national surveys conducted at similar times. Data from the Kaiser Health Tracking surveys and other national polls were analyzed to determine the factors that may affect party ID balance. Some of the hypotheses examined include the influence of cell phone sampling, field house effects, Spanish language interviews compared to surveys conducted in English only, survey topic and question wording, among other factors. Given that party identification is related to public opinion on a number of topics, understanding what may lead to over- or under-representation of a political party at any given point in time and any subsequent impact on reported results is particularly important.

Using Address Based Sampling with Challenging Populations.

Young Cell Phone Only or Mature Survey Responding Household? Exploring the Use of Geo-demographic Clusters in Sample Stratification.

Daniel Estersohn, Arbitron Inc (Dan.Estersohn@arbitron.com); Robin Gentry, Arbitron Inc (Robin.Gentry@arbitron.com); Vrinda Nair, Arbitron Inc (Vrinda.Nair@arbitron.com); Kelly Dixon, Arbitron Inc (Kelly.Dixon@arbitron.com)

The move to an address-based sample frame provides opportunities for new differential procedures by small geographies, such as Census block groups or ZIP+4s. These treatments can be administered in advance, without knowing the specific characteristics of the selected respondent. The treatments are based upon the demographic characteristics of the specific geography. Such treatments can include stratification by over-sampling and under-sampling or differential incentives. Arbitron has been investigating the potential of using geo-demographic clusters, such as Claritas' PRIZM, for use in stratifying and/or incentivizing sample. The PRIZM system allows for easy clustering of block groups by demographic characteristics. The system also makes it possible to generalize results from one market to many markets.

To date Arbitron has performed two separate analyses and one experiment using custom PRIZM groups. The first analysis used the PRIZM-level response rates of 887,000 mail-out cell phone screeners and block group-level Census demographic data. Cell phone-only households tend to be younger than landline households and younger households tend to be less responsive to surveys than older households. So clusters that had both low response rates and high young adult (18-34) penetration were

identified. Likewise, clusters that had both high response rates and high older adult (55+) penetration were identified. The low response clusters represent about 13% of U.S. households. The high response clusters represent about 14% of U.S. households. A test of differential premiums, for these clusters, was performed.

A second analysis was done to determine if the landline sample might benefit from over-sampling the young clusters and under-sampling the older clusters. This produced a different set of clusters than the national study. A model was created to simulate the results of over-and under-sampling. The paper will report the results of the two analyses and as well as the incentive experiment.

Seasonal Yield Variation and Related Response Patterns in Address-Based Mail Samples.

Charles A. DiSogra, Knowledge Networks, Inc. (cdisogra@knowledgenetworks.com); Erlina Hendarwan, Knowledge Networks, Inc. (ehendarwan@knowledgenetworks.com)

The probability-based online KnowledgePanel® run by Knowledge Networks maintains size and representative diversity through ongoing mail recruitment. In 2010, four national address-based samples, drawn from the USPS Delivery Sequence File frame, were fielded in two replicate waves per sample from end of January through November. Each sample, drawn without replacement, consisted of two strata with 40% of the sample targeting Hispanic census blocks in one stratum and 60% allocated to a second stratum of all remaining blocks.

All mailings were virtually identical in size, strategy, timing and sample design. Every wave or mail drop had 22,500 residential addresses covering all 50 states. Each had an initial packet sent first class to Current Resident with a \$2 incentive enclosed, a reminder postcard mailed one week later to the entire sample, and a reminder letter mailed to non-responders two weeks after the postcard. All materials are in English and Spanish.

Three response modes were in effect at all times: postage-paid reply envelope, online Web site portal, and toll-free telephone number. Total yield is measured daily as response from all modes. Mode-specific yields are also tracked daily. Each response is linked with its respective mail sample cohort through a unique PIN.

Mailings were at the ends of January and February, early April and May, end of June, mid-July, and ends of August and September. Total respondent yields show variation across mailings with lowest yield occurring with the May mailing and highest with the July mailing. The low May mailing yield is consistent for reply-mail returns and telephone call-ins. Online respondents, however, demonstrate a very different pattern showing progressively higher yields from January right through the August mailing (September mailing results are in progress at this writing). Statistics on yield variation and response patterns plus plotted seasonal response curves are presented.

Sampling College Students to Improve Coverage in University Towns.

Michelle A Cantave, Arbitron, Inc (michelle.cantave@arbitron.com); Robin Gentry, Arbitron, Inc (robin.gentry@arbitron.com)

In some metropolitan areas (e.g. Ithaca, NY, State College, PA and Ann Arbor, MI) as much as 40% of the population between the ages of 18 and 24 live in college dormitories or other university-owned housing (known as university group quarters). This population usually cannot be reached through traditional RDD landline or address-based sampling approaches because these dwellings are almost entirely excluded from frames provided by commercial sample vendors. As a result, campus dwellers are

a major source of under-coverage for studies of the general population which are meant to cover respondents in both group and non-group dwellings. In December 2009, a pilot study was fielded in three metro areas with large university populations using an address-based sample of dormitory rooms and university-owned apartments. The sample consisted of 3000 surveys mailed to students randomly sampled from university directories. Two-thirds of the students were sent a diary and a media survey; the remaining students were sent a short screener survey. In March 2010, we contacted the students that returned the screener survey from the pilot test and attempted to place diaries with all of the students in the respondent's dorm room/apartment. We will present results of the pilot tests—including mail usability rates and response rates as well as characteristics of the responding students. The presentation will include lessons learned and next steps for incorporating university/college students into surveys of the general population.

Improving REACH U.S. Sampling and Operations by Using Demographic Flags.

Ying Li, NORC at the University of Chicago (li-ying@norc.org); Michael Davern, NORC at the University of Chicago (Davern-Michael@norc.org)

The Racial and Ethnic Approaches to Community Health Across the U.S. (REACH U.S.) Risk Factor Survey is a CDC-sponsored community survey that monitors the progress and achievements of 28 communities as they implement programs to eliminate health disparities among African Americans, Asian Americans, Hispanics, and Native Americans. The survey targets racial and ethnic minorities in specific geographic areas using an address-based sampling (ABS), mixed mode data collection protocol involving telephone, mail, and face-to-face interviews. Because the survey targets specific ethnic and racial minorities within each REACH community, eligibility rates for the survey vary by community and we face the challenge of finding cost effective ways to locate and interview eligible households during our sample design and operational planning phases. At the end of the second year of data collection, we purchased a set of market research demographic flags and appended the flags to our sample frame. The flags included variables such as age, gender, ethnicity, household income, number of adults, education level, and dwelling type. In this research, we explore the congruence between these market research flags and respondent-provided interview data and the operational utility of these flags. The distributions of the flags among unresolved, eligible, and ineligible cases will be examined to see if the flags can inform weighting adjustments. Preliminary results indicate that the percentage of agreement between the market research flags and respondent interview data are significantly high. Operationally, the demographic flags potentially could be used to locate cases and target the use of bi-lingual interviewers to cases most likely to be eligible and yield interviews.

Friday, May 13, 2011

10:00 a.m. - 11:30 a.m.

Concurrent Session B

2010 Mid-Term Elections: The Western Battleground States.

Panel Abstract – Western Battleground States.

Floyd Ciruli, Ciruli Associates (fciruli@aol.com); Mark DiCamillo, Field Research Corp.; Stuart Elway, Elway Research; Mike O'Neil, O'Neil Associates

Public pollsters from western states will deconstruct the 2010 mid-term election. The national mid-term wave of Republican voting swept eight Congressional seats in 6 western states to the Republican side of the aisle. Yet, the wave stopped at the Colorado River and no western Senate seats moved. Arizona, California, Colorado, Nevada, New Mexico and Washington were midterm battleground states that will be discussed. Most will be front-line states in 2012 presidential race. The behavior of western states voters in the midterm will include discussion of:

- The nationalization or resistance to nationalization of the election.
- How high profile senate races shifted from referendums to contests
- How candidates and campaigns made a difference
- Eight congressional seats in the 6 western states changed parties but no change in California, why?
- Tea Party participation, independent and Hispanic voters, impact of illegal immigration and healthcare legislation.
- The likelihood the states will be swing states in the 2012 presidential election.
- Accuracy and impact of public polls reported in the mid-term election within the Western battleground states.

Comparative Analyses – When Does ABS Matter?

A Study of Health Measures Within Cell Phone Populations: Differences and Similarities of Cell Phone Respondents Attained by RDD Versus ABS.

David Dutwin, Social Science Research Solutions (ddutwin@icrsurvey.com); Sharon Long, Shadac (slong@umn.edu); Timothy Triplett, Urban Institute (TTriplett@urban.org); Susan Sherr, SSRS (ssherr@ssrs.com)

There are presently two primary methods by which researchers can interview cell phone only (CPO) households: 1) via a dual frame telephone survey or 2) the utilization of an address-based (AB) sample. It is yet unclear, however, whether for the cell-owning population generally and CPO population specifically, which of these methods is better at attaining relatively unbiased estimates. As well, there have been a number of studies (Keeter et al 2008; Lee et al 2010) that have shown stark differences in the demographics of cell phone mostly households attained from a landline compared to a cell phone. What is as of yet unexplored is how these match up to cell mostly populations attained using an address-based design, and further, whether there are similar differences in cell mostly households attained from telephone-matched address sample compared to unmatched sample. We propose a comprehensive comparison across two health surveys both done in the state of Massachusetts; 1) The dual-telephone frame MA Health Reform Survey which calls both landline and cell phones numbers; 2) The MA Health Interview Survey which uses an AB sample to reach all households, including cell mostly and CPO households. We compare cell-owning respondents across a number of principal measures, and across a host of overlapping and/or unique samples: RDD cell phone interviews; ABS cell phone owners; RDD cell phone owners; cell mostly interviews between both frames and within frames (landline to cell phone and matched to unmatched address); and CPO interviews between both frames. Both surveys were conducted in relatively the same timeframe, with the same survey vendor, and with the same telephone call rules and attempts. The overarching goal is to determine whether there are differences on health and other estimates between the two sampling methods, and if so, where such differences are most prevalent.

National Survey of Fishing, Hunting, and Wildlife-Associated Recreation (FHWAR) Cell Phone and Debit Card Test.

Elke McLaren, U.S. Census Bureau (elke.m.mclaren@census.gov); Denise Pepe, U.S. Census Bureau (denise.p.pepe@census.gov)

The FHWAR is an address-based sample with interviews conducted by Computer-Assisted Telephone Interviewing (CATI) and Computer-Assisted Personal Interviewing (CAPI). The 2011 FHWAR budget will only fund 5,154 CAPI interviews out of approximately 34,064 households with no available telephone number. Decreasing our sample from 34,064 to 5,154 introduces additional variance in our survey data.

The purpose of the Cell Phone and Debit Card Test is to research alternative survey designs that could increase the number of CATI interviews while reducing the variance associated with conducting fewer CAPI interviews.

This test contains three panels of 500 households with no available telephone number. Each advance letter, although unique to the panel, requests that a respondent calls the telephone center to conduct a telephone interview. The panels are described below.

1. Advance Letter and Cell Phone – For Communication Between the Household and the Census Bureau
2. Advance Letter with a \$25 Debit Card Incentive – PIN Received Upon Completion of Interview
3. Advance Letter Only – Wording Impresses that CATI interviews Save Government Dollars.

We will compare the results of our test panels to our production CATI and CAPI samples to evaluate whether any of the test options are viable alternatives to collecting data through a personal visit when we have an address but no telephone number. If this study proves successful, it will provide an option for future surveys to decrease interviewing costs while equalizing the difference between households with a telephone number and households without a telephone number.

At the conference, the authors plan to present study logistics, preliminary results, and suggested improvements for further study.

A Direct Comparison of ABS and Telephone Sampling in a Pilot Study of Children's Health.
Mary E Losch, UNI Center for Social & Behavioral Research & Dept. of Psychology (mary.losch@uni.edu); Pete Damiano, University of Iowa Public Policy Center (peter-damiano@uiowa.edu); Jean Willard, University of Iowa Public Policy Center (jean-willard@uiowa.edu); Anne Bonsall Hoekstra, UNI Center for Social & Behavioral Research (anne.bonsallhoekstra@uni.edu); Duoc Nguyen, UNI Center for Social & Behavioral Research (duoc.nguyen@uni.edu)

Address-based sampling (ABS) is playing an increasing role in survey data collection and has been shown in a number of studies to increase coverage. However, the approach requires a longer field period and is typically more expensive than utilizing traditional telephone sample frames. Moreover, few published studies using ABS have provided information on the feasibility of using ABS for targeted population subgroups. During the spring of 2010, we conducted a direct comparison of mixed-mode (phone and web) ABS and more traditional telephone sampling (RDD and targeted landline telephone frames) in a pilot study of children's health in a Midwestern state. Results supported previous findings that ABS improves coverage. Additional comparisons of the overall costs, time, and response rates will be presented along with analyses of mode effects within the ABS data. The advantages and disadvantages of using ABS for surveys of subpopulations will be discussed.

A Comparison of Address Based Sampling and Dual Frame Sampling for National Telephone Surveys.

John Boyle, Abt SRBI Inc. (j.boyle@srbi.com); Anna Fleeman, Abt SRBI Inc. (a.fleeman@srbi.com); Courtney Kennedy, Abt SRBI Inc. (c.kennedy@srbi.com); Faith Lewis, Abt SRBI Inc. (f.lewis@srbi.com); Charles D. Shuttles, Abt SRBI Inc. (c.shuttles@srbi.com); Andy Weiss, Abt SRBI Inc. (a.weiss@srbi.com)

Address based sampling (ABS) has been proposed as an alternative for dual frame sampling in order to represent cell phone only households in telephone surveys. This study compares a national ABS sample of cell phone only households to a national sample of cell phone only households from a dual frame sample.

A sample of approximately 4,000 addresses was drawn in the first stage of the ABS sample. These addresses were electronically matched with listed telephone numbers. Approximately half of the sample could not be electronically matched to listed telephone numbers. The unmatched sample was sent a short mail questionnaire which included telephone status, demographic questions, and requested telephone number. A total of three mailings were made to obtain a completed questionnaire. Those who returned the questionnaire with telephone number were re-contacted by telephone in a ten call design over three weeks for a short telephone survey. The demographic characteristics from those who completed the mail and telephone portions of the ABS sample are compared to the same demographic questions from another national dual frame telephone survey conducted by the authors.

The study addresses three important research questions related to telephone survey sampling alternatives. First, what type of response rate can be expected from a moderate level of effort (three mailings) for an ABS survey among unmatched telephone numbers. Second, how does the household telephone status in an ABS survey compare to the known distribution of that population. Third, how does the characteristics of cell phone only households from an ABS sample compare to cell phone only households from a dual frame sample. These findings should help to inform researchers about the choices in sampling frames for telephone surveys at the national level.

Cross-National Comparative Survey Findings.

Comparing Opinions of “Clashing” Civilizations Between Western and Islamic Worlds.

Lauren A. Walton, Gallup Research Center University of Nebraska-Lincoln (lauren.walton@huskers.unl.edu); **Brian Wells**, Gallup Research Center University of Nebraska-Lincoln (bmwbrian85@gmail.com); **Ann Arthur**, Gallup Research Center University of Nebraska-Lincoln (ann@yarthur.com); **Ana Lucia Cordova-Cazar**, Gallup Research Center University of Nebraska-Lincoln (al.cordovacazar@huskers.unl.edu); **Tarek Albaghal**, Gallup Research Center University of Nebraska-Lincoln (talbaghal@gmail.com); **Allan McCutcheon**, Gallup Research Center University of Nebraska-Lincoln (amccutch@unlserve.unl.edu); **Meryem Ay**, Gallup Research Center University of Nebraska-Lincoln (aymrym@gmail.com); **Lei Wang**, Gallup Research Center University of Nebraska-Lincoln (leihuang03@hotmail.com)

Samuel Huntington’s “Clash of Civilizations” theory (1993a, 1993b, 1996) has been given substantial attention due to media attention and discussion about this idea of a ‘clash’ between Western and Islamic civilizations since the 9/11 terrorist attacks. The theory contends that, due to differences in culture, division and conflict in the post-cold war are arising between distinct civilizations, including Western and Islamic ones. Are people making assumptions about the strength of the clash of civilization occurring between the Western and Islamic world? Huntington argued that previous quantitative research on the clash of civilizations has ignored the post-cold war nature of this new basis of conflict (see e.g. Russett, O’Neal, & Cox 2000), while other studies have only used pre-9/11 data (Midlarsky 1998; Chiozza 2002). Past studies have focused on ethnicity (Gurr 1994), religious minority status (Fox 2001), intrastate conflict (Gurr 1994), and interstate conflict (Russett, O’Neal, & Cox 2000; Henderson & Tucker 2001) in order to examine the clash of civilizations.

Using the Gallup World Poll, a probability based multinational survey, this research will try to examine the clash of civilizations hypothesis by comparing relevant post-9/11 public opinion data from the United States to Pakistan, Indonesia, and Saudi Arabia. Cross-national survey responses, which directly measure perceptions of the clash between Western and Islamic civilizations, will be looked at. Specifically, this research will look at the public’s perceptions on potential and current conflict, respect, commitment to improvements in interaction, and reasons for tensions between these two groups as they relate to overall perceptions of clash. For instance, preliminary analysis indicate differences in the perceptions of respect: only 15% of American respondents believe the Muslim world respects the West, while this is the case for 73% of Saudi Arabians, 48% of Pakistanis, and 70% of Indonesians.

Is Globalization Undermining the “Clash of Civilizations”? A Test of Huntington among the Publics of Greater Asia and the Pacific.

Christian Collet (collet@icu.ac.jp)

Samuel Huntington’s influential “clash of civilizations” hypothesis Huntington (1992, 1997, 2000) has been widely debated, but empirical tests of his ideas have been rarely undertaken at the micro-level among comparative publics (e.g., Norris and Inglehart (2002), Norris and Inglehart (2004); Carlson and Listhaug (2006)). In this paper, we bring new evidence to bear, focusing on attitudes in the ‘cauldron of civilizations’: Greater Asia and the Pacific. Using pooled data from four waves of the Asia Barometer, a cross-national, multilingual survey project that includes over 32,000 respondents in 31 Asian Pacific countries (including Russia and the US), we examine the extent to which publics identify with the core states of their civilization and the factors that may influence those perceptions. We give attention to the role of globalization and Westernization (operationalized as English skill), and whether they may be subsumed by religiosity and nationalism as Huntington suggests. Our evidence affirms the fault lines between the West, China and Islam identified by Huntington as the areas of potential conflict. At the same

time, we find important exceptions to the “clash” framework as a whole, particularly within Sinic societies and the role of “lone countries,”

Japan and Russia, in their expected alignment with China and the West, respectively. Generally, we find that attitudes toward core states are not zero-sum: peripheral state publics often see rival core states as mutually good influences. Specifying an individual-level model to account for competing effects, interactions and standard control factors, we find that, *ceteris paribus*, increases in an individual’s overall exposure to foreign cultures leads to a more positive assessment of the United States, China and Russia. Greater facility in English also contributes to better perceptions of US influence among Asian Muslims (and Japanese impressions of China), although the same does not hold for Chinese, where foreign exposure, standard

Measuring Tolerance, Religiosity and National Identity in the South Caucasus.

Robia Charles (Robia@crrccenters.org)

How is it possible to measure tolerance, religiosity and national identity in multinational and hard-to-reach populations? This paper examines these measures in Georgia, Armenia and Azerbaijan using data from the Caucasus Barometer (CB)—an annual nationwide survey in these three countries that is conducted by the Caucasus Research Resources Centers (CRRRC). The paper first gives an overview of the survey design employed by the CB, including questionnaire design with translations, sampling, face-to-face household interviews and analysis. The paper also discusses specific measures of tolerance, religiosity and national identity that are included in the questionnaire. To measure tolerance, the CB asks about attitudes towards various nationalities or ethnic groups with respect to doing business, having friendship or marriage. Results show that Georgians have the highest approval ratings for Italians, Americans and Russians and have the highest disapproval for Chinese and Kurds in all three categories. In Armenia, Russians and Americans garner the most approval and in Azerbaijan, Turks and Russians have the highest approval. Azerbaijanis and Turks have the highest disapproval rating in Armenia, while Armenians have the highest disapproval rating in neighboring Azerbaijan. With regard to religiosity, CB data shows that religious institutions are one of the most trusted institutions in the region. While there is a strong association between ethnicity and religious affiliation throughout the South Caucasus, a high level of religious importance in one’s daily life coincides with a low rate of religious attendance in all three countries. This indicates that religion is an important element of national and cultural identity in the Caucasus. Additionally, the CB asks if people feel that their way of life needs to be protected from European, Russian and American influences and finds that over 50% of the population in each country says yes (Armenian approval of Russian influence is an exception).

Public Opinion Perceptions in the Israeli-Palestinian Intractable Conflict.

Jacob Shamir (jshamir@mscc.huji.ac.il)

The proposed paper is based on a unique project, the Joint Israeli Palestinian Poll (JIPP). For over ten years since the year 2000, Prof. Jacob Shamir from the Hebrew University of Jerusalem and Prof. Khalil Shikaki from the Palestinian Center for Policy and Survey Research (PSR), Ramallah, are conducting joint surveys among Palestinians and Israelis on the conflict and the various attempts to break the cycle of violence. This project provides a rare opportunity where public opinion of two sides to an intractable conflict is studied in depth in a coordinated and joint manner. In these surveys both Israelis and Palestinians are asked questions about their attitudes, perceptions and expectations regarding a large range of issues which are at the heart of the conflict.

At times public opinion is accurately perceived, but at other times we observe pluralistic ignorance at varying levels, both within one’s society and with regard to the opponent’s public opinion. We propose to study jointly the two publics’ pluralistic ignorance, and look for the explanation of it in the information

environment. We juxtapose the two major approaches to public opinion dynamics: the approach which focuses on social cues and information (Noelle-Neumann 1993), and the approach which emphasizes substantive information carried in events and political developments (Page and Shapiro 1992). While we consider both approaches relevant to public opinion stability and change, building on previous work (Shamir and Shamir 1997, 2000), we expect that the major source of pluralistic ignorance and its disintegration will be a function of substantive information, similar for both sides to the conflict.

Variation in Perceptions of Democracy: The Influence of Procedural Fairness.

Matthew Miles (sandramaxim@gmail.com)

In the current World Values Survey, the mean response for the level of democracy in China is nearly the same as the mean "level of democracy" score in the US. This paper examines evidence from 51 countries using data from the World Values Survey, Freedom House, and the Worldwide Governance Index. Using a multi-level model to test the effect of national institutions on individual perceptions of democracy, this paper finds that governments that control corruption, protect the rule of law, and have effective government services, also have publics that perceive their systems as democratic; regardless of regime type.

Methodological Briefs: Survey Participation & Response Bias.

Driving Respondents to the Web: Experimental Trial of Benefit Appeals and Impacts on Survey Completion.

Danna L. Moore, Social and Economic Sciences Research Center (moored@wsu.edu)

This study empirically examines the effects of two different prominently displayed appeals in combination with set confidentiality assurances and other survey statements in letters on completion rates for agricultural screening questionnaires. The experiment was carried out on a sample of 13,000 farm units, using person name and business entity (farm name) addressed pre-notification letters and mail back questionnaires. While web surveys have become increasingly more common in research, this methodology has not been thoroughly investigated for agricultural populations and for USDA sponsored surveys. Whether web can be an effective methodology for reducing more expensive in-person interviewing and land visits for determining eligibility for the Agricultural Census is evaluated. A critical question in this research was whether a pre-notification letter asking respondents to complete a short 5 minute web screening questionnaire about their agricultural involvement could be effective. Also of interest was whether benefit appeal differences would hold across mixed mode (web, mail, and telephone) administration and whether systematic differences would contribute to coverage differences and response effects.

Early Response Bias in a Northern Colorado Community Health Survey.

Michael Dorssom, Wyoming Survey & Analysis Center (WYSAC) (mdorssom@uwyo.edu)

One problematic issue facing survey research is designing data collection strategies to most effectively balance the desire for an acceptable overall response rate with data gathering efficiency. In some cases, efforts to keep data collection timelines as short as possible may not yield the best final sample obtainable. The research question addressed in this methodological brief asks, "Do systematic differences exist between those who respond earlier and those who respond later in the data collection process?" The presented research shows that early response bias exists, and that extending the data

collection period may increase responses obtained from harder to reach groups, such as young adults and Hispanics.

In fall, 2010 the Wyoming Survey & Analysis Center conducted a large scale general population community health survey in Larimer County, CO. The data collection mode (utilizing Don Dillman's Tailored Design Method) was a four-sequence straight mail effort including first a pre-notification letter, a second mailing which included the questionnaire and cash incentive, a reminder postcard, and finally a reminder/replacement questionnaire sent as the fourth step in the sequence. All told, data collection was open for eight weeks. Pre-notice and cover letters were written in both English and Spanish, with Spanish questionnaires available upon request.

Data from this project show significant differences between those responding early in the collection effort (before the second questionnaire mailing) and those responding later in the collection effort (after the second questionnaire mailing) in key demographic variables, namely home ownership, race, educational attainment, and age. An inverse relationship between response latency and age, as well as a greater response from Hispanics later in the data collection effort, indicate that extending the collection period and focusing more effort on late responders should have a positive effect on obtaining data from these typically difficult groups.

Exploring the Attitudes, Styles & Policy Preferences of Top Legislative Leaders in the 50 States.
Debbie Borie-Holtz, Rutgers University (borieholtz@aol.com)

A national panel study of top legislative leaders in the states presents unique methodological concerns when studying both small and elite populations. In order to address concerns about coverage, response rate and generalizability in a 2010 census of leaders serving from 1997 through 2010, a material culture and document review was completed prior to going into the field. In addition to reducing coverage issues to less than 1 percent, the data were coded by key demographic, electoral and governmental service factors. This dataset was later used to conduct a Chi-square test of no difference in order to examine if those who responded were different from those who did not respond and in turn, help assess if the survey results were generalizable. To increase the response rate, the census was conducted in at least three modes in two phases; mailings were timed and forwarded to dual addresses when available; the cover was designed to maximize appeal; and personal "pitch" letters were sent with the replacement questionnaire. The customized "pitch" letters relied on information culled from the document review. The response rate for those who received limited follow-up was 34%; 48% for those receiving 3-4 follow-ups, and 62% for those receiving a replacement questionnaire with a personalized "pitch" letter. This paper describes the contact and follow-up approaches used to target a small group of political elites (n=393) and highlights the implementation strategies believed to have contributed to the overall success of the study.

How to Survey All 14 000 Swedish Local Political Representatives And Get 10 000 Responses.
Mikael Gilljam, Dept of political science, University of Gothenburg (mikael.gilljam@pol.gu.se); Donald Granberg, University of Missouri (GranbergD@missouri.edu); Bengt Holm, Detector (bengt.holm@detector.se); Mikael Persson, Department of political science, University of Gothenburg (mikael.persson@pol.gu.se)

The paper reports the survey practice from a unique study covering all local and regional political representatives in Sweden To our knowledge there are no previous survey conducted in any other country targeting all local political representatives. In this paper, we first report our successful multi-mode design that combines web-survey and paper questionnaires to maximize the response rate. We then

compare the response patterns in the results from the web-survey and the paper questionnaires via post to evaluate whether the data collecting method affect the responses.

The political representatives were first contacted via e-mail and invited to fill in the questionnaire in our online web-survey. Already during the first day, we had managed to achieve a response rate of over 10 percent. Thereafter, we approached the non-responding representatives several times, via e-mail and then via telephone up to ten times). Finally, each and every one of those who lacked an e-mail address were targeted and offered a paper questionnaire via post. All in all, after four and a half month of fieldwork, we had achieved nearly 10 000 responses and a response rate over 70.2 percent.

This multi-mode design employing a combination of web-survey and paper questionnaire gave us the opportunity to test how the data collecting method affect the responses. We are therefore able to present comparisons of the different response patterns observed as regards item non-response and the response scale variance in the web and paper questionnaires respectively. In the web-survey result we find that non-item response is gradually increasing as respondents make progress in the questionnaire, while the non-responses in the paper questionnaire stem from questions on controversial and private issues. As regards the respondents use of scale end-points and variance response patterns we find significant differences between the two modes used.

Can Information as Part of a Pre-Notice Help to Improve Response Rates in Telephone Surveys?
Alicia Catherine Tomaszczyk, York University (atomaszc@yorku.ca)

Gaining respondent compliance to participate in a survey is a major problem for survey researchers as large numbers of non-responders can lead to unrepresentative samples and inaccurate results. This study examines whether the inclusion of a pre-notice for a telephone survey in the form of an information flyer designed to persuade individuals to adopt positive attitudes toward the survey topic, along with an advanced letter, will help to improve response rates. Data for the study were obtained from a wind energy survey that was administered using telephone interviews. In the survey, a random sample of Waterloo and Cambridge, Ontario, area residents were asked questions about their impressions of wind energy and the construction of wind turbines in Waterloo. Analysis of the effect of type of pre-contact (flyer/no flyer) on the survey response behavior of the respondents showed that respondents who received a flyer were as likely to respond to a survey request as respondents who did not receive the flyer ($p > 0.05$). To investigate possible explanations for this null result, analyses of items that measured respondent interest in the survey topic, and of demographic measures, were conducted. These analyses indicate that the sample consisted of respondents high in survey topic interest. Thus, it is possible that high interest respondents were not affected by the information flyer because they were already knowledgeable about the topic. Future research should examine individuals with low and high interest in the survey topic to clarify the effect of an information flyer on interest.

Pick up the Phone! An Experiment using Conversational Answering Machine Messages.

Jennifer Vanicek, NORC at the University of Chicago (vanicek-jennifer@norc.org); Ashley Amaya, NORC at the University of Chicago (amaya-ashley@norc.org); Edward Sipulski, NORC at the University of Chicago (sipulski-edward@norc.org)

Declining response rates pose a challenge to survey research, and answering machines contribute to falling response rates. With the ability to screen calls, potential respondents may choose not to participate in a study after hearing answering machine messages that are scripted and therefore sound recorded and impersonal. More personal-sounding, unscripted messages may entice respondents to participate in the survey, leading to higher resolution and response rates.

The Racial and Ethnic Approaches to Community Health Across the U.S. (REACH U.S.) Risk Factor Survey, a project sponsored by the Centers for Disease Control and Prevention to eliminate health disparities among racial and ethnic minority populations, conducted an experiment to test the effects of conversational answering machine scripts on resolution and completion rates. A random half of the telephone sample received conversational answering machine messages. For these cases, interviewers were provided with bullet point text outlining necessary information to include in the message (organization name, study topic, and toll-free number), and they were instructed to leave an unscripted message. Text displays for cases in the control group were fully-scripted answering machine messages.

This paper will outline the experimental design, describe texts and training materials used, and present preliminary results from data collection. Analysis will focus on key indicators, such as resolution, screener completion, and member interview completion rates, as well as operational-related measures such as the number and length of dials and number of days in the field. The analysis will be conducted for the overall sample as well as controlling for potential subgroup effects including language and/or community.

Methodological Issues in Questionnaire Design.

Research Based On Satisficing Theory: A Systematic Review of Methods and Results.

Caroline Roberts, FORS, University of Lausanne (caroline.roberts@unil.ch); Emily Gilbert, University of Essex (eegilb@essex.ac.uk); Nick Allum, University of Essex (nallum@essex.ac.uk)

In 1987, Krosnick and Alwin published an article about response order effects in survey measurement in which they presented a cognitive theory for why some respondents might exhibit such effects based on Herbert Simon's (1957) concept of 'satisficing'. The approach was later elaborated in an article by Krosnick (1991) to account for a range of other response effects often observed in attitudinal data, attributing them to respondents shortcutting cognitive processes necessary for reporting answers accurately (Tourangeau, 1984). In the two decades since, Krosnick's article has become one of the most frequently cited in the field of survey methodology, and satisficing theory has become a popular framework for investigating the occurrence of measurement errors associated with the response process. Despite its popularity, however, there has been considerable variation in the methods used in applications of the theory, and comparatively few studies that have explicitly aimed to test it. Furthermore, relatively mixed empirical evidence suggests that the theory may hold true for certain types of response effects in some settings and not for others. In this paper, we present the results of a systematic review of research that has adopted the satisficing framework, the aim of which has been to document the variety of methods researchers have used to construct indicators of different forms of satisficing (including acquiescence bias, non-differentiation, 'no opinion' responding, and primacy and recency) and to evaluate the empirical evidence generated in relation to the theory's central claims. Based on our

analysis, we draw conclusions about the validity of the theory in relation to different types of response effect, and make recommendations regarding optimal methods for future research into survey response effects.

Complete Satisficing in Surveys: An Exploratory Investigation.

Jon A Krosnick, Stanford University (krosnick@stanford.edu); Bo MacInnis, Stanford University (bo@macinnis.org)

Satisficing, cognitive compromise in optimally responding to survey questions, has remained a public opinion research topic with enduring interests since its conceptualization by Simon (1957) and elaborate theorization by Krosnick (1991). The research on satisficing has centered around two notions of satisficing—weak satisficing and strong satisficing—based on the absence of or compromise in the four stages in optimized response as delineated in Tourangeau and Rasinski (1988): fully interpret the meaning of each question, search memory for relevant information, process and integrate retrieved information into summary judgment, and report judgment. Weak satisficing and strong satisficing are defined as the second and third stages of the response process being incomplete and/or biased, and skipped altogether, respectively. In this study, we propose a new and stronger form of satisficing, “complete satisficing”, where the first step of fully interpreting the question is compromised or omitted. We first provide an innovative method to devise questions that allow the detection of complete satisficing, proposing that such questions should possess the following design characteristics: a) questions are devoid of the need of information retrieval and integration, b) questions are free of content domains requiring no domain-specific knowledge or interests in question interpretation, c) questions for which there are correct answers and the selection of the correct answer solely depends on the question comprehension, and d) questions are of relatively low cognitive demand, bearing much resemblance to instructions. Second, we assess the statistical properties of questions in which complete satisficing might be manifested, including reliability and validity tests. Third, we investigate the potential instigators of complete satisficing, and its relationships to common representations of weak and strong satisficing in a variety of survey questions. Finally, the potential impact of complete satisficing on data quality and implications for better survey designs are discussed.

Generalizing What We Know about Respondents Who “Don’t Know”.

Rebekah Young, The Pennsylvania State University (rly116@psu.edu)

From researchers debating the inclusion of filter questions at the survey design stage to secondary data users grappling with non-substantive responses in statistical analysis, the “don’t know” (DK) response option is a survey research dilemma. Addressing this problem requires an understanding of the relationship between non-substantive response and individual demographics, question characteristics, and survey context. Unfortunately, a great deal of what we know about DK responses has not been generalized across topical domains in nationally representative data. In fact, most research has dealt with political topics; as a result, findings may not be easily generalized to other types of survey data, particularly for special populations. Examinations of DK responses in non-political question contexts have often occurred using convenience sample data that are not nationally representative. These limitations suggest that what we “know” about the types of respondents who provide DK answers may not be as well understood as the literature suggests.

In this paper I will generalize what is known about DK responses across topical domains from multiple sets of nationally representative data. Using individual factors, survey question factors, and survey context factors, I will employ a multilevel (hierarchical) model to identify predictors of DK responses. This approach offers the opportunity to separately estimate individual and contextual effects. An exploration of

individual, question, and survey context factors is consistent with recent research highlighting the importance of understanding how these factors interact in the assurance of survey quality (Couper, 2008; Wagner, 2010). A better understanding of who uses the DK option, and under what circumstances, will enable us to optimize the use of this option in questionnaire design and analysis.

Does Mentioning “Some People” And “Other People” In An Attitude Question Improve Measurement Quality?

David Scott Yeager, Stanford University (dyeager@stanford.edu); Jon Krosnick, Stanford University (krosnick@stanford.edu)

Researchers often measure attitudes and beliefs using “some/other” questions (“Some people think that ... but Other people think that...”) instead of asking simpler “direct” questions. This is done to, presumably, decrease social desirability or acquiescence response bias by communicating that either response is normative. The present study tested the alternative hypothesis that implicitly telling respondents that the public’s opinion is split 50/50 on the issue would be confusing for some respondents, thus reducing the quality of their data. Meta-analyses of thirteen original experiments embedded in national surveys of adults provided no evidence that the some/other form improves response validity. Direct questions yielded more valid reports than did some/other questions when using a conversationally natural response order (which was found to be more valid than the conversationally unnatural response order). Because some/other questions produce less valid measurements and involve more words than direct questions - so they involve more cognitive burden for respondents - the direct form using a conversationally conventional response order seems preferable.

The Impacts of the Zone of Ambivalence: Considering the Impact of Ignoring the Measurement of Apathy, Indecision and Lack of Information.

Don Levy, Siena Research Institute (dlevy@siena.edu)

Pollsters that measure public opinion towards current political and social issues exist in a complex relationship with politicians, the public and the press. Often times, politicians and the press consume and disseminate public opinion in ways that minimize the degree to which the public is not informed on issues and have not developed a firmly held position on those issues. The public, from whom opinion was initially measured, receive the statement of their opinion as fact in which nearly all are said to either support or oppose the issue at hand.

This paper describes research that sought to determine the degree to which the public either does not care or has not become sufficiently informed to express an opinion while juxtaposing that ‘zone of ambivalence’ to the position that the same sample expressed through more typical forced choice questioning. Two samples of over 500 respondents each were queried on a series of national and local issues. In both cases the samples were given both an opportunity to respond to a forced choice version of each issue as well as one of two different scaled response options, one that was rationally based and one that was emotionally based. In both cases but even more so with the emotionally based response scale, the percentage of respondents for every issue that held firm and considered opinions dropped precipitously. The zone of ambivalence ranged from 17 to 54 percent. Moreover, this research shows that respondents tend, through cluster analysis, to fall into one of three groups – those that are informed and hold views on most issues, those that select a few issues to care about, follow and consider and those that reside within the zone of ambivalence on nearly every issue.

Mixed-Mode Methods of Data Collection.

Testing Mail Notification Strategies for an Internet Response Option in the American Community Survey (ACS).

Mary Frances E. Zelenak, U.S. Census Bureau (mary.frances.zelenak@census.gov); Mary C. Davis, U.S. Census Bureau (mary.c.davis@census.gov); Rachel Horwitz, U.S. Census Bureau (rachel.t.horwitz@census.gov); Megha Joshipura, U.S. Census Bureau (megha.p.joshipura@census.gov); Michelle Ruiter, U.S. Census Bureau (michelle.l.ruiter@census.gov); Jennifer Guarino Tancreto, U.S. Census Bureau (jennifer.guarino.tancreto@census.gov)

Internet use has become more common over the last decade as people use it for everyday activities such as shopping, financial transactions, gathering information, and general communication. In the survey world, declining response rates as well as the benefits of using an automated mode have inspired survey organizations to investigate the use of the Internet to collect data. Currently, the Census Bureau collects ACS data using three modes: mailout/mailback of a paper questionnaire, Computer-Assisted Telephone Interview (CATI), and Computer-Assisted Personal Interview (CAPI). The CATI and CAPI modes are nonresponse follow-up operations. In April 2011, the Census Bureau will conduct a test to evaluate the feasibility of providing a fourth response mode - an Internet response option - to addresses selected for the ACS. The main objective of this test is to determine the best way to present the Internet response mode in the ACS mailing pieces to maximize self response. In the 2011 ACS Internet Test, the Census Bureau will test different notification strategies using modified versions of the ACS mailing materials – letters, postcards, and questionnaires. The materials were updated to reflect variations of choice and push strategies. This paper will discuss the notification strategies and present preliminary findings from the test.

Employing Mixed-Mode Methodology for the National Teacher Survey on Children: An Analysis of Response Rates & Cost.

Kristin Fuller, RTI International (kristinf@rti.org); Keith Smith, RTI International (kesmith@rti.org); Orin Day, RTI International (oday@rti.org); Sarah Harris, RTI International (harris@rti.org); Jennifer Wallin, RTI International (wallin@rti.org)

Researchers often employ mixed-mode designs as a method to improve survey response rates and to conserve costs. However, the literature related to the success of these designs, most notably web/mail designs shows mixed results. This paper compares the effects of a mixed-mode design (web/mail) versus a single mode (mail only) on response rates for a survey of teachers.

Our analyses are based on an experimental design implemented on the National Teacher Survey on Children, a study designed to collect data from teachers on the health and well-being of children. Teachers of Kindergarten to 12th grade students are asked to complete a short questionnaire that focuses on the child's school achievements, development, behavior in school, social skills, relationships with classmates and teachers, and participation in school activities. For this experiment, teachers were randomly selected into one of two groups. The first group was given the choice of completing the survey by web or mail, while the second group was only offered the mail survey. Both groups received a small monetary incentive, as well as a series of reminder mailings and follow-up calls to increase participation.

Preliminary analyses suggest that providing teachers with a choice of response mode does not increase overall response rates. Further analyses will consider disparate teacher characteristics and how grade level and class size may have influenced web and mail responders. In addition, a cost analysis will be performed to determine whether any cost savings were achieved between the two groups. This paper will

provide insight into the efficacy of a web/mail mixed-mode design with a national sample of K-12 teachers.

Using Predicting Spanish Preference to Target Bilingual Mailings in a Mail Survey with Telephone Follow-up: A Randomized Experiment.

Marc Elliott, RAND Corporation (elliott@rand.org); Amelia Haviland, RAND Corporation (haviland@rand.org); David J. Klein, Children's Hospital of Boston, Harvard Medical School (david.klein@childrens.harvard.edu); Carol Edwards, RAND Corporation (carola@rand.org)

Response rates for Spanish-speaking individuals in population surveys are often low when budgets do not permit bilingual mailings to all. The Medicare CAHPS Survey of healthcare experiences targets >600,000 individuals annually via two mailings and telephone follow-up. Because administrative data lack reliable indicators of language preference, beneficiaries outside Puerto Rico currently receive only English mailed surveys and Spanish-language completes occur almost entirely via the more expensive telephone mode.

We used surname and address information to generate predicted probabilities of Spanish-language preference (97% predictive concordance) for a subset of 167,000 selected individuals. Of the 10,000 (6%) with the highest predicted probabilities of Spanish preference, half (control) received the usual English-only mailings and half (intervention) bilingual mailings.

Mail response rate for this target population was higher for intervention (28.7%) than control (23.9%), but much lower than for the general population (46.6%); $p < 0.0001$ for all. Phone response rates among mail nonrespondents were similar in intervention and control (15.8% vs. 15.7%, $p = 0.60$), resulting in a net increase in response of 12% (4.2 percentage points). 53.1% of intervention mail responses were in Spanish, as were 77.3% (76.1%) of intervention (control) telephone responses. In net, evidence suggests that targeted bilingual mailings- (1) caused 6% of those who would not have responded to respond by mail and (2) caused 43% of those who would have responded in English to respond in Spanish, possibly improving data quality by allowing use of the preferred language. Subanalyses that exclude the bottom 40% of predicted Spanish preference from the randomized sample found increases in response rates 1.5x as large as those reported here, a higher proportion of responses in Spanish, and lower control response rates. Thus targeted bilingual mailing of mixed mode surveys using commonly available surname and address information can cost-effectively increase representation of this underrepresented group.

Emerging Methodologies for Capturing Long Distance Travel in Household Travel Surveys.

Julie L Paasche, NuStats (jpaasche@nustats.com)

Transportation planners create sophisticated travel demand models to forecast travel in their regions built upon a foundation of data from regional households. Traditional household travel surveys involve a two-phase survey approach via multiple modes of data collection (CATI, mail and CAWI): households are recruited to participate in the survey and asked to keep track of all travel by each member for 24 hours. Survey materials are sent to the household containing diary booklets for each person to record travel and travel information is retrieved from the household after the assigned travel day. An important component of the travel demand models is long distance travel; however, because of the low incidence of long distance travel, it is difficult to capture and measure as only a very small percentage of households will travel long distances (50 miles or more from home) during their assigned travel period. Thus, transportation planners must augment the long distance travel data and various emerging methodologies are being explored and utilized. This paper examines the challenges of collecting long distance travel

data including identifying the appropriate survey method, assessing added respondent burden, operational details, and issues of the available sample recruitment approaches, such as video license capture, targeted sampling, or extending the travel period beyond the typical 24-hour travel diary duration. The Front Range Travel Counts Survey ultimately captured retrospective long distance travel via mixed modes of data collection (mail-only after collection of the household travel data versus collecting the long distance travel in the two weeks preceding the assigned travel day and as a part of the household travel survey). A statistical analysis of the long distance travel data collected in the 24-hour travel diary survey and the companion long-distance data captured via the two different data collection methods will be presented.

Mobile Phone Mode Effects at Event Based Sampling.

Daniel E Williams, Opinionology (dwilliams@opinionology.com); Laura Haber, Opinionology (lhaber@opinionology.com)

The Tour of Utah (TOU) has been described as “America’s Toughest Stage Race.” High elevation, challenging climbs, diverse terrain, and breathtaking scenery make the race compelling for both participating cyclists and spectators. From its roots as a local event the Tour of Utah is growing rapidly in international appeal. The rapid growth of this event has created the need for understanding the spectator audience. We used mobile technology to survey these spectators and get immediate onsite feedback.

Mobile phone surveying presents challenges in mode choice with smart phones providing mobile web access, SMS text messaging and regular phone options such as recorded voice interviews (IVR) or regular phone interviews. Question format limitations, survey length limitations, and coverage issues have prevented existing research methods from being ported to mobile interviewing. The potential benefits of using mobile phones for at event surveying cannot be ignored. Mobile interviewing provided immediate response during the Tour of Utah and allowed us to test mode preference of the multiple response channels. This means that respondents can choose to participate in a way that they are already comfortable with.

We gave respondents the option of SMS text messaging, a mobile web format or an interactive voice recording as potential ways of responding to the survey. We explore the preferences of individuals and the corresponding potential mode effects.

Overcoming Challenges in Sensitive Topic Health Surveys.

An App a Day Could Keep the Doctor Away: Quantifying the Use of Health and Prevention Related Smartphone Apps Among a National Sample of iPhone Users.

Trent D Buskirk, Saint Louis University (tbuskirk@slu.edu); Mark Gaynor, Saint Louis University (mgaynor@slu.edu)

Need to figure out how many steps you walked yesterday while carrying your iPhone? “There’s an app for that.” App-enabled Smartphone devices currently represent nearly 10% of all types of mobile phones being used by the nearly 276.6 million subscribers in the United States (CTIA, 2009). Emphasis on using apps for translation/dissemination/prevention in the public health setting is just beginning to emerge. And while research studies reporting development and testing of specific health/wellness related Smartphone apps are emerging slowly, the number of apps in the public sector is growing rapidly (i.e. the number of iPhone apps is reported to be in excess of 85,000). This boom in development sparks several important user-related questions including: (a) are Smartphone users regularly downloading and using such apps

as intended?; (b) are Smartphone app users sharing information collected via these apps with researchers, medical professionals or other healthcare providers?; (c) what is the shelf life (i.e. time saliency) of such apps?

This paper will present the results from an app-user survey administered to a national sample of iPhone users that includes questions about types of apps used, frequency of use, money spent on purchasing apps and experiences using the most common health/wellness/prevention apps. The survey will also inquire about users' attitudes and receptivity to the idea of sharing information collected via health related apps with healthcare providers and researchers.

To take full advantage of the Smartphone's multiple mode capabilities, an experiment was implemented that randomly assigned a fraction of participants to CAPI completion of the survey via voice on the iPhone and assigned the remaining participants to completion of a web-based (app-like) version of the survey intended to be completed via the iPhone web browser. Comparisons in the experimental paradata such as number of completes, completion times and item nonresponse will be also be presented.

Respondent Willingness to Grant Record Access or Provide Identifying Information in a Survey.
Jenna A Fulton, Joint Program in Survey Methodology, UMD (jfulton@survey.umd.edu); Roger Tourangeau, Joint Program in Survey Methodology, UMD (rtourang@survey.umd.edu)

More surveys are requesting respondents' permission to link survey responses to their administrative data. Access to these records requires respondent consent; in many cases, this also requires personally-identifying linking information such as an SSN. Evidence from several national surveys shows that the public is becoming less willing to consent to these requests. There is potential for bias in linked data sets if some respondents fail to consent, particularly if those who consent are not representative of the entire sample. A better understanding of what identifiers and records respondents find prohibitively sensitive can be used to inform and improve consent requests included in future surveys to increase consent rates, decrease data collection costs, and reduce respondent burden. This presentation will investigate influences on respondent consent through questions included on an RDD survey of 2000 respondents sponsored by the Census Bureau and done as part of the 2010 JPSM Practicum. A battery of questions in this survey elicited respondents' perceptions of the sensitivity of requests for different personal identifiers used for linkage (e.g., SSN, email address, Medicare number), and their willingness to provide these identifiers in a survey. In addition, items evaluating respondents' willingness to grant access to different personal records were included (e.g., tax, employment, and Medicare records). This presentation will describe the kinds of personal information respondents are and are not comfortable providing, and the kinds of personal records that they are and are not comfortable granting access to in a survey context, as well as what influences their decision to consent to these requests. This research will explore whether some identifiers and records are universally sensitive in a survey context, if the perceived sensitivity varies by person, identifier, or record type, or if some people are generally hypersensitive to these requests.

Interview Privacy and Social Desirability Effects in Cross-Cultural Survey Research: The World Mental Health Survey Experience.

Zeina Mneimneh, Institute for Social Research-University of Michigan (zeinam@umich.edu)

To encourage accurate responses, many surveys that address sensitive topics require that the interview take place in a private setting so that the interviewer is the only individual to hear the respondent's answers. Maintaining such privacy is always a challenge but becomes more so in a cross-national study where there are wide cultural variations in the norms and patterns of social relations and economic and social development that could influence the ability to maintain interview privacy. This presentation focuses on the investigation of interview privacy measures across different cultures in a cross-national survey, the World Mental Health Survey Initiative (WMH). The WMH has conducted nationally or regionally representative surveys in 30 countries around the globe. These surveys have been coordinated by one central body and have used comparable methods including comparable probability sampling methods, the same data collection instrument (Composite International Diagnostic Interview 3.0), and similar interviewer training content and techniques. These surveys are designed to measure the prevalence and correlates of mental health disorders as well as impairment and patterns of help-seeking. Using the WMH data, variations in interview privacy across different cultures are explored. These include measures of who was present during the interview and the duration of this presence. The impact of interview privacy on reporting sensitive mental health conditions and social outcomes is then investigated and compared. Outcomes of sensitive nature include suicidal behaviors, rating of relationship with partner and children, and partner physical violence. Interaction effects between interview privacy and level of respondent's social desirability (measured through an adapted Marlowe Crowne Scale) are also investigated and compared across cultures. The presentation ends with practical strategies for addressing interview privacy across cultures.

The Effects of Neighborhood Level Characteristics on the Reliability and Validity of Self-reports of Health Data.

Jennifer Benoit-Bryan, University of Illinois, Chicago (jbenoi2@uic.edu); Timothy P Johnson, University of Illinois, Chicago (timj@uic.edu); Geon Lee, University of Illinois, Chicago (glee29@uic.edu)

There is considerable research now available that has investigated the reliability and validity of self-reported health behaviors. Little research, however, has examined how neighborhood context may influence the quality of these self reports. This study evaluates the relationship between several neighborhood characteristics and survey response quality using the nationally representative Add Health survey. Specifically, we employ waves three and four of the Add Health study, which includes a sample size of more than 15,000 respondents. We focus on neighborhood characteristics that can be constructed using data predominantly drawn from the U.S. Census, including neighborhood segregation, social disorganization, linguistic isolation, civic participation, race and ethnic composition, unemployment and socio-economic status. We examine the effects of these constructs on multiple indicators of self report reliability and validity, including test-retest concordance, self report concordance with biological assays, and the recanting of socially sensitive behaviors. Among the variables examined for test-retest concordance are reported age of first use of tobacco and alcohol, and first sexual experience. Self report concordance with biologic assays compares reported recent tobacco use with serum cotinine. Recanting is examined for lifetime reports of tobacco and alcohol use, first sexual experience, and HIV status. We employ hierarchical linear modeling (HLM) to examine the effects of neighborhood characteristics on these indicators of self-report reliability and validity after controlling for individual level demographic characteristics also known to predict the quality of self-reports.

Accuracy of Self-reported Condom Use Assessed by the Semen Y-Chromosome Biomarker.
Janet Rosenbaum, University of Maryland (janet.rosenbaum@gmail.com)

Many adolescents use condoms inconsistently but may over-report condom use. Biomarkers offer a potential solution to the over-report, allowing interventions to target adolescents at risk from inconsistent condom use. This study uses semen Y-chromosome biomarker from a vaginal swab, sensitive for 14 days post-coitus with specificity of 92%. Respondents are 715 sexually active African-American adolescent women ages 15-21, not trying to become pregnant, participating in a safe sex intervention in urban Atlanta in 2002-04. Respondents were surveyed by ACASI at baseline, 6 and 12 months and tested for sexually transmitted infections, pregnancy, and Y-chromosome DNA. Among the 29% of respondents who claimed perfect condom use in the past 14 days at wave 1, 30% tested positive for Y-chromosome, so they are suspected over-reporters of condom use; 20% and 15% of respondents are suspected of over-reporting condom use at waves 2 and 3, respectively. Suspected over-reporters of condom use were more likely to be pregnant at the following wave than any other category of condom users, including adolescents reporting never using condoms in the past 60 days: 19% of wave 1 suspected over-reporters were pregnant at wave 2 versus 14% of never-users, and 16% of wave 2 suspected over-reporters were pregnant at wave 3 versus 12% of never-users (Fisher's exact $p=0.001$, $p=0.08$). Suspected over-reporters were more likely to be pregnant than any other category of contraceptive users at last sex, including those reporting no contraception (Fisher's exact $p=0.02$, $p=0.05$). Suspected over-reporters reported fewer episodes of coitus, and few condom never-users reported using other contraceptive methods, so the higher pregnancy rate among suspected over-reporters is puzzling. Many adolescent women in a reproductive health intervention are suspected to have over-reported condom use. Suspected over-reporters may be at higher pregnancy risk than other adolescents, including adolescents who report never using condoms.

Public Perception of Hispanics in the U.S. – Attitudes & Impact.

A Before and After Comparison Investigating the Effects on Hispanics of the April 23, 2010, Signing of the Arizona Immigration Bill.

Paul J Lavrakas, Independent Consultant (pjlavrakas@hughes.net); Trevor N Tompson, The Associated Press (TTompson@ap.org)

A national survey of 1,521 Hispanics living in the United States was conducted by NORC University of Chicago for a consortium of researchers from The Associated Press, Univision, The Nielsen Company, The Bill and Melinda Gates Foundation, and Stanford University. Data were gathered from one head of house from a nationally representative sample of Hispanic households in the United States from March 11, 2010, through June 2, 2010. The designated sample of Hispanic households was provided by The Nielsen Company and taken from their national enumeration survey of Hispanic households in the U.S. Data were gathered via mail, the Internet, or the telephone. The AAPOR RR3 was 44%. In terms of language, 53% of the completions were done via the English questionnaire and 47% were done via the Spanish questionnaire. The survey questionnaire gathered data about Hispanics' experience with and aspirations for their own educational attainment and that of their children, their ethnic-identity, matters pertaining to household financial well-being and health, media usage, political attitudes, and a variety of demographic and other background characteristics. In the middle of the field period for this survey, the so-called Arizona Immigration Law was signed by Arizona Governor Jan Brewer on April 23, 2010. At that stage of the survey's field period, 50.5% of the Hispanics had completed the questionnaire. Thus, 49.5% of the Hispanic completions occurred after the law was signed. Our paper will report findings about various attitudes, behaviors, and behavioral intentions reported by the Hispanics who provided data

before the Arizona law was signed versus those Hispanics who provided data after the law was signed. Of note, multivariate analyses using a host of variables as predictors indicated no meaningful demographic differences between the two groups of Hispanic respondents.

Undocumented or Illegal? A Content Analysis of the Arizona Immigration Law Issues in Leading National Newspapers during 2010.

Thomas B Christie, University of Texas-Arlington (christie@uta.edu); Andrew M. Clark, University of Texas at Arlington (amclark@uta.edu)

This study examines coverage of the Arizona immigration law (Arizona Senate Bills 1070 and 2650) in leading national newspapers during January – November 2, 2010. This study uses the agenda-setting and agenda-building traditions of mass communication research as a foundation for understanding the possible influence of a state issue on national public opinion in the early stages of a major general election. This study builds upon earlier studies of mass media coverage of major public policy issues and their influences on public opinion formation. It examines the nature and extent of this coverage in light of the dominant news coverage of the major public opinion issues of jobs and the national economy. Key research questions driving this study were:

1. What were the components of the Arizona Immigration law issue that received coverage by three major national newspapers leading up to the 2010 general election?
2. How did coverage of the immigration issues raised by the Arizona law compare with issues noted in public opinion polling?
3. What are the differences in coverage of the Arizona immigration law among the three major national newspapers examined?

The study begins with the introduction of Arizona Senate Bill 1070 in January 2010 and concludes Nov. 2, 2010—the date of the U.S. general election. This timeframe allows for an analysis of immigration issues associated with the law and coverage of the major legal and political developments up until the date of the general election. Results reveal major differences in the nature of coverage of The Wall Street Journal, The New York Times, and Washington Post. An analysis of the coverage of these issues compares them to the public's ranking of issues in major public opinion polls and reveals how these issues were covered over the period leading up to the general election.

Anti-Hispanic Racism and Immigration Attitudes.

Wendy Gross, Stanford University (wtgross@stanford.edu); Jon A Krosnick, Stanford University (krosnick@stanford.edu); Trevor Thompson, Associated Press (ttompson@ap.org)

Racial divisions have shaped American society since colonization. In the last half-century, social science researchers have proposed a variety of theories about the origins of racism (defined here as the dislike of a racial group), including the impact of psychological, sociological, and political factors on these opinions. From this body of work, we have learned a great deal about racism, and these findings have been used in attempts to alleviate racial discrimination (e.g., Brown vs. Board of Education, 347 US 483 (1954); Dubofsky 1969; Massey and Denton 1993; Bowen and Bok 1998; Hurtado 1999).

A great deal of the research on racism has focused on the black-white divide. However, a host of societal changes has made evident that this dichotomy is no longer the only politically relevant racial division in the United States. Hispanics have been the largest minority group in the United States since 2000, and the growth of this ethnic group has outpaced that of every other racial or ethnic group (U.S. Bureau of the Census 1990, 2000, 2008; ACS 2008; Johnson and Lichter 2008). This ethnic group is diverse in its political opinions, but its size is not yet reflected in voter rolls (Shaw, de la Garza, and Lee 2000; Jackson

2003; Alvarez and Garcia Bedolla 2003; Kelly and Kelly 2005; Pantoja, Ramirez, and Segura 2001), making Hispanics a “sleeping giant” in American politics.

The changing ethnic and racial composition of American society makes obvious an exciting opportunity for the academic literature on racism – the exploration of anti-Hispanic racism and its connection to opinions on public policies. In this paper, we develop measures of implicit (AMP) and symbolic anti-Hispanic racism and explore the sources and consequences of anti-Hispanic racism among non-Hispanic whites.

Exploring Public Opinion of U.S. Hispanics About Education.

Trevor Tompson, The Associated Press (ttompson@ap.org); Jennifer Agiesta, The Associated Press (jagiesta@ap.org)

While Hispanic enrollment in K-12 schools is surging, the percentage of Hispanics who go on to enroll in college and complete an undergraduate degree remains low, even compared with other minority groups. This paper will explore the findings of a rarely done survey of the Hispanic population which had a special focus on measuring attitudes about education. The survey, co-sponsored by The Associated Press, Univision Communications Inc., The Nielsen Company, and Stanford University, with support from The Bill and Melinda Gates Foundation, was conducted in the spring of 2010 by NORC at the University of Chicago using a rigorous multi-mode methodology. Nearly half of the survey interviews were completed in Spanish. Among the findings: Hispanics have tremendous aspirations for their children when it comes to education. Attitudinally they see even more value in a college degree in terms of its ability to help someone get ahead in life than non-Hispanic whites, so the low level of attainment isn't due to low desire for education among this population. This survey helps reveal what barriers get in the way of most Hispanics achieving those goals. Economic barriers are by far the biggest concern. Family responsibilities are also often cited as a barrier. A high proportion of parents who speak mostly Spanish at home complain of language barriers in communicating with the schools their children attend, and far more have difficulty helping their children with their homework than parents who speak mostly English. We explore this rich dataset in depth to illuminate what is one of the key social problems for the U.S. Hispanic community.

Friday, May 13, 2011

1:45 p.m. - 3:15 p.m.

Concurrent Session C

All Things Obama.

A One-Term President? Obama's Prospect in the 2012 Election.

Helmut Norpoth, Stony Brook University (hnorpoth@notes.cc.sunysb.edu)

The staggering losses of the White House party in the 2010 midterm election, the slow pace of economic recovery, noisy opposition to the president's policies, along with middling approval ratings have raised doubts whether or not Obama may win reelection in 2012. Leaving aside these factors, I offer a forecast for 2012 that is derived from a cyclical model of U.S. presidential elections. The statistical estimation of the model includes elections as far back as 1828. The model forecast gives Obama a two-in-three chance of being reelected.

Faith in the President: How Public Perception of Barack Obama's Faith Shape Views of him and his Presidency.

Daniel Cox, Public Religion Research Institute (dcox@publicreligion.org); Robert Suls, Pew Research Center for the People and the Press (rsuls@pewresearch.org)

In August 2010, the Pew Research Center caused a stir when they found that an increasing number of Americans reported believing that President Obama was a Muslim. From March 2009 to August 2010, the number of Americans who thought Obama was Muslim increased from 11% to 18%. Further, the number who (correctly) thought he was Christian dropped 14 points, from 48% to 34%. A few months later, a survey conducted by Public Religion Research Institute (PRRI) found that a majority (51%) of the public reported believing that Obama had somewhat different or very different religious beliefs than their own. In the Pew poll, believing Obama was Muslim was highly correlated with negative views about his performance as President. In the PRRI poll, believing that Obama had different religious beliefs was highly correlated with negative evaluations of him personally.

In a country as highly religious as the U.S.—as evident in reported levels of religious attendance, prayer and belief in God—the positive correlation between views about the President's faith and his job performance or personal favorability are not surprising, but the direction of influence (is favorability being influenced by perceptions of Obama's religion or the other way around) is an important question that remains unresolved. In this paper we will use data collected by both Pew and PRRI to more thoroughly explore this relationship.

In this paper, we will begin by providing a profile of Americans who view Obama as religiously different. We will then develop regression models to control for other factors (i.e. political ideology, partisan affiliation) to provide a clearer picture of just how or how much perceptions of Obama's faith influence political judgments of him and his Presidency.

The Obama Effect vs. the Clinton Effect. An Analysis of the Youth Vote in the 1992, 1996 and 2008 Presidential Elections.

Emanuel Boussios, Nassau Community College (emanuel.boussios@ncc.edu)

The 2008 U.S. Presidential election was a voting booth landslide for candidate Barack Obama among 18-to-29 year old voters. When comparing to previous presidential elections the disparity among young voters was substantial and nearly two times the advantage presidential candidate Bill Clinton experienced in the 1992 and 1996 contests. This research seeks to explain why candidate Obama did so well among younger voters in comparison to democratic candidates in previous elections. This cross-sectional analysis studies groups by age, party identification, and political ideology, in their general likelihood to support candidates in presidential elections in the 1992, 1996 and 2008 presidential elections. This research studies two competing hypotheses to explain the popularity of Barack Obama among young voters: young voters favor candidates that reflect who they are— more liberal on political and social issues and Democratic; the young cohort voted for these candidates based on their personal characteristics and stylistic/charismatic appeal. This research finds that personality theories on voting behavior may provide the explanation for young voters' strong preference for candidate Obama in the 2008 election. In other words, young voters favored Obama less because he was a Democratic Party candidate and liberal on political and social issues, but more so of his personal characteristics and broad personal appeal. In addition, this research seeks to study whether or not Barack Obama has maintained his popularity among voters a year after his election to the presidency.

***An Examination of Partisanship During the Obama Era.* Jeffrey M. Jones, Gallup (jeff_jones@gallup.com)**

Americans' party affiliation has undergone significant change in the last few years. When President Obama took office, the Democratic advantage was the highest in decades. During his two years in office, that advantage has disappeared, mostly due to independents shifting toward the Republican Party. The paper examines the changes in partisanship both at the national and state levels.

Comparative Research on World Suffering, Extremism & Evangelicalism.

A World Suffering Metric Using the Gallup-Healthways Well-Being Index.
Ronald E Anderson, Univ. of Minnesota (rea@umn.edu)

Attempts to quantify worldwide pain and suffering have relied mostly upon health statistics such as those of the United Nations. The Gallup-Healthways surveys of well-being in 100+ countries included the Cantril Ladder scale to measure life well-being, and categorized those choosing a rung adjacent to "worse possible life" as "suffering." In the 20 richest countries, the average percent responding as suffering was 3.8% while in the 20 poorest countries, it was 20%. In several countries as many as half self-classified themselves as suffering.

This paper addresses two research questions: one is the validity of this suffering metric, and the other is what it may tell us about why suffering is not uniform across the globe.

Validity is explored by comparing the Gallup opinion indicator of suffering with pain surveys and compilations of official statistics on health. Of particular utility are such measures of the average occupancy of "acute care" beds and death rates from illnesses with a prolonged period of pain. For selective countries with ongoing war or armed conflict, it is possible to use estimates of both civilian and military injuries, deaths, and perhaps rapes as indicators by which to evaluate the utility of the suffering opinion metric.

The level of national suffering, of course, is related to measures of average wealth, but the suffering metric allows us to explore the extent to which the unequal distribution of wealth serves as a propellant to

suffering. A variety of other demographic factors, measured at the national level, will be used as controls. Modernization theory and cultural values theory will be evaluated using data from the World Values Survey. As the meaning of suffering differs greatly by religion, we explore whether there is more suffering in countries where the dominant religion emphasizes suffering as a good.

Modernizers vs. Fundamentalists: How Religion, Politics and Economics Shape Attitudes in the Muslim World.

Juliana Menasce Horowitz, Pew Research Center (jhorowitz@pewresearch.org); Neha Sahgal, Pew Research Center (nsahgal@pewforum.org)

The rift between moderate or modernizing Muslims on one end and Islamic extremists or fundamentalists on the other is a much discussed phenomena both in academic and policy circles. Under the Clinton administration and especially in the wake of 9/11 western leaders vowed to work with moderate voices in the Muslim world to combat terrorism, while the idea that moderate Islamist parties should be co opted and engaged by both domestic and western governments is popular in academic writings (for example, Esposito 1992). Yet, surprisingly, there is little understanding of what particular attitudes and values define “moderates” and “fundamentalists” in the Muslim world. Scholars studying Islamist movements have argued that movements may be “moderate” in some realms, such as their economic views, but not in others, such on their attitudes toward gender issues (Schwedler 2007).

Using public opinion data from the Pew Research Center’ Global Attitudes Project 2010 survey of Muslims in Turkey, Egypt, Jordan, Lebanon, Indonesia, Pakistan and Nigeria, we analyze the extent to which religious, political and economic attitudes influence Muslims’ identification with modernizers or fundamentalists in their countries. The data suggest that support for democracy and economic liberalization among Muslims in these countries often coexists with support for harsh punishments associated with the Shariah code and with acceptance of suicide bombing in defense of Islam. Using a logistic regression model of all seven countries, we find that self-identification as a modernizer is more strongly correlated with attitudes toward democracy and economic development than with views about Sharia law or support for suicide bombing. Still, an analysis of each country individually shows that the attitudes that define modernizers and fundamentalist vary across countries. These findings present a challenge for Western policymakers as they try to shape the attitudes of and garner support from Muslim publics.

The Predictors of Latent Support for Extremist Ideology in Pakistan.

Tarek Baghal, University of Nebraska (mtalbaghal@hotmail.com); Allan McCutcheon, University of Nebraska - Lincoln (amccutch@unlserve.unl.edu)

Pakistan is a major area of concern for policy makers and academic researchers interested in the extent of support for extremist and militant Islamic sub-national groups. Recent research has used latent class models (LCM) to estimate the population size for varying levels of support for these groups among Pakistani men (Albaghal and McCutcheon 2010). This research fit a model with four latent classes, using four indicators of support for extremist ideologies, each with three response options. The largest estimated percentage of the population belonged to the oppositional class (53.1%), followed by those leaning oppositional (26.6%), leaning supportive (16.3%), and supportive (4%) classes. Although the supportive group is the smallest class, 4% of the Pakistani male population equates to a large number at risk of becoming involved in extremist activity. Further, including age and education categorical variables into the model showed that among the young and less educated, Pakistan’s largest population segment, the supportive class made up 9% of the population. This prior research was limited in the extent covariates could be used, however, by including them directly into the LCM. Covariates had to be few in

number and partitioned into broad categories to ensure that the data did not become overly sparse and the model was estimable. The current research extends the prior research by first assigning respondents to one of the four latent classes based on their response distributions (Goodman 2007). By doing this, ordered logistic regression models can be employed which allow for more covariates to be examined without constraints of limiting categorizations. The results are suggestive as to the characteristics of those displaying varying levels of support, which can be important to both researchers and policy makers in continued efforts to understand the situation in Pakistan.

A Global Portrait of Evangelicals: A Survey of Evangelical Leaders from the Third Lausanne Congress of World Evangelization.

Carolyn L. Funk, Virginia Commonwealth University (clfunk@vcu.edu); Allison Pond, Pew Forum on Religion & Public Life (apond@pewforum.org); Timothy Samuel Shah, Georgetown University (timothyshah@mac.com)

One of the major fault lines for societal conflict in the U.S. centers on religious divides, especially between evangelical Christians and secular institutions and citizens in society. Little is known about evangelicals in other parts of the world and the ways in which societal conflict between evangelicals and other groups in those societies are similar to or potentially quite different from the conflict seen in the U.S. We report the results of a unique survey of religious leaders attending the Third Lausanne Congress of World Evangelization. The religious leaders invited to the Congress were intended to serve as a “global parliament” of evangelicals and were chosen in approximate proportion to the population size of evangelicals living in each of 13 regions around the world. The survey conducted by the Pew Forum on Religion & Public Life draws together the opinions and perspectives of these evangelical leaders from 167 countries around the globe. The survey, conducted by web and paper questionnaire from August through December 2010, included perceptions of intergroup relations, what it means to live as an evangelical in society today, potential threats to evangelicalism, views about the role of men and women in the family, and beliefs about moral and religious issues. The analysis highlights similarities and differences in the perspectives of U.S. evangelical leaders with those from other parts of the world. We also draw on Pew Forum surveys of the general population in the U.S. to compare religious beliefs of U.S. evangelical leaders with evangelicals in the general population. The findings illustrate the diversity of backgrounds and views among evangelicals around the world. These results provide a more nuanced portrait of the relationship between evangelicals and other groups in society as seen from the perspective of this group of global leaders active in the evangelical movement.

Women’s Opinion on Women’s Religious Freedom in Iran And Turkey.

Meryem Ay, Gallup Research Center, University of Nebraska-Lincoln (meryemay@huskers.unl.edu); Tarek Albaghal , Gallup Research Center, University of Nebraska-Lincoln (talbaghal@gmail.com); Allan L. McCutcheon, Gallup Research Center, University of Nebraska- Lincoln (Allan_McCutcheon@gallup.com); Ana Lucia Cordova-Cazar, Gallup Research Center, University of Nebraska- Lincoln (al.cordovacazar@huskers.unl.edu); Ann Arthur, Gallup Research Center, University of Nebraska-Lincoln (ann@yarthur.com); Lei Wang, Gallup Research Center, University of Nebraska - Lincoln (leihuang03@hotmail.com); Lauren Walton, Gallup Research Center, University of Nebraska -Lincoln (lauren.walton8008@gmail.com)

Women's rights have frequently received particular attention because they are often targets of inequality. In the Middle East, where the dominant religion is Islam, and societies tend to be patriarchic, gender related religious obligations are frequently more strictly enforced.

The present study focuses on religious freedom from the perspective of the Middle Eastern women by examining public opinion in Iran and Turkey. Both interfere in women's religious expressions: in Iran,

certain religious practices (e.g. covering one's hair) are mandated by the government; in Turkey, religious women are prevented from carrying out many of these same practices. Iran and Turkey presents an interesting comparison due to their political and demographic situations, such as the institutionalization of religion in government, the majority sect of Islam, and level of Westernization. Previous research has suggested that there is important nuance in the secular-religious debate in Turkey, particularly regarding women's issues, while women in Iran appear to be more orientated towards religious obligations (Gallup 2006, 2007).

By using the data from Gallup World Poll, the current research examines several measures of women's attitudes toward their personal and religious freedom. Preliminary results indicate that 46.5% of Turkish women feel satisfied with the amount of freedom in their lives, while this is the case for 56.7 % of Iranian women. When asked whether they would add freedom of religion to a new constitution if given a chance, 88.9% of Turkish women agreed compared to 71.8 % of Iranian women. Women who reported that religion is important for them comprise about 80% of women in both countries. These results suggest that women in Turkey may be more dissatisfied with their religious freedom than those in Iran. Additional analyses will examine possible explanations, including methodological, why women in these countries feel the way they do.

Cross-National Public Perspectives on Immigration.

Attitudes toward Immigrants in the US and Europe: Do Skills Matter?

Claudia Diehl, University of Goettingen (cdiehl@gwdg.de); Nyiri Zsolt, The German Marshall Fund of the United States (ZNyiri@gmfus.org)

While there is ample empirical evidence that people with higher levels of education have more liberal attitudes toward immigration, the mechanisms linking these two factors are less clear. Theoretical arguments suggest that varying degrees of competition for jobs and social benefits play a role in this regard but empirical support is mixed. This study focuses on what types of natives (highly or lowly skilled) are opposed to what types of migrants (highly or lowly skilled) to provide further insight into the relationship between skill level and attitudes toward immigrants. Prior research on this topic has been hampered by the fact that most surveys measure attitudes towards immigrants and immigration in general without taking the skill levels of immigrants into account.

Using survey evidence from the United States, Canada, France, Germany, the UK, Italy, the Netherlands, and Spain, we analyze the role of skill level in explaining natives' attitudes about immigration. We expect the attitudes of unskilled and skilled natives to differ more strongly when related to admitting migrants with low rather than high levels of education. Likewise, skilled and unskilled natives are expected to differ in their assessment of the economic consequences of immigration but less so when it comes to its non-economic impact. Furthermore, we assume that a worsening personal financial situation is more closely related to anti-immigrant attitudes for low skilled than for highly skilled natives because the latter are less likely to see themselves as competing with migrants in the labor market.

This research also provides a rigorous test of the link between skill levels and immigration related attitudes by analyzing it in a cross-national context. We test country effects toward immigrants based on the overall skill level of the immigrant population and on the degree of labor market protection in each of the eight countries.

Cross-Cultural Views on Immigration: Hispanics and Non-Hispanics in the Wake of Arizona's Immigration Law.

Trevor Thompson (TTompson@ap.org)

No topic sparked greater debate about cross-cultural interactions than Arizona's attempt to implement immigration laws at the state level. This paper will examine views on the bill and broader immigration policy across cultural lines. As SB1070 made its way through the legislature and the details became clear, a public debate raged. Shortly after the bill passed, The Associated Press and Univision Communications, Inc. partnered to conduct a survey of 901 Hispanics to match a general population survey of 1,002 adults, exploring views among Hispanics and the public in general on politics, discrimination, and immigration policies. The RDD telephone survey was conducted in English and in Spanish by GfK Roper Public Affairs and Media. Overall, about equal proportions of both Hispanics and non-Hispanics felt illegal immigration was a serious problem for the United States today, but they differed significantly in how the problem should be handled. Most Hispanics felt that the issue should be reserved for the federal government, while most non-Hispanics felt that state and local governments should have a role in enforcing immigration laws. Nearly all Hispanics favored a way for illegal immigrants to become citizens, while just a slim majority of non-Hispanics favored that idea. Three out of four Hispanics believe that illegal immigrants mostly contribute to society, while a majority of non-Hispanics believe that illegal immigrants are a drain on society. Eight in 10 Hispanics felt there was a significant amount of discrimination against Hispanics in the United States, compared with 6 in 10 non-Hispanics who perceived such discrimination. The AP-Univision poll also examined views on the details of the Arizona bill and the impact it would have on life for Hispanics in the state. The paper will explore the findings from this unique study on these and other topics, comparing the views of Hispanics and

Leaving Home: Current Motivations Behind Latin American Migration to the United States.

Ana Lucia CORDOVA CAZAR, Gallup Research Center, University of Nebraska-Lincoln (al.cordovacazar@huskers.unl.edu); Lei Huang, Gallup Research Center, University of Nebraska-Lincoln (leihuang03@hotmail.com); Tarek Albaghal, Gallup Research Center, University of Nebraska-Lincoln (talbaghal@gmail.com); Ann Arthur, Gallup Research Center, University of Nebraska-Lincoln (ann.arthur@huskers.unl.edu); Allan L. McCutcheon (amccutcheon1@unl.edu); Meryem Ay, Gallup Research Center, University of Nebraska-Lincoln (aymrym@gmail.com); Lauren Walton, Gallup Research Center, University of Nebraska-Lincoln (lauren.walton@huskers.unl.edu)

By 2007, Hispanics -the fastest growing minority in the United States (Marotta & Garcia, 2003)- represented almost one-half of the US foreign-born population (Grieco, 2010). Apart from Mexican-born immigrants, who make up the largest proportion of the Hispanic population in the United States (64.4%), sizable portions of migrants come from other Spanish-speaking countries from South and Central America, such as Cuba, Dominican Republic, Puerto Rico, Guatemala, Honduras, El Salvador, Colombia and Ecuador (U.S. Census Bureau, 2007).

This Latin American exodus started around the late 1950s and seems to be unstoppable. Indeed, data from the Gallup World Poll, a probability-based multinational survey carried out by the Gallup Organization, indicates that by 2008 this trend continues: more than one-third of the population of these countries express the intent to move to a different country, with the largest majority (43.54%) choosing the United States.

The political and economic strains facing these countries have been put forward as the main reasons for Latin Americans leaving their home countries (Oboler, 2005). Using the Gallup World Poll data, this paper will empirically examine whether these hypotheses explain the current migratory patterns of Latin Americans. In particular, this paper will analyze the degree to which aspects such as personal economic circumstances, expectations of thriving in life, and the perception of national political and economic conditions influence the intent of Latin Americans to either stay or leave their countries of origin.

References

- Grieco (2010). Race and Hispanic origin of the foreign-born population in the United States: 2007. ACS Report #11. U.S. Census Bureau.
- Marotta & Garcia (2003). Latinos in the US in 2000. *Hispanic Journal of Behavioral Sciences*, 25(1), 13-34.
- Oboler (2005). 50 years of South American immigration (1950-2000). *Latino Studies*, 3, 45-52.
- U.S. Census Bureau. (2007). The American community - Hispanics: 2004. U.S. Census Bureau.

Exploring the Effects of Political Ideology and U.S.-Mexico Border Proximity on American Attitudes toward Illegal Immigration.

Timothy B. Gravelle, Gallup (tim_gravelle@gallup.com)

The issue of illegal immigration to the United States from across the U.S.-Mexico border is the source of much political debate in the U.S. It is also at the root of contentious policy actions designed to counter illegal immigration such as SB1070 in the State of Arizona. This paper contributes to an understanding of American public opinion toward illegal immigration by asking two related research questions. First, how do political ideology and proximity to the U.S.-Mexico border influence attitudes toward illegal immigration? Second, does border proximity moderate the effect of political ideology on attitudes toward illegal immigration - that is, do border proximity and ideology interact? This paper draws on recent survey data from the Gallup Poll to answer these questions with a particular focus on the rated importance of the issue of illegal immigration, concern over the effects of illegal immigration on the U.S., and concern for the rights of those affected by stricter immigration laws. Methodologically, an innovation presented in this paper is the application of geocoding tools to public opinion data, thereby allowing for distance and proximity calculations to be performed.

Methods to Improve Data Quality.

Dear Diary: Improving Data Quality and Creating an Engaging Experience Through Video.

Miriam Gerver, VS Research (mgerver@gmail.com); Vivienne Segal, VS Research (vivienne@vsresearch.com)

As we move to a more technological age, using video diaries in place of paper diaries may be the wave of the future as a standard research methodology. We will examine two studies which used diaries: one was a paper diary on food consumption, and the other was a video diary on medication use. The studies were part of larger projects that used diaries as precursors to, and informed the direction of, subsequent focus groups or individual interviews. Both studies were conducted in the second half of 2010, and participants were given similar instructions about the type of information to include in their diaries. Additionally, we are beginning a study for an intimacy enhancement product comparing paper to video diaries, and may have the potential to include that research as well.

Comparing the two types of diary research, the video diaries had two clear advantages to the researchers: Firstly, the videos were uploaded by respondents immediately, so the researchers could view them and provide feedback to respondents on a timely basis if they were not following the instructions correctly, thereby minimizing measurement error. In contrast, the paper diaries were mailed back, and if there were problems with them, it only became apparent after the diaries were completed. Secondly, the video diaries elicited richer in-depth respondent entries (compared to the paper diaries), and insight from respondents, similar to the type of data that would be obtained from qualitative interviews.

We will focus on the video diary study, discussing the content of the study, the instructions provided to the respondents, lessons learned along the way, and advantages of using this method, both to the researchers and to the respondents. We will also show examples of video clips (either original clips, or simulated clips using original responses, due to privacy and confidentiality concerns).

Ensuring Data Quality: Monitoring the Accuracy and Consistency Among Telephone Interview Monitors.

Joe Baker, Mathematica Policy Research (jbaker@mathematica-mpr.com); Claudia Gentile, Mathematica Policy Research (CGentile@mathematica-mpr.com); Jason Markesich, Mathematica Policy Research (JMarkesich@mathematica-mpr.com); Shawn Marsh, Mathematica Policy Research (SMarsh@mathematica-mpr.com); Rebecca Weiner, Mathematica Policy Research (RWeiner@mathematica-mpr.com)

Although interviewer monitoring is central to survey quality assurance, little research has been devoted to examining the accuracy and consistency of monitors' behaviors. In 2010, Mathematica Policy Research conducted an exploratory study that found a high degree of accuracy and consistency among monitors' overall interview ratings. However, monitors tended to use only the middle of the rating scale and sometimes used criteria not included in the scale. We conducted a follow-up study to further examine the following questions: (1) How accurate and consistent are monitors' overall ratings of interviews? (2) What criteria do monitors use to rate interviewers? (3) How accurate and consistent are monitors in identifying nonstandardized interviewer behaviors and providing constructive feedback to interviewers?

We examined the behaviors of two groups of monitors: 3 senior monitor supervisors and 12 active monitors who evaluated 20 digitally recorded interviews from six projects. Monitors used our five-point Interviewer Rating Scale to assign an overall rating and a behavioral coding system to highlight both positive and nonstandardized interviewer behaviors. To measure accuracy and consistency, we compared monitors' ratings within and across the two groups and compared their behavior codes at the question level. To examine the extent of monitor variation within each monitor over time, we asked monitors to reevaluate three interviews from the pilot study. We conducted focus groups in which monitors discussed their decision-making processes, revealing how they applied the criteria when assigning ratings and providing feedback to interviewers. Lastly, we collected information from the interviewers about their views of and experiences with the monitoring system.

Our plan is to use the data from the study to improve our internal monitoring process, as well as to encourage others in the industry to review their own processes and conduct more in-depth research on monitoring behaviors and ways to improve survey quality assurance.

Memory Jogging Tool to Reduce Recall Bias: The Use Of A Calendar Tool To Help Store Owners Recall Retrospective Information.

TraShawna Boals, The Nielsen Company (trashawna.boals@nielsen.com); Lorelle Vanno, The Nielsen Company (lorelle.vanno@nielsen.com); Jennie Lai, The Nielsen Company (jennie.lai@nielsen.com); Michael Link, The Nielsen Company (michael.link@nielsen.com)

Recall bias is known to be problematic in retrospective reporting especially when measures are not taken to reduce it. The effective design of data collection methods or tools is critical in reducing potential recall bias. In order to reduce the recall bias of store owners in a retail panel, The Nielsen Company designed a calendar instrument for recalling retrospective purchase information. This tool collects minimal information on purchases and business related landmark events. These data are then used by the interviewer as a

memory aid to collect more detailed information in an in-person interview. We will conduct qualitative interviews with store owners and interviewers to assess the simplicity of the tool and evaluate if the tool was an effective memory jogging aid. These discoveries will help survey researchers in their development of memory aid tools for reducing respondent recall bias.

Redesigning Contact Materials for the National Survey on Drug Use and Health.

Douglas B. Currivan, RTI International (dcurrivan@rti.org)

Designing survey contact materials presents a challenge for survey practitioners to determine the content and features most likely to encourage participation among recipients. For general household surveys, limited research exists to inform decisions on whether specific text and graphics in the contact materials are likely to be effective in facilitating cooperation and avoiding refusals. As a result, designing contact materials requires survey researchers to combine relevant data, knowledge, and experience to construct effective documents.

This paper will summarize recent efforts to redesign the primary contact materials for the National Survey on Drug Use and Health (NSDUH). The NSDUH provides national, state, and substate data on substance use and mental health in the civilian, non-institutionalized population age 12 and older. Data are collected on a quarterly basis each year, with approximately 140,000 household screenings and 67,500 interviews completed annually.

Three methods were used in redesigning the advance letter envelope, the advance letter, and the question and answer (Q&A) brochure for the NSDUH. First, the researchers developed alternative versions for each of these contact materials to address potential limitations of the current materials. Second, we submitted the current and alternative versions of the contact materials for expert review and feedback. Third, we conducted 17 focus groups with members of the target population in different parts of the United States to discuss participants' reactions to the different versions of the materials.

This study will identify contributions each of the three methods made toward determining the final text and graphics for each of the three types of contact materials. We will also highlight important themes that emerged from this research, especially from the focus groups. Based on these research findings, we will discuss how various content and feature elements in contact materials are likely to facilitate cooperation in household surveys in the United States.

Logical Edits of Health Insurance Coverage in the ACS and the CPS ASEC.

Joanna M Turner, University of Minnesota - SHADAC (turn0053@umn.edu); Michel Boudreaux, University of Minnesota - SHADAC (boudr019@umn.edu)

The American Community Survey (ACS) added a question about health insurance coverage to the 2008 questionnaire, establishing itself as the only other federal survey, besides the Annual Social and Economic Supplement to the Current Population Survey (CPS ASEC), to produce annual state-level health insurance estimates for all age groups. In 2009, logical coverage edits (also called consistency edits) were added to the editing routine for Medicaid, Medicare, and TRICARE/other military coverage. The goal of logical editing is to improve the accuracy of coverage variables at the individual level by deterministically assigning coverage to people who do not report it. Rules are based on eligibility and enrollment procedures and utilize other information provided in the survey. Implementing the logical coverage edits reduced the uninsured rate from 15.1 to 14.6 percent in 2008. Lynch et al. (2010) discuss the overall impact by key subgroups. In this paper we expand on this work and discuss the logical coverage edits in relation to their impact for states.

The ACS edits were developed by the U.S. Census Bureau in consultation with an independent technical advisory group and were based on edits used in the CPS ASEC. While these edits likely reduce overall reporting errors they have never been empirically validated. The benefit of logical editing is to assign coverage to people that report no coverage, but actually have coverage. There is also a cost to editing; assigning coverage to people who are not actually covered. To explore this trade-off, we present findings from a multi-phase research project (SNACC 2008) that linked CPS ASEC data to Medicaid administrative records. While administrative records have their own source of error this provides an initial evaluation of the benefits and costs of the Medicaid portion of the CPS ASEC logical coverage edits.

Multi-Mode Comparative Data – When Does Mode Matter?

Is There a Trade-Off Between Quality and Cost? An Experiment Comparing Telephone and Face-to-Face Responses to the National Beneficiary Survey (NBS).

Matt Sloan, Mathematica Policy Research (msloan@mathematica-mpr.com); Debra Wright, Mathematica Policy Research (dwright@mathematica-mpr.com); Kirsten Barrett, Mathematica Policy Research (kbarrett@mathematica-mpr.com)

Multimode surveys offer several benefits, such as providing multiple ways to reach respondents and allowing respondents to reply in the mode most convenient for them. However, there is a risk that different data collection modes will not produce equivalent results. For example, research suggests that telephone interviewing places a greater cognitive burden on respondents than face-to-face interviewing, resulting in higher rates of acquiescence and satisficing. This effect may be more pronounced for individuals with lower cognitive ability. Multimode survey methods may also involve a cost trade-off. It is clear that in-person interviewing is significantly more expensive and time consuming than telephone interviewing. It is less clear how these single-mode cost estimates compare to the marginal cost of attempting a phone interview followed by an in-person interview.

Mathematica Policy Research conducted a study to determine whether multimode methods led to cost and quality trade-offs in the NBS, a longitudinal survey of approximately 7,000 disability beneficiaries. Sponsored by the Social Security Administration, the NBS is conducted via telephone, with face-to-face follow-up for nonrespondents. For this study, Mathematica randomly assigned NBS respondents to either telephone or face-to-face modes to examine differences in data quality and cost. Mathematica also compared the data from a single mode of interviewing with data from respondents who received both telephone and in-person outreach. In comparing the data across modes, Mathematica examined item nonresponse, the number of options checked for check-all-that-apply items, nondifferentiation between items in a series, the proportion of “agree” or “yes” responses, the length of responses to open-ended items, and the distribution of responses or means for sensitive items. In addition, Mathematica examined the mode effects by type of disability (cognitive or physical) and determined the average per-case cost to complete the survey for each type. Finally, Mathematica developed methods to reduce mode effects for disability beneficiaries.

Re-Examining the Validity of Different Survey Modes for Measuring Public Opinion in the U.S.: Findings From a 2010 Multi-Mode Comparison.

Stephen D. Ansolabehere, Harvard University (sda@gov.harvard.edu); Bernard Fraga, Harvard University (bfraga@fas.harvard.edu); Brian F Schaffner, University of Massachusetts, Amherst (b.f.schaffner@gmail.com)

In 2010, a task force commissioned by AAPOR produced a report that provided recommendations regarding how the survey research industry should approach the use of opt-in Internet panels. Their primary recommendation stated, "Researchers should avoid nonprobability online panels when one of the research objectives is to accurately estimate population values" (AAPOR 2010, 52). Yet, the AAPOR task force also noted that, "Despite the widespread use of online panels there is still a great deal that is not known with confidence" (54). Indeed, the mode study that arguably attracted the most attention in 2009 was conducted in 2004 (Yeager et al. 2009). Yet, studies conducted using data collected several years ago are unlikely to shed much light on the utility of opt-in Internet surveys today (or the increasing problems with RDD telephone polls). After all, Internet usage has increased significantly during the past several years while the use of landline telephones has declined. More importantly, the methods used to recruit panelists and generate representative samples from opt-in panels are undergoing constant innovation.

In this paper, we present data from a four-mode study carried out in 2010. National surveys were fielded at the same time over the Internet (using an opt-in Internet panel), by telephone with live interviews (using a national RDD sample of landlines and cell phones), by telephone with IVR (landline only), and by mail (using a national sample of residential addresses). Each survey utilized a nearly identical questionnaire soliciting information across a range of political and social indicators, many of which can be validated with surveys fielded by other organizations at the same time. Comparing the findings from the modes to each other and the external indicators, we demonstrate that carefully executed opt-in Internet panels are an increasingly valid way to measure public opinion in the United States.

Does Survey Mode Matter? Comparing Consumer Satisfaction Results across Internet and RDD Telephone Samples.

Barbara Everitt Bryant, University of Michigan (bryantb@umich.edu); Forrest V Morgeson, ACSI, LLC (morgeson@acsi.org); Reginald Baker, Market Strategies, International (reg.baker@marketstrategies.com)

In this paper, we utilize a multi-mode sample of customer satisfaction data and structural equation modeling (SEM) statistical techniques to investigate similarities and differences in survey responses across samples drawn using random-digit-dial telephone and internet interviewing. The large-n samples examined cover customers of multiple NAICS economic sectors, including food processing (manufacturing/nondurable goods), automobiles (manufacturing/durable goods), fast food restaurants (food services), airlines (transportation), commercial banks (finance), and wireless telephone service (information). The data come from the American Customer Satisfaction Index (ACSI), an economic indicator of customer satisfaction with the goods and services available to household consumers in the United States.

Results from the estimated ACSI structural equations suggest that while survey mode marginally impacts the mean and modal distribution of the survey responses, an effect observed in other studies, the latent variable factor loadings from the SEM measurement model and the coefficient estimates from the SEM structural model suggest considerable stability. The implications of these findings, and particularly their effect on survey mode selection in consumer survey research, are discussed.

Age Matters: Differential Mode Effects in Vulnerable Populations.

Tracy A. Keirns, UNH Survey Center (tfowler@unh.edu)

Considerable research has been conducted examining the effects of survey mode on sensitive or socially desirable responses. This research finds that respondents are more likely to provide socially desired

responses to sensitive questions when surveys are conducted by an interviewer either in-person or by telephone than in self-administered questionnaires (Kreuter, Presser, and Tourangeau 2008; Tourangeau and Yan 2007). In satisfaction surveys the same pattern emerges, respondents report more extreme satisfaction by telephone than by mail (Dillman et al. 2009). Most of this research examines mode effects in general population or student samples with little if any research on mode effects in specialized populations.

This study examines whether mode effects are likely to impact populations differently. For instance, vulnerable populations may interpret the sensitivity of questions differently due to the perception of fear that their responses would be disclosed. Furthermore, this difference may be compounded by demographic characteristics such as age. We analyze differential mode effects between severe and persistently mentally ill youth and adults on satisfaction with mental health services in a multi-mode survey (paper and telephone). Preliminary analyses of three years of surveys show significant mode effects: participants who responded by telephone were more likely to give higher satisfaction scores than those who responded to mail surveys. In addition, the mode effects were impacted by the age of the respondent. For adults, significant differences were present in the intensity of agreement to specific items, but no significant difference when items were combined into domains such as general satisfaction, access or outcomes. For youth, the mode effects were stronger in both intensity and significant differences in domains.

These results call into question the use of multi-mode methodologies for special population surveys, particularly those which survey youth.

Sampling the American Community Survey to Collect Vaccination Data: Results of a Mixed-Mode Study by Collection Mode, Household and Demographic Characteristics and Key Vaccination Indicators.

M. Christopher Stringer, U.S. Census Bureau (m.christopher.stringer@census.gov); Andrea L Piani, U.S. Census Bureau (andrea.l.piani@census.gov); Philip J Smith, NCIRD/CDC (pzs6@cdc.gov); Stacie M Greby, NCIRD/CDC (sig6@cdc.gov); James A Singleton, NCIRD/CDC (xzs8@cdc.gov)

The Centers for Disease Control and Prevention and the U.S. Census Bureau partnered to conduct a feasibility study of the use of the American Community Survey (ACS) as a sampling frame for assessing childhood vaccination coverage. A secondary, alternate sampling source was used to supplement the ACS cases in one geographic stratum in order to achieve the sample size requirements for the study. Use of the dual sample frame allowed for comparison of results by sample source within the study.

The household data collection plan incorporated a multi-mode approach, using telephone and in-person interviewing. ACS sample households identified as having age-eligible children were contacted and parents/guardians were interviewed to obtain household-reported vaccination histories as well as child, maternal, and household demographic characteristics. At the end of the interview respondents were asked for permission to contact their vaccination providers. Vaccination providers were then contacted to collect immunization data on the sample children to verify vaccination status. The provider record check also incorporated a multi-mode approach using mail, telephone, and in-person visits.

This presentation provides a preliminary evaluation of differences in survey outcomes for landline vs. non-landline households, by mode of interview, sample source, and household mover status. Unweighted distributions of child, maternal, and household demographic characteristics, as well as vaccination coverage rates for MMR and Varicella obtained from household-reported vaccination histories, will be

compared by telephone status, mode of interview, sample source, household mover status and select demographic characteristics. These findings provide additional information comparing health characteristics of landline and non-landline households in a survey with relatively high response rates.

Public Opinion on Healthcare Policy & Reform.

Public Opinion on Health Reform Post-Midterm Elections: Continuities and Contradictions.

Claudia Deane, Kaiser Family Foundation (ClaudiaD@kff.org); Mollyann Brodie, Kaiser Family Foundation (MollyB@kff.org); Sarah Cho, Kaiser Family Foundation (SarahC@kff.org)

On March 23, 2010 President Obama signed the Patient Protection and Affordable Care Act (PPACA) into law, marking the most significant government action on health care since the enactment of Medicare and Medicaid in 1965. Since its passage, the Kaiser Family Foundation has been in the field every month with an in-depth, national sample survey of Americans' views toward, and understanding of, the new health reform law, and it is on this rich base of data that this paper will draw. In this paper, we will analyze the fundamentals of public opinion on health care policy and health care reform as they existed before passage, and show how these views have translated in the post-passage era, with a particular focus on Americans' preexisting feelings toward their own health care arrangements, their varying appetites for reform, and their interest in a variety of policy solutions. We will also analyze: the malleability of opinion on health reform, the way public views of reform as a package square with views of individual provisions of the new legislation; and how partisan affiliation has become the basis of a vast public divide in views of reform. Finally, as the newly Republican-controlled House of Representatives begins their first months of work, we will examine the public's complicated views toward possible repeal of the legislation.

Representing and misrepresenting public opinion about health reform.

Gerald M Kosicki, School of Communication, The Ohio State University (kosicki.1@osu.edu)

The narrative surrounding public opinion and health care has been characterized by a number of simple errors that have underestimated support and exaggerated the opposition.

The first problem is that the issue was characterized as a simple two-way narrative. That is, as a contest between liberal and conservative solutions. It is more realistic to view the issue as a three-way contest among moderates, liberals and conservatives. Two-way narratives tended to count as opposition to the Democratic plan those who were opposed to any form of health reform with those who preferred a stronger plan.

A second error was exaggerating the opposition. The Tea Party represents a small slice of the population but received heavy media coverage for its disruptive and sometimes shocking tactics in summer 2009.

A third problem with the health care debate is that the public showed multiple signs of being severely misinformed about the issue. Even months after passage of the historic legislation substantial proportions of the public admit being confused about the issue are not aware of important provisions of the bill. Large minorities of the public continue to believe a number of health care myths.

A fourth failure was misinterpretation of the results of the Massachusetts Senate race. It is important to recognize that Massachusetts was one state in the union that already had the kind of universal health

reform envisioned in the federal law and so the local questions about the legislation turned on issues unique to Massachusetts.

The paper systematically reviews a variety of polling data related to health care reform over the 23-month period from January 2009 to November 2010. Of particular interest are the media polls, although the paper also relies heavily on data from Kaiser, Pew, New York Times, and Gallup and selected multivariate analyses of public opinion data.

Testing a Model of Public Support for Healthcare as a Government Priority.

Michael G. Elasmr, Boston University (elasmar@bu.edu)

Debates on both the national and international levels raise the issue of whether preventing, controlling and treating health risks/diseases is the responsibility of the state. This paper asks: What are the antecedents that explain the variation in people's support for healthcare as a government priority? This paper proposes a model designed to explain the variation in individuals' support of healthcare as a government priority. This model is tested using structural equation modeling techniques applied to survey data drawn from the 2007 Pew Global Attitudes Project. (The Pew Global Attitudes Project bears no responsibility for the analyses or interpretations of the data presented in this paper). The resulting "Model for Public Support for Healthcare as Government Priority" suggests the following interpretation: individuals who hold fundamental beliefs in democratic rights are more likely to become sensitive to the potential for and severity of health issues affecting their communities. When it comes to health, this sensitivity results in more support, not less support, for government intervention regarding the procurement of basic health needs. Although the impact of gender was weak, it is also possible that women more than men are likely to support the prioritizing of healthcare by government.

Measuring Health Insurance Premium Assistance and Program Participation in Social Surveys.

Dianne Rucinski, Institute for Health Research and Policy (drucin@uic.edu)

The difficulty of accurately measuring health insurance coverage is well established. Especially challenging is the measurement of publicly sponsored program participation. In the private health insurance arena, participants' lack of awareness of premium subsidies paid by employers and the "true costs" of health care animate discussions of consumer-driven health plans and health savings accounts. People with employer-sponsored coverage may not be able to provide accurate reports of premium assistance. These challenges to accurate survey measurement of health insurance coverage will be exacerbated by the implementation of the Affordable Care Act of 2010. The mechanisms through which people will be able to acquire insurance coverage will proliferate, including tax subsidies, publicly-funded premium assistance, and expansions of Medicaid. Adding to the complexity, considerable variation in health insurance premium assistance and publicly sponsored programs is expected across the states. To monitor the impact of these new mechanisms on health insurance coverage, access to health care, and on health outcomes, survey researchers will need to understand how best to elicit accurate information from assistance recipients and program participants. In this paper we report the results of a survey in which two random samples of respondents were asked about health care payment assistance from five sources: employers/unions, professional associations, federal government, state government, and local government. The two samples, an RDD statewide telephone sample and a random sample of respondents drawn from an administrative list of caregivers of children currently enrolled in a state sponsored health insurance program, permit exploration of the level of premium assistance awareness among those with employment-based coverage and those with publicly sponsored coverage. We will report on respondent characteristics associated with premium assistance awareness. We will discuss the

implications of the results for measuring premium assistance and suggest alternative approaches to assessing health insurance coverage assistance.

The Impact of Healthcare Utilization on Satisfaction with Health Insurance Plans.

LinChiat Chang, LinChiat Chang Consulting, LLC (contact@linchiat.com); Jon A. Krosnick, Stanford University (krosnick@stanford.edu); Trevor Tompson, Associated Press (TTompson@ap.org)

This paper explores the impact of healthcare utilization on perceived quality of health insurance plans, using data from a landline-cellphone bilingual (English/Spanish) telephone survey conducted in November 2009 with a nationally-representative probability sample of 1,502 adults. Although 72% of adults rated the quality of their private plans as "good" or "excellent", those who had received more inpatient care over the past 5 years gave significantly lower quality ratings, as did people who reported poorer state of health. In contrast, the negative relationship between healthcare utilization and satisfaction with insurance plans did not emerge among Medicare or Medicaid recipients, and was reversed among beneficiaries of military health plans.

Further analyses revealed that the negative impact of inpatient care utilization on satisfaction with private insurance was more pronounced among the less empowered segments of the population - namely, ethnic minorities, women, and people in lower income groups. In contrast, the amount of outpatient care over the past 5 years was associated with higher quality ratings for private health plans only among White Americans and people in higher income groups. The impact of healthcare utilization on perceived quality of private health plans was mediated by whether the plans had refused to pay for healthcare services in the past.

Taken together, these results suggest that one reason the majority of Americans are contented with their private health insurance plans is because they have not yet had to push the boundaries of their coverage, whereas the minority of Americans who have had to cope with substantial healthcare utilization, particularly utilization of medical services in hospitals, are not quite as satisfied with their private plans.

Survey Research in the Arab Gulf: Lessons from the Region.

Political Attitude Research in the Middle East.

Jill Wittrock, Center for Political Studies (jillwitt@umich.edu); Mark Tessler, Department of Political Science (tessler@umich.edu)

The Arab world has experienced a critical transition in the availability of systematic and objective political attitude survey data. Until recently, research on individual-level political attitudes, values and behavior patterns was limited to a handful of American and Arab social scientists. In addition, the available data were constrained by the small number of countries in which political surveys could be conducted, the degree to which genuine representative samples could be drawn, and the extent to which sensitive topics could be investigated. This situation is now changing. At present, it is possible to carry out systematic and objective political attitude surveys in at least ten Arab countries, and a growing number of social scientists in the Arab world are acquiring the skills, the resources and infrastructure necessary to conduct this research. In addition, noteworthy cross-national survey collaborations have administered politically-focused surveys in the Arab world, including the World Values Survey (WVS) and Arab Barometer. This paper reviews the past challenges of conducting such research in the Arab world, describes the recent

developments and data collection efforts of the WVS and Arab Barometer, and discusses the future opportunities for political attitude research in the region.

Building Survey Research Capacity in the Arab Gulf Region: the Case of the Social and Economic Survey Research Institute (SESRI) at Qatar University.

Darwish Alemadi, Social and Economic Survey Research Institute, Qatar University (dalemadi@qu.edu.qa); Hanan Abdul Rahim, Social and Economic Survey Research Institute, Qatar University (hanan.arahim@qu.edu.qa)

Over a relatively short period of time, Qatari society and other societies in the Arab Gulf Region have been transformed from predominantly tribal communities to modernizing countries with significant energy-based wealth. With a substantial proportion of expatriate labor force as a predominant feature of these countries, the rapid economic changes are expected to have had significant effects on the organization of social relationships and on cultural values and attitudes. However, advances in social research methodology have not kept up adequately with the pace of change in Arab Gulf societies or with the opinion of nationals and residents about these changes. It is vital that such issues be documented and tracked over time. This presentation will describe a new initiative at Qatar University, the national university of Qatar, called the Social and Economic Survey Research Institute (SESRI), which was established officially in September 2008. The initiative is in line with Qatar University's and the political establishment's drive to emphasize research as a basis for decision making. The mission of this new Institute is to strengthen national capacity for studying social and economic issues in the country, including public opinion research. High quality social surveys will enable the Institute to contribute to evidence-based policy formulation and decision making in the country and to understand and explain societal trends. The presentation will describe in detail the rationale and process of setting up this Institute, including its partnership with the Institute for Social Research (ISR) at the University of Michigan and the challenges and opportunities associated with this initiative in the Gulf Arab region. The model of the Institute (university-based and independent) will be explained as will its research agenda.

Survey Data Collection in the Arab Gulf Countries.

Abdoulaye Diop, Social and Economic Survey Research Institute (SESRI) (goumalo@yahoo.fr); Elmogiera F. E. Elawad, Social and Economic Survey Research Institute (elmogiera@qu.edu.qa); Kien Trung Le, Social and Economic Survey Research Institute (kienle@qu.edu.qa)

Survey data collection in Arab countries, like in any developed or developing countries, is affected by many factors, including the socio-cultural aspects of the survey environment. This environment, which is characterized by strong cultural customs and traditions, needs to be taken into account when planning for survey data collections. The Gulf Cooperation Council (GCC) countries are not an exception within the Arab and Islamic countries. In addition to the Arab and Islamic traditions, these GCC countries are experiencing a rapid growth of economic development associated with a massive presence of foreign workers in their countries, a population that is sometimes larger, even much larger, than the population of citizens. These elements, put together, create a number of challenges to any survey organization conducting wide, representative public opinion polls. Do Arab nationals allow women in their households to be interviewed by outsiders? What are the gender and interviewing preferences (mode of interview, day, time, nationality of interviewer) for both male and female Arabs? What are the fieldwork implications of the presence of a large pool of low-skilled and high-skilled labor migrants? What are the challenges associated with the many languages in use in the Gulf? Are sampling and interviewing strategies different, or differently received, among citizens, blue collar expatriate workers in communal living situations, and white collar expatriate workers? This paper presents the field survey operations methods used in the first public opinion surveys conducted by the Qatar University Social and Economic Survey

Research Institute (SESRI), representative surveys that were collected not only in Qatar, but in other GCC countries. The paper highlights the lessons learned during the survey data collection process which could be very important for future surveys in the Gulf and Arab regions.

Handling of Sensitive Social Topics in the Arab Gulf.

Mohammad Nizam Khan, Social and Economic Survey Research Institute (SESRI) (nizamkhan@qu.edu.qa); Hanan Abdul Rahim, Social and Economic Survey Research Institute (SESRI) (hanan.arahim@qu.edu.qa)

Conducting surveys on social issues considered controversial according to local standards can be especially problematic in a conservative and relatively closed society such as the Arab Muslim population in the Arab Gulf. Yet the definition of what is controversial can be context-specific and requires intimate knowledge of the society under study. An illustrative example of this can be found in the study of marriage, family, and gender roles. Qatari society is at a unique intersection of modernity and conservatism, where a high proportion of women are educated and working, but where Islamic norms and tribal traditions are still observed. We asked a representative cross-section of Qatari nationals a module of potentially controversial questions on beliefs and practices regarding freedom of mobility for women and wife beating. In this presentation, we will discuss the approaches undertaken to maximize response validity - avoiding social desirability, overcoming mistrust of interviewers, and increasing the comfort level of respondents. The approaches within include consideration of what and how questions can be asked according to social norms in the region, or that might be best avoided within the context of the Arab Muslim population in the region. The presentation analyzes the effect of respondent and interviewer gender, and interviewer nationality, on the questions being asked, as well as hypothesizing about impacts of the mode of data collection (in this case, CAPI). We are able to further leverage the data through the fact that the gender roles questions were based on questions found in the Demographic and Health surveys (DHS), allowing for cross-cultural analysis and comparisons.

Challenges and Solutions to Traditional Western Data Collection Methods in the Middle East.

Vadim Volos, GfK (vadim.volos@gfk.com)

Multinational corporations and foreign governments that commission market and public-opinion research in the Gulf expect data collection to meet survey quality standards perceived as such based on the experience and knowledge that was mostly accumulated from research projects conducted in Western Europe and North America. In particular, these expectations presume that probability samples of the general public can be accomplished via random selection of households, using a door-to-door contact approach. Local researchers in the Gulf often claim that such approaches are prohibitive due to local culture and insist on snow-balling or seed sampling. The presentation discusses a recent research project executed by GfK in the Gulf in which we tried to avoid both of these tracks and employ instead a custom method of face-to face interviewing that we believe provides inclusive probabilistic sampling of the target audience (upper income citizens in the KSA, UAE and Qatar) while taking into account local cultural sensitivities.

The Full Monty on RDD Samples – Enhancement, Weighting, Abandonment.

RDD Sampling in the 21st Century – Construction of a True Landline Frame.

Mansour Fahimi, Marketing Systems Group (mfahimi@m-s-g.com); David Malarek, Marketing Systems Group (dmalarek@m-s-g.com); Gerry Holzbaur, Marketing Systems Group (gholzbaur@m-s-g.com); Jerry Oberkofler, Marketing Systems Group (joberkofler@m-s-g.com)

Virtually, all RDD sampling frames are comprised of 100-series telephone banks that include at least one listed residential number (1+listed frame). This traditional methodology can result in massive undercoverage of households in the following three distinct ways:

1. A growing number of households are abandoning their landlines and vanishing from the frame – such households have markedly different profiles as compared to the general public;
2. With advances in the telecommunication technologies, “techie” households are flocking to quasi-landline alternatives that rely on Voice over Internet Protocol (VoIP) – many of the VoIP-based telephone numbers remain unlisted; and
3. Landline RDD frames carry geodemographic data at various levels of aggregation for targeted sampling – such data pertain to all the US and not landline households and can result in miss-targeted samples with poor coverage properties.

In order to address these evolving challenges, Marketing Systems Group has developed a new methodology for RDD frame construction that relies on the census of landline and quasi-landline residential telephone numbers for each 100-series bank, as well as 1000-series block level counts of cellular numbers. While eliminating the shortfalls associated with the existing RDD frames, this methodology provides a host of possibilities for improving the efficiency of the resulting samples. Such efficiencies include PPS sampling based on correct residential density of banks, appendage of exchange-level geodemographic data relevant to landline households, as well as providing estimates for the number of cell-only households at small levels of geography.

Improving Landline RDD Coverage and Dialing Efficiency Through The Selective Use Of 100 Banks With No Listings.

Andy Weiss, Abt SRBI Inc. (a.weiss@srbi.com); Mike Battaglia, Abt Associates Inc. (Mike_Battaglia@abtassoc.com); John Boyle, Abt SRBI Inc. (j.boyle@srbi.com); Donna Eisenhower, New York City Department of Health and Mental Hygiene (deisenho@health.nyc.gov); Stephen Immerwahr, New York City Department of Health and Mental Hygiene (simmerwa@health.nyc.gov); Tara Merry, Abt SRBI Inc. (t.merry@srbi.com)

The increase in Cable/VoIP phone service over the last several years has resulted in an increase in working landline phone numbers that fall outside the landline RDD sampling frame of 100 banks with one or more listed number. Expansion of this type of phone service and the increase in working numbers in banks with zero listings has varied substantially by geography making the impact difficult to measure. Expanding the frame to 1000 blocks with at least one listing produces an inefficient sample.

Building on a previous study that measured the scale of the zero bank coverage problems in New York City, we have targeted a subset of zero banks and added them to our sampling frame. We are reporting data across several studies using this sample design.

This design has resulted in a self-weighting sample with improved coverage that is more efficient than a frame limited to banks with 1 or more listings. In one county we have completed 15% of our interviews from records that would not have been included in a frame that did not include zero banks. In some neighborhoods the zero bank completes account for more than 30% of the total landline completes.

Improving Landline RDD Coverage with an ABS Supplement.

Robin Gentry, Arbitron (robin.gentry@arbitron.com); Michelle Cantave, Arbitron, Inc. (michelle.cantave@arbitron.com)

Arbitron currently uses a hybrid sampling approach to cover populations in metro areas it surveys. A 2+ list assisted RDD landline sampling frame is used for households with one or more landline numbers, and an address frame is used to identify cell phone only and cell phone mainly households. While utilizing this hybrid frame approach has significantly improved coverage, recent research studies have raised questions about the coverage of households with landline telephones when using a list-assisted RDD sampling approach. With the expanding phone service options for consumers, including cable companies and voice over internet protocol (VOIP), the traditional RDD landline sample frame has come under increasing scrutiny. Estimates of the number of residential phone numbers that may be outside the 1+ 100-banks frame have ranged from 5 to 20 percent (see www.surveypractice.org/ September 2008, January 2009, April 2010) and an empirical study conducted by Arbitron estimated the undercoverage to be between two and six percent.

Arbitron conducted a follow-up study in September 2010 to determine whether zero and 1-listed 100 bank landline numbers could be identified within the ABS frame and used to improve coverage in the RDD frame. This study looked at two groups of landline numbers identified as being in a zero or 1 banks—addresses where a phone number was matched based on a telephone directory and/or commercial database and addresses where no match was initially found but a respondent reported a landline telephone number on their returned screener. Results from this study will provide guidance about the utility of zero and 1 bank numbers from the ABS frame and will also include information about the characteristics of households and persons who reside in zero and 1-banks.

Weighting for Enhanced RDD Sampling.

Vicki J Pineau, TNS (vpineau46@juno.com); Margaret E Strickland, TNS (margaret.strickland@tns-global.com)

Population and household coverage in standard residential RDD sampling is declining substantially due to increases in households using cell only telephone service. The latest official figures from the National Health Interview Survey (NHIS: July – December 2009) projected the incidence of cell-only households at 24.5% in the U.S. Survey researchers around the world are struggling with issues around declining response rates and changes in the composition and coverage of telephone sampling frames. The quality of our survey data both in terms of bias and variance are at risk if we are not aggressive and holistically attacking this problem. Enhanced RDD refers to sampling from both landline and cell-phone telephone exchanges to improve population coverage and is being embraced tentatively. In Enhanced RDD, post-stratification weighting is generally required to account for the incidence of cell only sample primarily due to the differential sampling among the frames relative to the population. In this paper, we will discuss decision criteria for when post-stratification weighting should be applied independently within Landline and Cell-Only sample, and not just applied in total. We will discuss both bias and variance impacts leading to our recommendations.

Conversion of CATI Surveys to Mail Mode.

Pat Dean Brick, Westat (patdeanbrick@westat.com); Douglas Williams, Westat (douglaswilliams@westat.com); J. Michael Brick, Westat (mikebrick@westat.com)

During the past decade, Random Digit Dial (RDD) surveys have experienced a precipitous drop in response rates and coverage. The decline of RDD is a loss for survey research since RDD provided a battery of rigorous and sophisticated methodologies including: within household sampling schemes;

complex survey instruments with multiple paths through the instrument, intricate skip patterns and derived variables generated during the interview; and sophisticated sample management tools that captured detailed call histories used for making decisions on calling protocols at the individual case level. The decline of RDD has left survey practitioners scrambling for ways to reproduce these sophisticated and effective survey methods and tools from RDD studies.

Some studies previously conducted by RDD have been reincarnated as mail surveys, and this process of migrating RDD to mail mode is the focus of our paper. Here we challenge the common assumption that the advantages of RDD must be sacrificed when switching to mail mode. We describe how some of these benefits can be retained in mail surveys by using new technologies for handling paper and by different survey methods. Screening and within household sampling done by CATI can be completely transferred to a mail screener. The screener identifies the eligibility of the household for the survey (e.g., children in the household), and it can be used to identify eligible persons so that the sampling of persons can be done rigorously. We present design principles for a number of different types of mail screeners. The screener information can also be used to tailor the main questionnaire to a specific respondent— thus duplicating the multiple paths through the instrument of CATI. Furthermore, there are now systems for mail surveys that can mimic the sample management and post-collection processing of CATI.

Friday, May 13, 2011

3:15 p.m. - 4:15 p.m.

Poster Session 2

Poster Session 2.

The Electronic Fax: Transitioning to a New Mode of Transcript Collection.

Jamie Wescott, RTI International (jwescott@rti.org); Tiffany Mattox, RTI International (tmattox@rti.org)

Survey methodology is continually evolving to reduce costs, improve data quality, and decrease the burden of participation. Consequently, researchers must balance advances in survey technology with the technological capacity of all participants. In the realm of large-scale data collection from institutions, fax is sometimes the most technologically advanced transmission option available for participants to transmit data.

This poster presentation will explore the use of Rightfax® electronic fax software, which stores incoming data as image files on a server, for the 2009 Postsecondary Education Transcript Study (PETS:09), a large-scale postsecondary transcript data collection which concluded in spring of 2010. Our challenge was to develop automated and efficient collection methods while working within the technological capability and comfort zone of respondents. Prior transcript collections conducted by RTI utilized hardcopy transmission modes, such as traditional fax. For the most recent collection, we introduced new electronic transmission and receipt modes, each designed to ensure the security of the data. The most popular of these modes was electronic fax.

The challenges and efficacy of implementing electronic fax will be assessed based on the results of the collection of approximately 56,000 postsecondary transcripts for PETS:09, which is the transcript collection component for two studies sponsored by the National Center for Education Statistics: The 2008/09 Baccalaureate and Beyond Longitudinal Study (B&B:08/09) and the 2004/-09 Beginning Postsecondary Students Longitudinal Study (BPS:04/09). Specific issues to be presented include evaluation of our goal to seamlessly replace hardcopy fax machines with electronic fax, considerations

related to data receipt staffing and data storage costs, the feasibility of using the software for automating processing; and customization of the Rightfax® software.

The Influence of Public Radio on Public Perception, Civic Engagement, and Community Involvement.

Lisa McQuighan, Arbitron (lisa.mcquighan@arbitron.com)

In the Spring of 2010 Arbitron Inc. collaborated with the leaders in the public radio industry to conduct a study examining the influence of public radio on its listeners with respect to public perception, civic engagement and community involvement. Specifically, this study attempted to determine how involved public radio listeners are in philanthropic, political, and community-oriented activities and what role public radio plays in their decisions to do so.

Arbitron sampled former diary-keepers from the Syndicated Radio Service who were age 25 and older and who identified a public radio station as their most listened to station. In a short telephone interview, respondents were asked whether in the past six months if they were involved in any philanthropic or community-based activities such as donating money, volunteering, attending local town meetings, or contacting government officials about a public issue. Then, respondents were asked to rate how much influence certain factors (such as friends and family, television, the internet, commercial radio, and public radio) had on their decision to become involved in these activities. Last, respondents were directly asked about their likelihood to share information they heard or to become involved in issues they've heard about on public radio, what influence public radio has on their voting decisions and their opinion of public radio in general. The results highlight the important role that public radio plays in their listener's lives and the considerable influence that media can have on both an individual and a community.

Co-Authors: Neal Bonner, Anna Fleeman, Nicole Wasikowski

Use of Passive Electronic Measurement to Complement More Traditional Data Collection Methods.

Anh Thu Burks, Nielsen Company (anh.thu.burks@nielsen.com); Ana Melgar, Nielsen Company (ana.melgar@nielsen.com); Rosemary Holden, Nielsen Company (rosemary.holden@nielsen.com); Yvonne Olivares, Nielsen Company (yvonne.olivares@nielsen.com); Jeff Scagnelli, Nielsen Company (jeff.scagnelli@nielsen.com); Norm Trussell, Nielsen Company (norm.trussell@nielsen.com)

In traditional television audience measurement, household participants are actively required to participate by logging in and out of a device to determine who is watching what programs. As a complement to this more traditional data collection method (i.e., logging in and out) Nielsen is investigating the utility of a non-traditional electronic collection measurement tool that requires no "active" involvement from participants. More specifically, this device is designed to count the number of people looking at a television, with no interaction from the viewer. It uses image sensors to capture the reflection of Infrared (IR) light from the unit. Special software determines if a face is detected, and counts each face as one person. This paper will: 1) Report on the general acceptance of this type of counting technology by the general public gathered through focus groups, 2) Discuss the types of recruitment materials and communication strategies needed to best describe this very new collection tool (i.e., infrared technology) while simultaneously increasing the acceptance of it, 3) Report on the panel's acceptance and willingness to have this type of technology in their home, and 4) Discuss the accuracy of the collection tool. The implications and the use of this type of technology will be discussed in terms of how it can be used to complement and further improve on the data quality gathered from more traditional (existing) data collection methods.

Collection of Biological Specimens in a Three-year Survey of Detroit Residents.
Christine Cowles, Abt SRBI, Inc. (c.cowles@srbi.com); Mark Morgan, Abt SRBI
(m.morgan@srbi.com)

Researchers are increasingly combining biological and survey data to greatly enhance our understanding of interactions between multiple biologic and behavioral factors associated with disease. A high level of participation in the biologic sample phase of studies is critical to the success of the research goals.

The Detroit Neighborhood Health Study is a three-year survey investigating the interaction of mental and physical health, genes and the environment. The DNHS cohort consists of 1,500 residents of Detroit recruited at baseline, and approximately 500 participants were added in year two, for a total of 2,000 in year three. Various types of biological specimens were collected from participants in each of the three waves.

This is a follow up to a 2010 paper which presented partial results. It will describe the characteristics of the study participants, the types of biological samples collected and the success of that effort over three years. We will share best practice recommendations to encourage high participation rates for biological specimen collection studies.

Employing a Process-Tracing Methodology in a Questionnaire: Evidence from a Statewide Poll.
Virginia Tangel, Eagleton Center for Public Interest Polling, Rutgers University
(vtangel@sociology.rutgers.edu); David P. Redlawsk, Eagleton Center for Public Interest Polling,
Rutgers University (redlawsk@rutgers.edu)

In October 2010, New Jersey Governor Chris Christie withdrew New Jersey's support for the Access to the Region's Core (ARC) tunnel project from New Jersey to Manhattan under the Hudson River, citing cost overruns. In response to the cancellation of the project, a political firestorm ensued between the Governor, state legislature, federal government, and co-signing agencies, but public opinion was on the Governor's side: an October 2010 Rutgers-Eagleton Poll found that 51% of New Jersey voters supported Governor Christie's decision to axe the project, with 39% opposed.

In the December 2010 Rutgers-Eagleton Poll, respondents were presented with a linear series of questions pertaining to scenarios surrounding the ARC project. The questions measured baseline support of public transit, support of the Governor's decision to cancel the project, support for the project in light of actual scenarios, and a hypothetical scenario about the project's potential effect on job growth. Using a dynamic process tracing methodology (Redlawsk et al. 2010, Lau and Redlawsk 2006) in the questionnaire, we gather a more nuanced picture of support or opposition to the largest current public works project in the country. Our research questions at what point do respondents switch their support or disapproval of Governor Christie's plan to cancel the ARC project, if at all? We hypothesize that respondents who supported Governor Christie's decision will switch their support for the project's completion when asked if they would support it if the cost overruns were absorbed by other agencies; we also hypothesize that respondents who initially opposed Governor Christie's decision to cancel the ARC project will maintain support for the project throughout the series of questions. The data will trace individual respondents' views at the aggregate to measure the robustness of the support for Christie's decision, providing a possible "affective tipping point" (Redlawsk et al. 2010).

Societal Conflict Over Research and Treatment Using Human Stem Cells? Public Perceptions.
Mariah Evans, University of Nevada, Reno (mariahev2@gmail.com)

Research using human embryonic stem cells has roused widespread ethical debate for nearly two decades. NIH guidelines rule out funding for therapeutic cloning, a scientifically promising line of research, while accepting research based on stem cells extracted directly from embryos or sourced from adults, the last widely seen as an ethical panacea by elites. Societal conflict over stem cell research has recently erupted into a major federal law suit. The general public has largely been excluded from this debate, except for some useful polls tracing trends over time. But systematic studies of the public attitudes towards stem cells and related technologies remain terra incognita. To address these issues, I use the International Survey Center's May, 2009 survey "Attitudes toward Stem Cell Research 5" with 2295 cases, a representative national sample of US adults. Conceptualization, question development, and pretesting was based on our series of Australian studies beginning in 1993. We focus on two key dimensions: the nature of the procedures and the seriousness of the disease to be treated, using examples that are important in current research. Factor analysis shows that there are at least five distinct factors, which we measure with reliable multiple item scales. Structural equation models reveal the linkages among different attitudes and their social and cultural sources. Results show that the American public overwhelmingly approves three lines of stem cell research, and in almost equal measure: Therapeutic cloning is not controversial in their view, but neither is there any special ethical virtue in using adult cells. Contrary to the influential "slippery slope" argument, attitudes on therapeutic cloning are not closely related to attitudes to reproductive cloning. For stem cell issues, current public policy appears to be based on a misperception of public opinion.

How Can We Believe What They Say? The Role of Missing and Validating Data in Panelists Demographic Information.

Kumar Rao, The Nielsen Company (kumar.rao@nielsen.com)

The use of online panels (probability-based or volunteer opt-in) as a mode of data collection has become increasingly popular in market, social, psychological, and medical research (Callegaro and DiSogra 2009). As part of the sign up process, potential panel members are required to complete an online recruitment survey that asks for their basic demographic information. In this study, we analyze data from an opt-in panel in the United States that is composed of respondents who voluntarily sign up (opt-in) to become members of the panel. Like any other online survey, responses to some questions are optional (i.e. skip is permitted), while for others a response is required to advance to the next page. In this paper, we draw upon data obtained from the welcome survey to examine the issue of missing data (i.e., item nonresponse) in the optional set of demographic questions. Specifically, we examine the item nonresponse rate for these questions and dropout rates across all pages in the web survey questionnaire. Then, we attempt to answer one of the fundamental questions related to the practice of survey science "how can we believe what respondents say?" To answer this, we use panel members' behavioral data, which is captured from their online and offline activity on their computers, to validate some of their responses to the demographic questions. We construct an overall measure of data quality, defined in terms of item response rate, accuracy of responses, and completeness of responses, and identify demographic predictors of data quality in a multivariate context. Lastly, we discuss the findings from the study and its implications for survey science and conclude with recommendations for future research.

Respondent Conditioning: Effects on Participation and Data Quality.

Graciela Contreras, U.S. Census Bureau (graciela.contreras@census.gov); Kyra Linse, U.S. Census Bureau (kyra.m.linse@census.gov)

The Census Coverage Measurement (CCM) Person Interview was the main interviewing stage of the program conducted by the Census Bureau to measure the coverage in the 2010 Census. A sample of households was independently interviewed to determine who lives at the address now, who lived there on Census Day and if the people at the address had any other places where they lived. The data were collected using a Computer Assisted Personal Interview (CAPI) instrument that asked a series of questions to roster people and collect alternate addresses where these people could have been counted in the census. The results are used to determine the residency of the people and are compared to 2010 Census results to measure coverage.

Due to data quality concerns, a small number of completed cases were reworked during production using the same CAPI instrument. This paper will compare the results of the first and second interviews for these cases to evaluate the implications of respondent conditioning. We expect that the second round of interview attempts resulted in a higher noninterview rate. Also, for those who responded to both interviews we hypothesize that, compared to overall averages, these interviews will result in more missing data, more incomplete data (i.e., shorter answers to open ended questions) and less reporting of alternate addresses, which was a key data point of this interview. We will also investigate if conducting the second interview with the same respondent versus talking to a different person from the same household impacts the magnitude of any respondent conditioning. We suspect that the overall negative impact on data quality will be reduced when a different household member acts as the respondent in the second interview.

Cognitive Pretesting of Memory Aids for the National Crime Victimization Survey.

Alison K. Baldwin, NORC at the University of Chicago (baldwin-kate@norc.org); Lisa Lee, NORC at the University of Chicago (lee-lisa@norc.org); James Carr, NORC at the University of Chicago (carr-james@norc.org)

The National Crime Victimization Survey (NCVS) asks respondents to report on crimes they experienced in a six-month reference period. On behalf of the Bureau of Justice Statistics, NORC at the University of Chicago has been conducting research on the impact of a proposed 12-month reference period upon the NCVS. Changing to a 12-month period might increase the cost efficiency of the survey but could also have adverse effects on survey data quality. The longer ago a crime incident occurred, the more likely it is to be forgotten, leading to underreporting. Data quality can also be diminished by telescoping, in which errors in the recall of the dates of crime incidents increase with a longer reference period. NORC is designing two methods of improving recall with a lengthened reference period. One method adds crime-related attitude questions before the NCVS crime screening questions. A second method has respondents complete an event history calendar to use along with the crime screeners. These methods will be tested for accuracy and usability in a field experiment involving telephone interviews. In addition, we will use validation data from police reports to assess the quality of the data collected during the cognitive pretesting phase. In this paper we will discuss the challenges presented by simultaneously testing instruments for multiple experimental conditions to help determine the most useful methods for improving recall.

Twitter Content Capture and Analysis: A Case Study of the 2010 AAPOR Conference.

Elizabeth Dean, RTI (edean@rti.org); Joe Murphy, RTI (jmurphy@rti.org); Robert Furberg, RTI (rfurberg@rti.org); Anna Thompson, RTI (althompson@rti.org)

Twitter is a popular micro-blogging tool that allows users to broadcast their thoughts via the internet in posts of 140 characters or less. Over 145 million registered Twitter users post over 90 million tweets per day. Twitter is used as a news source, a tool for social interaction and networking, and a marketing

platform. Recent research has indicated Twitter may have predictive value in forecasting elections results and market performance.

Twitter has become a popular communication tool at academic and professional conferences and meetings. Attendees use Twitter to maintain real-time online conversations alongside live spoken remarks, obtain and distribute logistical information, advertise presentations and booths, facilitate in-person meetings, extend debate about conference topics, and register assessments on the quality of the conference.

At the 2010 annual meeting, AAPOR embraced and supported the use of Twitter by displaying a board updating the conference Twitter stream and by including a Twitter feature in the official conference iPhone app. The official #AAPOR2010 hashtag (a text tag included within each tweet to associate the tweet with the conference stream) was used 309 times by 80 different Twitter accounts between May 3, 2010 and May 16, 2010.

This paper will analyze the overall volume and nature of Twitter content associated with the AAPOR 2010 conference before, during, and after the event. Fluctuations in tweet volume will be analyzed by date, time and user profile. Tweet content will be coded and presented by topic, originality or redundancy, and interactivity. This presentation will enhance understanding of conference use of Twitter by AAPOR, demonstrate techniques for analyzing Twitter data, and provide recommendations for adding to the value of Twitter use by the AAPOR community. Preliminary analyses suggest Twitter may be a useful tool for evaluating conference quality, coordinating social logistics, and improving efficiency of planning by attendees.

Public Opinion Toward Police Enforcement of Illegal Immigrant Policy: The Case of Prince William County.

Kien Le, Social and Economic Survey Research Institute, Qatar University (kienle@qu.edu.qa); Thomas Guterbock, Center for Survey Research, University of Virginia (tomg@virginia.edu); Abdoulaye Diop, Social and Economic Survey Research Institute, Qatar University (adiop@qu.edu.qa); Deborah Rexrode, Center for Survey Research, University of Virginia (dlr3r@virginia.edu)

The increasingly difficult economic conditions experienced in the US since late 2007 have brought the immigration debate back to the forefront of attention for voters and policy makers. In response to these conditions, some local governments have enacted policies to deny illegal immigrants access to services and to limit illegal immigration at the local level, despite the fact that immigration control is a federal mandate. While the literature is abundant on public opinion towards immigrants in general and illegal immigrants in particular, we know little about public opinion towards the police departments which enforce these laws. In this paper, we will study this issue using data from annual surveys (2006 to 2010) conducted in Prince William County, a large, Northern Virginia jurisdiction which implemented a policy in 2008 ordering its police to check on the immigration status of everyone they arrest.

Using ordered probit regression to take into account the ordered nature of the public opinion ratings and dummy variables to represent all possible combinations of years and races, we are able to conduct statistical tests on the change in the satisfaction for each racial or ethnic group while controlling for difference in individual characteristics.

We find a strong impact of the immigration policy on public satisfaction with the police and related survey items. In the year when the police first started enforcing the policy, there were significant decreases in satisfaction with police. Hispanics showed the largest decrease. Surprisingly, the decreases proved to be

temporary, as overall satisfaction with police bounced back in the following year, especially for Hispanics, and remained stable in the year after that. We discuss how public awareness of the policy together with police handling of the policy and population shifts might explain both the drop and the recovery of public satisfaction.

Studying Broadband Access: Consider the Non-Adopter.

Marc Weiner, Bloustein Center for Survey Research, Rutgers University (mdw@ejb.rutgers.edu); Chintan Turakhia, Abt SRBI Inc (c.turakhia@srbi.com); David Ciemnecki, Abt SRBI Inc. (d.ciemnecki@srbi.com); Orin T. Puniello, Bloustein Center for Survey Research, Rutgers University (otp@ejb.rutgers.edu)

In March 2010, the Federal Communications Commission made the establishment and user-adoption of a high-speed Internet infrastructure a national policy priority. A vital part of that effort, beyond the development of a broadband infrastructure, is understanding user adoption and usage propensities. A fair amount of recent research addresses attitudes and behaviors that lead to broadband adoption and the types of activities in which various populations engage over the Internet. However, research on the non-adopter has lagged.

Thus, rather than further exploring the broadband adopter, we focus on the non-adopter as the greater social and cultural barrier to full incorporation of the Internet into the American economic/commercial structure. We explore, both descriptively and inferentially, predictors of non-adoption at the statewide level in an effort to develop a best-practices model for collecting and analyzing data from broadband non-adopters.

Relying on original statewide data collected under the National Telecommunications and Information Administration's nationwide broadband mapping and planning effort, we describe the attributes and suggest models that predict non-adoption, a critical concomitant component to full, nationwide broadband adoption. From a total sample of 3,100, we analyze 1,240 non-adopters, 37.5% of which were captured in a cross-section and 62.5% of which were collected via stratified oversample. These 1,240 non-adopters are comparatively analyzed against 1,860 adopters, 25% of whom were captured by cell-phone and the remaining by landline.

In addition, we discuss the methodological challenges of data capture from broadband non-adopters, where the phenomenon-of-interest (broadband adoption and use) is inversely associated with advanced age, and where the propensity to respond to a landline survey increases with age.

In sum, then, this paper presents a full-overview of the substantive aspects and methodological challenges of describing and modeling broadband non-adoption and non-use. These findings have both methodological and policy implications.

Minimizing Item Nonresponse in a School-based Web Survey of High School Students.

Laura Burns, RTI International (lburns@rti.org); Jill Dever, RTI International (jdever@rti.org)

The High School Longitudinal Study of 2009 (HSLS:09), conducted for the National Center for Education Statistics (NCES), tracks more than 21,000 U.S. 9th graders through their high school years and beyond. The main focus of HSLS:09 is to understand students' academic goal-setting and decision-making regarding postsecondary educational destinations and career choices. The survey emphasizes especially the STEM areas (science, technology, engineering and mathematics).

School-based surveys of secondary school students are subject to strict time limitations not placed on many other types of surveys. The questionnaire and, in this case, math assessment must be completed by students within the time set aside during the school day by the school's administration. This creates a challenge for questionnaire developers who wish to maximize the amount of information that is collected while maintaining high response rates for items at the end of the questionnaire. HSLS:09 mitigated the degree of item nonresponse on the web-based questionnaire with two approaches. First, the order in which modules at the end of the questionnaire were administered was varied; students were randomly assigned to one of three groups. Second, students who had not been able to complete the questionnaire in the first part of the in-school session were able to resume work on the questionnaire if time allowed after the math assessment had been completed. In this paper we evaluate the effectiveness of these approaches by comparing item nonresponse rates with other NCES school-based surveys that used paper and pencil questionnaires and therefore were not able to employ these strategies.

Using Scanner Technology to Learn About Foods Acquired By Low-Income Households.

Holly H Matulewicz, Mathematica Policy Research (hmatulewicz@mathematica-mpr.com); Laura Kalb, Mathematica Policy Research (lkalb@mathematica-mpr.com); Nancy Cole, PhD., Mathematica Policy Research (ncole@mathematica-mpr.com); Nicholas M Redel, Mathematica Policy Research (nredel@mathematica-mpr.com)

Most studies of diet quality focus on food consumption, generally over a 24-hour period. However, people consume foods only after they choose where to shop, what to buy, and how much to pay. Food acquisitions have many dimensions. The National Household Food Acquisition and Purchase Survey (FoodAPS), sponsored by the U.S. Department of Agriculture, will collect data from a nationally representative sample of households about foods acquired from all sources during a one-week period. This includes "food-at-home" (acquired for home preparation and consumption) and "food-away-from-home," and will include foods purchased as well as foods obtained for free. To minimize respondent burden and obtain reliable, high quality data, survey respondents are given handheld scanners to scan barcodes on foods brought home. Scanned barcodes can be linked to product names and sizes, thus eliminating measurement error on these dimensions. Scanners also relieve respondents of the burden associated with manually recording information about each food item. To gather feedback on the scanners and the proposed data collection design, we conducted a pretest with 22 low-income households, including persons who may have had limited experience with technology (those with less than high school education and elders). Results were overwhelmingly positive. Each respondent gained the skills necessary to use the scanner and many reported that scanning was the most enjoyable task in the study. This paper describes how we alleviated respondents' apprehensions about this technology through hands-on practice and mock scenarios with actual groceries. It also describes respondents' feedback and experiences using this technology, and modifications to the study protocol which resulted from these findings.

Comparing Attitudes of Evangelicals in an RDD and Internet Sample: Adjusting for Non-random Bias.

Veronica Roth, The Pennsylvania State University (vlr124@psu.edu)

As the use of non-random sampling techniques increases, due in no small part to the advent of the internet survey panel, exploring the differences and similarities of random and non-random samples becomes vital. While haphazard samples are known to have problems with representativeness (Chang and Krsonic, 2009), the implications for multivariate analyses are less clear (Groves, 2006). Weighting may be helpful in making a convenience sample more representative (Groves, 2006), but it can introduce design effects and decrease statistical power. Using data from a 2008 poll regarding evangelical attitudes

towards American public policies, I investigated the differences and similarities between a random and non-random sample. The random sample, collected via RDD, has 133 evangelicals, aged 18-29. The non-random internet sample has 300 cases, collected from an internet opt-in panel. I used logistic regression to model the key demographic variables that differentiate the random and non-random sample. Then I compared a series of bootstrapped models, to look for differences in the distributions of certain outcome variables and key demographics. Finally, I compared the ability of the demographic variables to predict attitudinal outcomes on multiple political issues for both weighted and unweighted models. Due to the lack of Census Data on religion, I used the General Social Survey to construct inverse probability weights on key demographics. I conclude with a discussion of the benefits and challenges of this mixed-mode data collection, and how post-collection analyses may or may not be useful.

Chang, L. and Krosnic, J.A. (2009) "National Surveys via RDD Telephone Interviewing versus the Internet: Comparing Sample Representativeness and Response Quality" *Public Opinion Quarterly* 73(4): pp. 641-678.

Groves, R.M. (2006) "Nonresponse Rates and Nonresponse Bias in Household Surveys" *Public Opinion Quarterly* 70(5): pp. 646-675.

Mode Effects in Self-Administered Surveys Among Pregnant and Parenting Adolescents.

Azucena A. Derecho, RTI International (derecho@rti.org); Brian Head, RTI International (bhead@rti.org); Marni Kan, RTI International (mkan@rti.org); Olivia Silber Ashley, RTI International (osilber@rti.org); Cassie Williams, RTI International (cawilliams@rti.org); Sarah Jones, RTI International (sjones@rti.org)

Asking survey respondents questions about sensitive topics may elicit socially desirable answers. Mode of administration has been shown to affect the magnitude of social desirability bias. Comparisons of computer-assisted self-interviewing (CASI) and self-administered questionnaires (SAQs) among general adolescent samples have produced modest evidence of mode effects, with more accurate reports on sensitive items (e.g., alcohol or other drug use) when the mode of administration is CASI. Few studies have tested for mode effects among pregnant and parenting adolescents. We collected baseline data from a convenience sample of 1,036 pregnant and parenting adolescents enrolled in a cross-site evaluation of the Adolescent Family Life Program, which involved projects in 10 states. The purpose of this study is to examine mode effects among adolescents who completed the survey using CASI (32%) and those who completed using SAQs (68%) on responses to questions about contraceptive use, sexually transmitted disease (STD) prevention, number of pregnancies, desire to have another baby before finishing high school, marriage attitudes, educational aspirations, communication with their parents, and avoiding pressure from others. We explored associations between data collection mode and responses to these questions using multiple linear and logistic regressions, controlling for project site and other potential confounders (e.g., demographic and socioeconomic variables). Results showed that data collection mode was not associated with any of the items examined. These findings suggest that among pregnant and parenting adolescents, responses to SAQs and CASI are generally comparable. When planning a survey within this population, mode may be less important to consider than other factors, such as capacity or budget. Future research should examine other potentially sensitive items relevant to this population to determine whether mode effects exist.

Conducting Surveys on Sensitive Topics; Sexual Behavior and Health.

William C McCready, Knowledge Networks (bmccready@knowledgenetworks.com); Debra Lynne Herbenick, Indiana University, Center for Sexual Health Promotion (debby@indiana.edu); Michael

Reece, Indiana University, Center for Sexual Health Promotion (mireece@indiana.edu); Rick Li, Knowledge Networks (rli@knowledge networks.com)

KnowledgePanel provided the survey data for one of the largest nationally representative study of sexual-health behaviors ever fielded, the National Survey of Sexual Health and Behavior (NSSHB), conducted by Indiana University sexual health researchers from the Center for Sexual Health Promotion. These data provide an updated snapshot of contemporary Americans' sexual behaviors. It includes the sexual experiences and sexual-health-related behaviors of 5,865 adolescents and adults ages 14 to 94. This presentation will focus on the basic age-related findings and the various methodological approaches to panelists that resulted in the consent, cooperation and trust required to generate these important health-related data about a sensitive topic. Initial findings from the survey, presented in nine separate research articles, were published on Oct. 1 in a special issue of The Journal of Sexual Medicine, a leading journal in the area of sexual health. According to the study's findings, 1 of 4 acts of vaginal intercourse are condom protected in the U.S. (1 in 3 among singles). These data, when compared to other studies in the recent past, suggest that although condom use has increased among some groups, efforts to promote the use of condoms to sexually active individuals should remain a public health priority. The study helps both the public and professionals to understand how condom use patterns vary across these varying stages in people's relationships and across ages. A unique feature of the study was the inclusion of adolescent men and women. Many surveys of adolescent sexual behavior create an impression that adolescents are becoming sexually active at younger ages, and that most teens are sexually active. These data show that partnered sexual behaviors are important but by no means pervasive aspects of adolescents' lives. In fact, many contemporary adolescents are being responsible by abstaining or by using condoms when having sex.

Effects of Response Formats when Measuring Attitudes in Consumer Web Surveys Across Markets.

Elizabeth Nunge, Google (enunge@gmail.com); Mick Couper, University of Michigan (MCouper@umich.edu)

Previous studies suggest different response format variations (e.g., order of response options) produce different effects in different cultures. In addition, there are cultural variations in the use of extreme response options and acquiescence. Combined, these effects could produce quite different distributions when comparing results of satisfaction measures across countries. This study investigates the effects of horizontally flipping response scales in online consumer surveys across three different countries, China (n=1896), Japan (n=1222), and the United States (n=3673). Given the strong left-right orientation in Western societies, we expected stronger effects for the U.S. than for China and Japan. Users who opted into a survey linked from the Google Maps interface were presented with either one of two seven-point satisfaction ratings; 'extremely satisfied' to 'extremely dissatisfied' or 'extremely dissatisfied' to 'extremely satisfied.' Findings point to a higher selection rate for the first response options across all three countries, yielding higher rates of satisfaction when the positive response is presented first. The effect on the negative end of the scale ('extremely dissatisfied') appears to be similar across all three countries. However, in China and Japan the effect of response order is also apparent on the positive end of the scale ('extremely satisfied'), while the effect in the U.S. is considerably smaller. We report on the detailed results and discuss the implications for conducting cross-national or cross-cultural research.

Pretty Please, Complete my Survey: Refusal Conversion by Mail.

Ashley Amaya, NORC at the University of Chicago (amaya-ashley@norc.org); Alyson Croen, NORC at the University of Chicago (croen-alyson@norc.org); Karen Wooten, Centers for Disease Control and Prevention (kgw1@cdc.gov)

The increasing difficulty of gaining respondent cooperation in telephone surveys has led to decreases in response rates and increases in refusal rates, survey costs and the potential non-response bias. Respondents may refuse to participate at any time for a variety of reasons: they are reluctant to share personal information with an interviewer, are contacted at an inconvenient time, or are not sufficiently motivated to participate. Thus, it has become increasingly necessary to develop creative survey methods to increase participation among respondents who would otherwise refuse.

One technique of converting refusals is to offer an alternative data collection mode to the respondent. To this end, a mail survey was sent in August 2010 to 2,194 households that had previously refused via telephone to complete the National Immunization Survey (NIS), a list-assisted RDD survey of households with children aged 19 to 35 months and teenagers aged 13 to 17 years. The mail follow-up allowed households that did not complete the screener section and/or household interview to do so at their own pace and in their own time. The experiment included mailing a cover letter addressing the most common reasons respondents refuse to participate and confirmed the survey credibility, a questionnaire allowing respondents to view the entire survey up-front and complete it in their own time, and a small monetary incentive enhancing motivation to complete the survey.

The findings in this experiment address the importance of reaching randomly selected respondents for representative population surveys. It sought to 1) increase response rates and 2) mitigate non-response bias. To evaluate these goals, we conducted analyses on a variety of key rates, such as screener completion, eligibility, and interview completion rates. We also looked at the demographic distribution of these new respondents and their impact on the demographic distribution of the survey compared to population estimates.

Financial Stress, Self-Reported Health Status, and Personal Health.

Robert Kirby Goidel, Louisiana State University (kgoidel@lsu.edu); Dek Terrell, Louisiana State University (mdterre@lsu.edu); Stephen Barnes, Louisiana State University (barnes@lsu.edu); Todd Griffin, Louisiana Department of Health & Hospitals (todd.griffin@la.gov)

In the popular press, the link between financial stress and personal health is frequently noted. In academic research, however, the relationship is more tenuous. Lyons and Yilmazer (2005), for example, report that poor health appears to increase financial burdens but that “there is little evidence that financial strain leads to poor health.” Looking at macro and micro-level data, Ruhm (2000) similarly finds that, because obesity and smoking increase during economic upturns, recessions may be good for personal and public health.

In this study, we utilize 2009 BRFSS social context module to re-examine the relationship between financial stress, self-reported health status, and objective indicators of health in Louisiana. Louisiana, which consistently ranks last or near last in state health rankings and poverty, but has fared relatively well during the recent economic downturn, serves an ideal starting point for this investigation. We hypothesize that financial stress affects self-reported health status as well as the days the respondent reported that their physical health and mental health were not good but has little connection to other indicators of health.

Taking Comedy Seriously: The Effect of Political Comedy on Political Knowledge and Ideological Polarization.

Krysha Gregorowicz, University of Michigan (kryshag@umich.edu)

The relationship between entertainment and information is central to our understanding of the consequences of a high-choice media environment on public opinion. Political polarization and an increasing gap in political knowledge and engagement are particularly worrisome trends associated with the expanding media landscape. Merging information and entertainment, political comedy may reverse these trends by encouraging attention to politics and facilitating learning as a bi-product of entertainment. Conversely, critics complain that political comedy causes political polarization, distracts attention from important political issues, and promotes toxic levels of political cynicism. I argue that comedy is associated with unique and complex patterns of cognition which encourage elaboration and information retention. Further, this cognitive complexity facilitates the development of knowledge structures which shape and constrain political attitudes and beliefs. An experiment was designed to test these predictions. Participants were exposed to either a comedy or news video clip, meticulously edited to be content identical and vary only on the presence of jokes. Several surprising findings challenge the dominant by-product model of political entertainment effects. First, political comedy increases information recall relative to identical information presented in a hard news format. However, incidental exposure does not account for this effect. Instead, the relationship between comedy and learning is mediated by the degree to which one is amused. Lacking prerequisite knowledge and the processing abilities necessary to 'get' jokes, for those low in prior knowledge, comedy has little benefit relative to hard news. Learning is most pronounced among those with moderate levels of prior knowledge. Second, as opposed to political polarization, comedy seems to encourage ideological moderation. Relative to news and control, those exposed to comedy rate themselves as significantly more moderate and report significantly lower approval ratings for people and groups perceived to be ideologically extreme in either direction.

Exploring Record-of-Call Paradata in an RDD Survey.

Taylor H Lewis, National Center for Health Statistics (tlewis@survey.umd.edu); Meena Khare, National Center for Health Statistics (meena.khare@cdc.hhs.gov); James Singleton, National Center for Immunization and Respiratory Diseases (James.Singleton@cdc.hhs.gov)

The National Immunization Survey (NIS) is designed to collect vaccination information for U.S. children aged 19 to 35 months. Data collection consists of two stages: in the first, a random-digit dial (RDD) telephone sample is fielded to screen for households with one or more age-eligible children; in the second, consent is sought to mail an immunization history questionnaire to the child's provider(s) which collects data pertaining to the number and type of vaccinations received.

Paradata are data about the survey process, the utility of which have been increasingly investigated over the past decade. This poster visualizes data from the NIS 2009 call record logs for insights into potential data collection efficiencies and survey error assessments. We define a few level-of-effort metrics, including number of dials and total interviewer time expended to attain a respondent, and weigh the trade-offs between the costs attributable to obtaining the hardest to reach participants and various measures of data quality.

If Not Response Rates, What Alternatives Should Be Used For Survey Quality Indicator Measures.

Donsig Jang, Mathematica Policy Research (djang@mathematica-mpr.com); Flora F Lan, National Science Foundation, Division of Science Resources Statistics (flan@nsf.gov)

Response rate is often used as an indication of measuring the quality of the survey. However, it only tells one side of the survey story; the other side about the association between respondents and non-respondents are unknown. Researchers continue to seek the tools to assess and compare the quality of the response to different surveys. Introduced by Schouten et al. (2009), R-indicators are used to measure how well a respondent set represents the sample or population from which it was drawn. This measure

may be a better indicator of survey nonresponse bias than response rates for survey outcomes closely related to auxiliary variables used for R-indicator calculation. In this study, we will examine R-indicators as a possible alternative method in measuring survey quality.

Using the data from the 2008 NSRCG, we will examine R-indicator and response rate trends over a six-month data collection period to observe the relationship between the two measures. To understand the relationship between the two measures and potential nonresponse bias, we will examine any linkages between trends in key survey estimates during data collection and the two survey quality measures. Specifically, each week during the last eight weeks of the study, we will construct survey analysis weights. These weights will be used to produce survey estimates that only include sample members who have responded by that week. We will then compare these estimates with the estimates from all sample members.

Through this empirical investigation, we expect to identify the relationship between survey quality measures (R-indicators and response rates) and nonresponse bias. Our findings will help us to identify an alternative quality measures to use for the 2010 NSRCG and to control the survey process in such a way that a representative sample is obtained and decide when to end the data collection period.

Showcard Use And Mixed-Mode Administration In Longitudinal Surveys: Are The Data Comparable?

Beth A. Fisher, NORC at the University of Chicago (fisher-beth@norc.org); Ned English, NORC at the University of Chicago (english-ned@norc.org); Catherine Haggerty, NORC at the University of Chicago (haggerty@norc.uchicago.edu)

Making Connections is a study of ten disadvantaged US urban communities, funded by the Annie E. Casey Foundation. This paper examines one aspect of survey administration that has implications for longitudinal surveys using a mixed-mode design: differences in responses given in person with the aid of showcards compared to those given over the telephone without showcard use and the substantive impact of multi-mode administration.

In the first wave of data collection for Making Connections, all interviews were completed in person. Due to cost constraints, subsequent interviews were completed by telephone when a telephone number was available and the respondent was willing to participate by telephone; approximately a third of subsequent interviews were completed by telephone. When interviews were conducted by telephone, all response options for each question were read to the respondent. If the interview was completed in person, response options were read and respondents were shown a card that enumerated the response options. Questions selected for showcard facilitation were those with a long list of possible responses or those with a scaled response. Showcards are frequently used in surveys because they ease the burden of remembering response options and therefore improve data quality by reducing satisficing or primacy and recency effects (Holbrook, Green, and Krosnick, 2003).

We compare the substantive data collected within and across the two modes of administration between the in-person only group and the in-person/telephone group to explore the potential unintended consequences of attempts to ease respondent burden through the use of showcards and other interview aids. We draw from prior comparisons of interview modes (Roberts, Jäckle, and Lynn 2007, etc.) to investigate differences associated with telephone and in-person interviewing for questions where showcards were used and where they were not.

Tailoring Aided Recall Tools: Relating Task Difficulty And Need For Cognition To The Effects Of A Calendar Tool On Recall Accuracy.

Wander van der Vaart, University for Humanistics (wvdv@uvh.nl); Tina J. Glasner, Utrecht University (t.glasner@uu.nl); Wil Dijkstra, VU University Amsterdam (w.dijkstra@vu.nl)

Research demonstrates that data quality in retrospective surveys can benefit from employing aided recall tools. This is especially so, if these tools are tailor-made to the issues and population at hand. However, despite the popularity of aided recall applications, such as the event history calendar, it largely remains unclear what factors can be utilized in 'tailoring' these tools and improve their effectiveness.

We examine this issue by relating the impact of a calendar application to both respondent characteristics and characteristics of the recall task. Based on the Elaboration Likelihood Model we expect that the impact of the calendar method depends on the respondents' tendency to engage in effortful cognitive endeavors ('need for cognition'). Furthermore, we expect that 'need for cognition' is related to the number and types of landmark events that respondents use as calendar recall aids, and to the impact of these landmarks on data quality.

The data stem from a field experiment on a calendar application that was embedded in a standardized telephone survey regarding purchase behavior of clients of opticians. Respondents' retrospective reports about purchases of lenses and pairs of glasses, during a reference period of seven years, are compared to the optician database information regarding numbers, prices and dates. In addition we established the numbers and types of landmark events reported by respondents.

The impact of the calendar and the landmarks on recall accuracy is related to the difficulty of the recall task, the respondents' 'need for cognition' and their sociodemographic characteristics. Based on the results, conclusions will be drawn on what factors are important in developing tailor-made calendar methods and how this knowledge can be utilized for aided recall in general.

An Exploration of Racial and Ethnic Identity Constructs among U.S. Hispanics with Implications for Survey Design and Analysis.

Martin Cerda, Encuesta, Inc. (martin.cerda@encuesta.com); Ilgin Basar, Encuesta, Inc. (ilgin.basar@encuesta.com); Jessica Jamanca, Encuesta, Inc. (jessica.jamanca@encuesta.com)

There is a great deal of public and scholarly debate as to the best way to ask questions related to race and ethnicity for the purposes of population measurement and trending. While it is recognized that "there is no true gold standard for measuring race and Hispanic origin when they are, by definition, self-identification concepts" (Measuring Race and Hispanic Origin: Cognitive Test Findings Searching for "Truth", AAPOR Conference 2010, Jennifer Hunter Childs, and Rodney Terry; U.S. Census Bureau), it is evident that the current methods used by many create confusion among the general public and in particular for some Latinos ("Some Hispanics puzzle over race question on census form", The Arizona Republic, March 23, 2010) and sometimes confound practitioners in spite of great efforts (Overview of Results of New Race and Hispanic Origin Questions in Census 2000, Research Report Series, Survey Methodology #2007-5, Jorge de Pinal, Elizabeth Martin, Claudette Bennett, and Art Cresce). Furthermore, this current situation impacts how other population groups view their own notion of race and ethnic identity, and we surmise limits how some express their personal views on race and ethnicity as part of the survey research process.

To address this issue, we have designed a series of parallel Hispanic and non-Hispanic studies based on ethnographic research (in situ contextual interviews), diagnostic testing (using both mail and online

surveys to assess the potential use of current and emerging technology), and random probability survey sampling (nationally representative telephone surveys with household and cell phone level sample frames), to explore new ways of defining race and ethnic constructs and testing of possible new questions related to self-identification.

The findings will highlight differences in how the general public views race and ethnic identity and present possible solutions when conducting survey research among U.S. Hispanics and the population at large.

Do Extra Efforts Enhance Retention of Racial/Ethnic Subgroups in Longitudinal RDD Telephone Surveys?

Ashley J. Frear Cooper, University of Nebraska-Lincoln (afrear2@unl.edu); Julia McQuillan, University of Nebraska-Lincoln (jmcquillan2@unl.edu); Stacia Jorgensen, University of Nebraska-Lincoln (shalada2@unl.edu); Amanda Richardson, University of Nebraska-Lincoln (arichardson2@unl.edu); Andrew V. Bedrous, University of Nebraska-Lincoln (abedrou2@unlserve.unl.edu)

Does retention in longitudinal, population-based random digit dial telephone surveys differ by race/ethnicity in the contemporary United States? If so, can mail-based interventions differentially increase retention among racial/ethnic subgroups? To answer these questions we analyzed data from an ongoing health survey of women ages 25 to 45 from the first wave in 2004-2006 through the second wave in 2008-2010. The proportion responding to standard retention efforts--prenotification letters, follow-up mailings (e.g., holiday card with small gift, summary results), and repeated phone calls (up to 197)--was higher among white and Asian women than among Black or Hispanic women. To address this problem, we conducted two extra efforts. First, we randomly assigned subgroups of non-responders (stratified by race/ethnicity) to one of two conditions: (1) a letter with \$5 requesting participation or (2) a letter with \$5 and an offer of \$20 upon completion. Two weeks later, all of the remaining non-responders were offered the \$20 upon completion. Following this effort, the remaining non-responders were sent a 4-page critical-item mail survey that also requested updated contact information (specifically, current phone number). Upon return, those who completed the survey and provided a new phone number were contacted by phone to verify the number, thank them, and ask for participation in the phone survey, which many subsequently completed. This study shows that offering larger cash incentives and a mail option to a panel telephone survey can substantially increase retention (increasing the response rate by 12 percentage points) and is particularly effective for Hispanic women.

New Tricks: Cell Phone Adoption and Use Among Older Adults.

John Fries, AARP (jfries@aarp.org)

Initially, the growing prevalence of cell phone only and cell phone mostly households was limited to younger populations. However, there is increasing evidence that older adults have begun adopting some of these communication patterns. Consequently, organizations focusing on 50+ adults, such as AARP, need to begin exploring and understanding cell phone penetration and usage patterns among older adults. Using dual-frame samples, this research shares initial results from a series of analyses aimed at understanding the noncoverage bias resulting from reliance on landline-only surveys of older adults, including the impact on productivity and the substantive findings.

Public Activism and Public Opinion Among Mid-Life and Older Persons: Comparisons and Trends.

Robert H Prisuta, AARP (rprisuta@aarp.org)

This presentation examines the relationship between public activism and public opinion among mid-life and older persons from several perspectives. Findings are drawn from a telephone survey of a representative sample of individuals 45+ (n= 1,500). Using a three-item activism scale, it will report on the degree to which respondents participated in several active behaviors involving the expression of political opinion, including contacting public officials, taking part in demonstrations or protests, or contacting or interacting with the media. It will compare the demographic characteristics of those individuals with those who are less active, ascertaining the degree to which these participants are representative of the general public. It will also compare activists/non-activists on political ideology, voting behavior, and issues of concern. In addition, it will compare data collected immediately after the 2010 Congressional elections, in December 2010, with similar data collected over a year earlier (August 2009) and also with data collected in previous years. The findings will provide insight into the extent to which public participation in these types of activities represented either general public opinion or an increased generalized level of public activism.

Young Hearts And Old Brains: Does Ideology Change Over Time?

Zeljka Buturovic, Zogby International (zeljka@zogby.com); Leann Atkinson, Zogby International (leann@zogby.com); Joe Mazloom, Zogby International (joe@zogby.com)

Though common wisdom suggests that ideology changes over lifetime, our cross-sectional survey of 40340 likely American voters shows that ideology changes relatively little over one's lifetime. Though 71% of American voters say that their ideology has changed over their lifetime and they report becoming more conservative in almost 2:1 ratio, a closer look shows that these differences are small. Thus, over 70% of likely American voters report that their ideology, measured on a nine point scale, has changed two points or less over their lifetime with more conservative, very conservative and libertarian voters somewhat more likely to report changes of four or more points. Perhaps surprisingly, reported ideological change does not appear to depend on age, suggesting that whatever shift in ideology does occur, occurs early on. One's understanding of whether family or outside influences play the most important role in one's ideology is likewise unrelated to changes in his ideology. However, we find a significant change in the relative importance of economic issues for one's ideology over the lifetime. While economic issues are the most important determinant of ideology overall, and especially so for libertarian and conservative voters our data shows that their importance increases markedly over lifetime for all ideological categories. However, this shift in emphasis does not appear to influence ideological identity, corroborating the view that a one-dimensional ideological scale captures the essence of the phenomenon.

Taming Box's M.

Scott Van Manen, The Futures Company (svanmanen@aol.com); Amy E. Drew, The Futures Company (amy.drew@thefuturescompany.com)

Comparison of covariance matrices is especially valuable when modeling and segmenting data from different populations, e.g., different countries. Unfortunately, the commonly accepted statistics for comparing covariance matrices, Box's M, is volatile to changes in sample size, number of input variables, and even the shapes of the distributions of the input variables. The standard conversion to an approximate F value allows for significance testing across sample sizes and numbers of inputs, but does not provide a measure of effect size, nor does it adjust for bias introduced by non-normally distributed data.

This poster will show a method for "taming" Box's M, including an adaptation for non-normal data, and a conversion into an estimate of effect size in the form of an estimate of shared variance between 2 covariance matrices. Implications of unequal covariance matrices will be demonstrated in real world data, demonstrating the importance of this metric.

Evaluating a Dual-Response Design for the Collection of Salary Data on a Federal Survey.

Vincent E Welch, NORC at the University of Chicago (welch-vince@norc.org); Brianna Groenhout, NORC at the University of Chicago (groenhout-brianna@norc.org); Sarah Hernandez, NORC at the University of Chicago (hernandez-sarah@norc.org); Mireya Dominguez, NORC at the University of Chicago (Dominguez-Mireya@norc.org)

The current global and national economic situation makes the collection of financial data as important as it has ever been. However, the current climate of increased skepticism regarding surveys, generally, and financial data on surveys, specifically, on the part of the general public has made it increasingly difficult for survey researchers to collect these important data.

The Survey of Earned Doctorates (SED) has recently added an item assessing the expected salary of respondents with definite commitments for paid positions after graduation. In an effort to maximize item response rates and the analytic utility of the collected data, the SED allowed for dual-format responding on this item. Respondents were asked to provide an exact dollar amount of their expected salary for the upcoming year, but were allowed to also respond to a fixed range for what their salary for the upcoming year would be.

In the current work, we examine the characteristics of the data gathered from this dual format salary item. Particularly, we will examine if there is evidence of bias in the characteristics of respondents that choose to supply a dollar value versus those who elect to only choose a range on the basis of several important demographic and employment factors (e.g., citizenship status, race/ethnicity, sex, field of study, type of plans, employment sector). We will also examine whether there is evidence that the reported salaries of respondents who provided a dollar value differ from those who chose a range response after accounting for relevant demographic and employment covariates. Finally, we will examine the provided salary data for evidence of mode effects. That is, we will analyze the data to see if doctorate recipients who responded via PAPI, differed in their reported expected salary from those who responded via web and CATI.

The performance of this dual-format responding can have implications

The Politics of Identity in a Time of Economic Crisis, Social Inequality, and Social Change.

Gregory Holyk, Washington and Lee University (holykg@wlu.edu); Andrew Becker, Weststat (andrew.dot.becker@gmail.com)

The substance and importance of identity in American politics change with shifting circumstances and context. We argue that the volatile and uncertain times in which Americans find themselves have increased the importance of identity in political attitudes. Particular political issues and ideologies often serve as a means for individuals to express their core social identities. This motivation to express social identity is heightened in the current stressful political, economic, and conflictual climate. The recent success of the Tea Party movement is an example of this phenomenon.

Experiencing a personal crisis and hardship introduces a desire to enhance one's feelings of self worth/self-esteem. One of the most common self-esteem enhancement strategies is to praise social groups you identify and define yourself with, and derogate those groups that you do not identify with. The expression of political attitudes reflects this process. This practice has a deeply powerful effect on political behavior and thus on the trajectories of American politics more generally.

We use recent survey questions on important current and longstanding issues and policies such as gay marriage, immigration, economic bailouts, economic intervention and equality, gun control, abortion, and

the place of religion in politics to show that attitudes on these issues will reflect a motivation to express a basic social identity more than concerns over the policy itself.

We conduct confirmatory factor analyses and logistic regressions using the waves of the ANES 2008-2009 panel study. We show that indeed there is a large underlying identity component to the current expression of political attitudes that is strengthened by the current economic, political, and social conflict. The panel dataset also allows us to show an increase in the importance of identity-related attitudes over time as the economic crisis worsened and increased importance of social identity for those individuals more affected by these societal changes.

The Public Acceptance Of Nanotechnology: How Ideology And Confidence In Safety Systems Moderate The Effect Of Benefit Perceptions On Federal Funding Support.

Jiyoun Kim, University of Wisconsin, Madison (jkim294@wisc.edu); Peter Ladwig, University of Wisconsin, Madison (pladwig@wisc.edu); Ashley Anderson, University of Wisconsin, Madison (aaanderson3@wisc.edu); Dominique Brossard, University of Wisconsin, Madison (dbrossard@wisc.edu); Dietram Scheufele, University of Wisconsin, Madison (scheufele@gmail.com); Mike Xenos, University of Wisconsin, Madison (xenos@wisc.edu)

Previous research has shown that benefit and risk perceptions of nanotechnology are associated with support for nanotech research, including support for federal funding. However, little research has investigated the potential factors that could influence that relationship.

This study examines how people's perceptions of the benefits of nanotechnology are related to their approval of federal funding. Secondly, it investigates the moderating role of ideology and confidence in technology regulatory system in the relationship between benefit perceptions and support for federal funding. Based on the previous literature on public acceptance of nanotechnology, this study hypothesized that the effect of benefit perceptions on support for federal funding for nanotechnology would be moderated by 1) individuals' ideology, 2) their confidence in safety system.

To test the hypotheses, this study conducted a nationally representative survey that provided respondents with four different definitions of nanotechnology. Data was collected by Knowledge Networks from an online panel (N = 2,338, response rate of 54.2%) between 9 July and 23 July 2010.

Results show that perceptions of nanotechnology benefits are positively associated with support for federal funding for the technology. This relationship is moderated by both individuals' ideology and their confidence in technology regulatory system. Practical implications for policymakers are discussed.

Acknowledgement:

This material is based upon work supported by a grant from the UW-Madison Nanoscale Science and Engineering Center in Templated Synthesis and Assembly at the Nanoscale (Grant No. SES-DMR-0832760). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

Evaluating Change in DSF Coverage for a Representative Sample of Rural Segments.

Michael Latterner, NORC at the University of Chicago (latterner-michael@norc.org)

In 2003, NORC designed its National Frame of segments to supply sample for surveys during the upcoming decade. The USPS Delivery Sequence File (DSF) was used as the primary source of address listings in areas where the DSF was expected to yield coverage properties comparable to traditionally-

listed addresses. NORC conducted in-person listings in areas believed to have inferior coverage, including rural areas and those without geocodable city-style addresses. In the years since NORC's National Frame was constructed, research has demonstrated an increase in the coverage of the DSF. Specifically, there are rural areas where the percent of city-style addresses has increased substantially from 2003 to 2010 (Latterner, English, O'Muircheartaigh 2010), a factor highly correlated with coverage.

At question is whether the observed increase in percent city-style addresses means that less traditional listing will be required for address frame construction. To answer this question, we have examined rural segments in our National Frame that were traditionally-listed in 2003. The proposed poster would illustrate the findings of this research to reveal where we believe the DSF has improved the most, where improvements have not been observed, and beliefs for why differences in these improvements exist. The poster will also provide findings regarding Enhanced 911 conversion, a process by which rural route boxes have been converted to city-style addresses.

Changes in Marginal MSE Per Additional Costs in Multimode Surveys.

Carol Pierannunzi, Centers for Disease Control and Prevention (ivk7@cdc.gov); Hacı Akcin, Northrup Grumman/ CDC Contractor (hakcin@cdc.gov)

This research builds upon cost/benefit simulations by using data from the Behavioral Risk Factor Surveillance System (BRFSS) in 33 states to provide ratios between marginal changes in MSE and marginal changes in costs after the introduction of mixed modes. The BRFSS is the largest continuous telephone survey in the world, collecting health risk behavior and health status data in all 50 states, the District of Columbia and three U.S. territories. The BRFSS is a state-based survey coordinated by the Centers for Disease Control and Prevention. Traditionally a landline survey, the BFRSS introduced cell phone frames in 2008. Bias in landline surveys is well documented. Cell phones have been introduced for their potential to remove bias in the sampling frame. In 2009, BRFSS data show that cell phone surveys were approximately 1.5 to 3 times more expensive than completed landline surveys. Mean age estimates for each state are taken from the landline and cell phone administration of the survey and compared to Census data. MSE and costs per completed survey are used to calculate marginal change in MSE per increased costs due to addition of cell phone frames. Results illustrate the change in MSE for each additional dollar spent and optimal effectiveness of mixed modes.

Differences in Respondent Propensity to Disclose Child's Name in the NHES 2009 Pilot Study.

Cameron McPhee, American Institutes for Research (cmcphee@air.org); Alex Knecht, American Institutes for Research (aknecht@air.org); Sarah Grady, American Institutes for Research (sgrady@air.org)

Reporting of sensitive information has long been a concern of survey researchers. Much of the literature on sensitive questions compares data collection modes, finding that self-administered modes yield higher levels of reporting of sensitive information than interviewer-administered modes. This research shows that within self-administered surveys, reporting can vary widely depending upon the context of the survey.

The redesigned National Household Education Survey is a mail survey with two-stage sampling. The 2009 pilot study (n= 10,000) tested different versions of the first stage of the survey (the screener) by sending some respondents a short, stripped down survey that asked the bare essentials while other respondents were sent a longer, more comprehensive survey designed to engage them in the purpose of the survey. A third group of respondents received a screener survey that was a compromise between the stripped down approach and the engaging approach; it was a survey with the bare essentials plus some additional demographic questions. All versions of the screener asked for the names of all children in that

household. However, to accommodate sensitivity concerns raised by cognitive interview respondents, the surveys also gave the option of providing children's initials or nicknames.

Results indicate that about 21.1 percent of respondents from the 2009 NHES pilot study's shortest version of the screener used children's initials or skipped children's names entirely. However, respondents to the longest, engaging version of the screener used children's initials or skipped children's names entirely 38.3 percent of the time. Among the group that received the medium-length screener survey, 26.7 percent used children's initials or did not provide child's names.

This paper's goal is to explore the differences in the response patterns for child's name. We will be comparing differences in demographics of the respondents. The characteristics explored will include age, education, income, locale, and child's age and gender.

Selection Bias In The Deliberative Poll: Balancing Representation Of Opinion And Representation Of Socio-Demographics.

Sean J Westwood, Stanford University (seanjw@stanford.edu); Nuri Kim, Stanford (nurikim@stanford.edu)

The methodology of the Deliberative Poll is unique for a number of reasons, but in assessing the balance and representation of each poll, a key issue is the difference between those who come for the face-to-face event, those contacted in the initial poll who opt to not attend, and the public-at-large. Traditionally this analysis is performed through a series of variable-level t-tests and Chi-2 tests between participants, non-participants and the public. While useful in identifying differences, this analysis fails to fully engage the consequences of the observed differences on participation. We expand the traditional analysis through simple binary logistic regression models, with participation as the response variable and socio-demographic variables as the predictors. We find several socio-demographic traits significantly predict poll participation, specifically, race, education, and gender. These findings suggest that there are systemic selection biases imbedded the current methodology of the Deliberative Poll. Such results suggest that while great care is taken in the assembly of deliberators, oversampling and undersampling occurs across multiple socio-demographic categories across multiple polls. However, when the sample was weighted to match the census data, measures of opinion retained similar means that were statistically insignificant in their differences for all of the initial attitude indices. There were marginal differences for some of the opinion indices at time two, but these differences were quite small and did not change the interpretation of the results. This suggests that although there are problems in sample selection in the Deliberative Poll overall opinion and opinion change is generally not significantly influenced, which further legitimates the process of deliberative opinion assessment through the Deliberative Poll.

Child Care Policies and the Gender Gap: A Policy Response Model.

Dan Cassino, Fairleigh Dickinson University's PublicMind Poll (dcassino@fdu.edu); Yasemin Besen-Cassino, Montclair State University (beseny@mail.montclair.edu)

The gender gap in politics has been a persistent factor in American national elections for more than 30 years. Explanations for it have generally centered on supposed inborn or socialized differences in how men and women view issues and political figures, with women being seen as more concerned with social issues, and men more concerned with economic issues. In recent years, this view has come under fire from policy response models, which argue that the differences between men and women may result from differences in how society treats individuals based on gender. For example, women may prefer more liberal parties and politicians because they have greater need for the results of more liberal policies in the

form of health care and child care policies. Such views are given additional credence by the lack of a gender gap in many nations with more liberal social policies.

However, in order to determine whether or not this is actually the cause of the gender gap in the United States, we must take a comparative approach. If support for more liberal parties in the US among women is a function of social policies, there should be less of a gender gap in nations with more liberal social policies. To that end we examine support for parties in OECD countries by gender, the social policies of that country and the ideology of that party. Our results show that in countries with more liberal social policies, especially those relating to the burden of having children, liberal parties are expected to have less support among women than among men. Only in countries with an extremely high burden of having children – such as the US and Australia – is there expected to be a gender gap such that liberal parties have more support among women than among men.

Do the Reluctant and the Eager Responders Differ in the Medical Expenditure Panel Survey and Do They Impact Survey Estimates?

David Kashihara, Agency for Healthcare Research and Quality (david.kashihara@ahrq.hhs.gov)

The Medical Expenditure Panel Survey (MEPS) is a nationally representative study of health care use and expenses, sponsored by the Agency for Healthcare Research and Quality (AHRQ). The set of households selected for each year's panel of the MEPS is a subsample of households participating in the previous year's National Health Interview Survey (NHIS) conducted by the National Center for Health Statistics. The MEPS, like a number of federal surveys, focuses much time and expense to maximize survey response rates. As a result, some who initially refuse to participate are converted to a response using multiple survey protocols and/or increasing levels of persuasive follow-up efforts. This study examines the household, demographic, and socio-economic characteristics of eager versus reluctant respondents. We also investigate how eager versus reluctant respondents affect the quality of data by comparing key survey estimates from the two groups as well as item missingness rates for both groups. This paper also makes use of both NHIS and MEPS paradata to examine e.g., interviewer level of effort, number of broken appointments, and NHIS partial versus complete interview status for the two groups.

Public Health Accreditation in the State of Ohio: Developing a Methodological Model for Accreditation.

T. David Jones, Wright State University- Center for Urban and Public Affairs

(david.jones@wright.edu); Brittany D. Sumler, Wright State University-Center for Urban and Public Affairs (sumler.2@wright.edu)

Public health organizations in the State Ohio are moving towards a model of voluntary Public Health Accreditation, designed to provide standards for local health departments as it pertains to methods of data collection, strategic planning, and building of community partnerships in the arena of community health. One core component of this voluntary Public Health Accreditation is the development of a Community Health Improvement Plan, which includes multiple data collection methods, including modified versions of the CDC's BRFSS, YRBS, and National Survey of Children's Health. Beginning with the implementation of the Public Health Accreditation program in 2011, counties will be required to document data collection methods and regularly conduct Community Health Improvement Plans.

Wright State University's Center for Urban and Public Affairs (CUPA) is currently working with 17 local health departments in the State of Ohio to develop a Community Health Improvement Plan. Implementation of the survey instruments have posed many methodological challenges in some Ohio counties, based upon the conservative nature of some school boards and community leaders who believe that issues such as youth sexual activity, alcohol and drug use are not problems in their county. To further

complicate the issue, the Public Health Accreditation Board has specific standards and protocols that require counties to mirror national instruments and design strategic health plans that utilize these national benchmarks, this making the exclusion of these topics problematic.

This paper will discuss the process that CUPA has employed to gather consensus on methodology from a wide variety of community organizations that have never participated in a data collection process. The paper will also address the impact of methodological changes in some counties as it pertains to comparability of data with the state and national instruments, and the subsequent impact on obtaining Accreditation. Recommendations for improving this process will also be discussed.

Exploring Trends and Patterns of Nonresponse: Results from the supplementary German ISSP surveys, 1986-2008.

Volker Hüfken, University of Duesseldorf, Institute of Social Sciences (huefkenv@phil-fak.uni-duesseldorf.de)

As can be taken from an overview about nonresponse in International Social Survey Programme (ISSP) module response-rates vary about 50-80% (Couper and de Leeuw 2003). It is to be assumed that different technical procedures in investigating the field are responsible for much of such varying response rates. First, it seems to be relatively clear that the facultative presence of the interviewer is important for further participation (Sheatsley 1984). But little information is currently available about reasons, why respondents themselves refuse to continue with the supplementary interview. Hence, four groups of variables (socio demographics, data on interviewing and cooperation, behavior variables, and attitudes) are chosen for an analysis of supplementary nonresponse. Furthermore, there is yet little knowledge about the stability of nonresponse patterns across years. Our aims are: (1) Finding out whether there are trends or regularities in ISSP supplementary nonresponse (characteristics of each individual surveys) (2) Examination of influences which different surveys as such may have had on rates of nonresponse (3) Combining our various findings the construction of some more comprehensive model is aimed at. We use variables from the cumulative ALLBUS/GGSSs 1986-2008.

Friday, May 13, 2011

3:20 p.m. - 3:30 p.m.

Demonstration Session 2

Demonstration 2A.

Multi-media Ethics Education: Tools to Assist in Preparing Undergraduates to Collect Data from Human Subjects.

**Alexandra Cooper, Social Science Research Institute, Duke University (cooper@duke.edu);
Dwayne Dixon, Duke University (dwayne.dixon@duke.edu); Lorna Hicks, Office of Research
Support (lorna.hicks@duke.edu)**

Increasingly, higher education is both calling on and providing opportunities to undergraduates to conduct independent, mentored research. Many students pursue research into various aspects of human behavior, culture, and organization and, in conducting their research, collect data from individuals about their attitudes, beliefs, and/or behaviors. Novice researchers require careful guidance as they plan and

conduct research involving human subjects (particularly when those projects involve subjects from other cultures or who are part of vulnerable populations, including for example children, the mentally ill, the homeless, or illegal immigrants). In this demonstration we present and discuss examples drawn from a group of multi-media pedagogical modules we have developed with funding from Duke University's Center for Instructional Technology. Designed to facilitate students' critical thinking as they consider how to conduct what are often delicate – but also challenging and important research projects – these modules serve to assist both faculty and others who guide students in preparing to conduct research involving human subjects and the students themselves as they work to develop and implement their research protocols. By portraying applications of key concepts that guide ethical research in real-world settings, the modules are intended to prepare students to engage actively in mentored research projects - preparation that is essential to anyone collecting attitudinal or behavioral data from human subjects and opening themselves to the considerable responsibilities such projects entail.

Friday 3:35 p.m. - 3:45 p.m.

Demonstration 2B.

Building a Web-Based Tool to Aid in Training Users of a New Survey Monitoring System.

Erica Saleska, RTI International (esaleska@rti.org); Neelima Kunta, RTI International (nkunta@rti.org); M. Rita Thissen, RTI International (rthissen@rti.org)

Computer-assisted training has proven valuable for instructing interviewers, data managers and others who use survey-related systems. RTI International has been working with the U.S. Census Bureau to develop and deploy an application called the “CARI Interactive Data Access System,” a complex and new approach to behavior coding, quality assurance coding, and coding interviewer performance. CARI (Computer Audio-Recorded Interviewing) is a technique of capturing audio and/or image files during the interviewing process. Because this web-based system is new to its users, coder training is needed. Due to the number of coders to be trained, an online, self-paced method was requested for introducing coders to the software.

There are many varieties of online training, which differ in strengths. Choices range from textual help files to elaborate video presentations. At one extreme, training tools may require extensive preparation and a big budget, and may be difficult to modify once deployed. At the other extreme, very simple tools may not be effective. Usage may also vary, as some require the presence of an instructor or specialized equipment. The literature indicates that self-paced study can be highly effective, if adequately planned.

With these options and challenges in mind, the team implemented “eTraining” as a sequence of training pages through which the coder can learn independently, which enables coders to view presentations at their own pace, and gives them confidence in their ability to perform their jobs. This demonstration will present the “eTraining” application, which offers courses for all modules of the system, providing text and images with voice-over to reinforce content. A quiz after each course allows coders to assess what they have learned, and overall certification is offered at the end. “eTraining” offers flexibility in training newly-hired coders or can act as a refresher course for those who return to coding after an absence.

Friday 3:50 p.m. - 4:00 p.m.

Demonstration 2C.

A Demonstration of CARI Tools: Increasing the Efficiency of Survey Operations and Improving Data Quality.

Wendy Hicks, Westat (WendyHicks@Westat.com); Rick Dulaney, Westat (RickDulaney@Westat.com); Brad Edwards, Westat (BradEdwards@Westat.com)

As the sophistication of Computer Audio Recorded Interviewing (CARI) technology continues to improve (e.g., the CARI capabilities within Blaise [™] 4.8.2), the range of feasible applications within survey research also grows. This demonstration covers three different capabilities achieved through the linkage of CARI functionality and a web-based coding tool.

Linking CARI technology and a coding tool provides a means for increasing the efficiency of specific survey operations as well as improving data quality. The web-based coding tool uses CARI data to facilitate: 1) the identification of potential interview falsification by linking CARI data with other paradata such as contact history information; 2) the assessment of the potential for measurement error due to question design and the administration process, and; 3) more accurate capture and editing of survey response data. The coding tool pulls in the audio files collected during an interview, sorts and delivers those files according to the specific rules set for each function, and displays the 'live' screen shots captured during the interview along with the keyed response data and other survey paradata as appropriate. Within the coding tool, survey staff can generate a set of monitoring and analytical reports tailored to each function. Discussion will also point out some of the design considerations and implications for the system set-up and implementation. Additional efficiency and quality functions possible with the linkage of CARI and coding tools that are currently under development will also be discussed.

Friday, May 13, 2011

4:15 p.m. - 5:45 p.m.

Concurrent Session D

Are Measurement and Item Nonresponse Differences a Problem in Web & Mail Mixed-Mode Surveys?

Do Mail and Web Produce Different Answers? Mode Differences in Question Responses and Item Nonresponse Rates.

Morgan M. Millar, Washington State University (morgan_millar@wsu.edu); Don A. Dillman, Washington State University (dillman@wsu.edu)

Mixed-mode surveys have become increasingly important in today's changing survey climate. Recent research has examined the feasibility of combining mail/paper and web modes of response in a way that encourages a majority of respondents to respond via web, and then uses mail as a follow-up method to improve overall response rates. However, combining multiple modes in a single study begs the question of whether the two modes elicit similar types of responses. A good deal of prior research has demonstrated modes such as telephone and paper/mail tend to produce different types of responses. However, we know much less about mode differences between more similar modes such as mail and web, which are both self-administered and rely on visual stimuli. There is some evidence to suggest that maintaining a common visual layout for both paper and web questionnaires can help eliminate mode differences in responses. However, there has been little systematic examination of whether such differences arise when sampled individuals are randomly assigned to one mode of response or the other. In this paper we use data from two mail/web mixed-mode experiments, which were conducted in 2009 on a population that is completely accessible by both postal mail and email, to examine whether there are mode differences in question responses as well as rates of item nonresponse. Both of these studies contained both mail-only and web-only response panels, to which individuals were randomly assigned. The web and paper versions of the questionnaires were virtually identical in visual layout, allowing us to determine whether the two modes produce differences in responses when the visual stimuli are presented in similar fashion. Preliminary findings show there are no significant differences in rates of item nonresponse across modes, and we hypothesize that there are minimal differences in responses across modes as well.

Comparing Numeric and Text Open-End Responses in Mail and Web Surveys.

Jolene Smyth, University of Nebraska-Lincoln (jsmyth2@unlnotes.unl.edu); Kristen Olson, University of Nebraska-Lincoln (kolson5@unlnotes.unl.edu)

Recent research has shown that it is possible to improve coverage and reduce nonresponse by mixing web and mail data collection modes. Generally it is assumed that because both web and mail are visual modes, they will produce comparable data, but little empirical research has examined this assumption. Now that surveyors have the ability to relatively easily mix web and mail modes, we need to know whether or not measurement is, in fact, comparable. Open-ended questions (number boxes and text boxes) seem especially problematic because respondents often have full control over how they answer them; answers to these questions are not structured and guided in the same way as closed-ended questions with limited response options. This is especially true in self-administered surveys where there is

no interviewer to probe, ensure that the desired type of answer is provided, or convert the respondent's answer into the desired format. In this paper, we examine item-nonresponse rates, response distributions, and the effects of questionnaire design features on a variety of open-ended questions from the Quality of Life in a Changing Nebraska (QLCN) survey. Where possible, we also examine the effects of design changes by subgroups (e.g., by respondents expected to be more or less familiar with each mode). The QLCN was conducted between July and September, 2009 (N=566; AAPOR RR1 = 46%) and contained eleven open-ended boxes where numeric information was requested and two open-ended boxes where descriptive text was requested. In addition to being randomly assigned to either the web or mail mode, respondents were randomly assigned to one of two questionnaire design treatments. Questionnaire design experiments include a question order experiment, small versus large box size on both numeric and text questions, and presence versus absence of answer box labels on numeric questions.

Item non-response in Web and Mail Responses to General Public Surveys.

Virginia M. Lesser, Department of Statistics - Survey Research Center

(lesser@science.oregonstate.edu); Lydia Newton, Survey Research Center-Oregon State University (newton@science.oregonstate.edu); Daniel Yang, Survey Research Center-Oregon State University (dyang@science.oregonstate.edu)

Increasingly, the type of mode used to collect survey data is considered when determining how to best maximize response rate for the population and various subpopulations. Many studies are using several data collection modes as a method to increase responses from targeted demographic groups. In addition, surveys that are repeated over time may have used one mode initially in data collection, and then switched to a different mode as more complete frames become available for other modes. Our objective is to compare item nonresponse and the responses measured across different modes used for data collection. We wish to investigate if and how mode is influencing responses and item nonresponse in a questionnaire. In a study of the general population of Oregon households, a telephone survey was used biennially for 10 years to obtain opinions on transportation issues in the State. Experiments embedded in the 2006, 2008, and 2010 studies were conducted to compare measurements and item nonresponse collected by telephone, mail, Web, and mixed-mode approaches. We compare whether the different data collection methods result in different responses to the survey questions. Adjustments were made for the demographics obtained from the completed respondents for each mode to evaluate the impact of the demographic groups responding by different modes on the response. In addition, studies were also conducted for two specialized populations, individuals with hunting licenses and individuals owning a marine boat license. Experiments were conducted for each of these populations comparing responses to a mail, Web, and a mixed-mode method. Item nonresponse across these modes is also compared.

Determinants of Web and Mail Item Nonresponse in Address-based Samples of the General Public.

Benjamin L. Messer, Washington State University Social and Economic Sciences Research Cent (aluvial@gmail.com); Michelle L. Edwards, Washington State University Department of Sociology (michelle.edwards@email.wsu.edu); Don A. Dillman, Washington State University Social and Economic Sciences Research Cent (dillman@wsu.edu)

Item nonresponse in self-administered modes such as web and mail can be a major problem affecting survey data quality and, in some cases, may be as severe as unit nonresponse. Moreover, little is known about the most important sources of item nonresponse in web and mail household surveys. In this paper, we assess item non-response differences by web and mail modes, questionnaire attributes, and respondent demographics in three general public household surveys. For each of the three surveys, conducted in the northwestern U.S. in 2007, 2008, and 2009, respectively, we used address-based

sampling with the U.S. Postal Service's Delivery Sequence File, a near comprehensive database of postal addresses, and employed postal mail methods to send all contacts. Sampled respondents in each survey were presented with a) a mail-only response option, b) a mail response option with a web follow-up sent two weeks later, or c) a web response option with a mail follow-up sent two weeks later. The web and mail questionnaires in each survey were designed very similarly in order to minimize and control for effects from visual design and layout. We attempt to develop a perspective on total item nonresponse in using web and mail survey modes and the different patterns in which item nonresponse occurs in web and mail questionnaires, such as by question type (e.g. factual, attitudinal, behavioral) and format (e.g. nominal, ordinal, multi-item, open-end, etc.). Also, using data from two of the statewide household surveys, we are able to analyze item nonresponse differences by respondent demographic characteristics, including gender, age, education, race, and income. This paper serves to explicitly quantify and describe item nonresponse differences and the sources of those differences, and to identify potential ways of reducing item non-response in web and mail modes of data collection.

Item Non-response in a Client Survey of the General Public.

Glenn D Israel, University of Florida (gdisrael@ufl.edu)

Client surveys of the general public represent a significant segment of studies conducted by organizations and contact information (e.g., postal address, e-mail address, and telephone number) is often incomplete. Consequently, survey administrators must match the survey procedures to the available contact information to minimize coverage error. When there are multiple types of contact information, mixed-mode procedures can be to improve response rates and reduce costs. The quality of data from self-administered client surveys can, however, be affected by item non-response but little research has focused on assessing how implementation procedures and response mode affect the extent of item non-response. In this paper, I examine item non-response by clients who have received educational information from a government agency, the Cooperative Extension Service. The analysis uses survey responses from clients in three strata (based on providing contact information for postal address only, e-mail only, and both mail and e-mail). For clients in the strata with both mail and e-mail addresses, four experimental groups were created, including two mixed-mode groups, a mail only group and an e-mail only group. I focus the analysis on non-response differences between mail and web modes, question format and question type. Additional analysis is conducted to explore the role of response mode and client attributes (including gender, age, race and education) on the item non-response rate. The analysis shows that some item-by-item comparisons between web and mail responses have significant differences. Survey responses via the web have a significantly lower overall item non-response rate than those via postal mail. These results suggest that as long as the web and mail responses are substantively equivalent, efforts to increase the proportion of web responses can improve data quality and reduce costs at the same time.

Correlates & Impacts of Unit Nonresponse.

Longitudinal Response Rates and the Economy.

Tricia McCarthy, NORC at the University of Chicago (mccarthy-tricia@norc.org); Hildie Cohen, NORC at the University of Chicago (cohen-hildie@norc.org)

The recent economic conditions in this country have greatly affected the populace of the United States. Studies like the National Longitudinal Survey of Youth sister studies allow us to trace members of this populace throughout this economic downturn, and track the tangible life changes they might face because of it. Just as it is important to collect high quality data that will allow us to track the changes in

respondents' lives, it is also important to track the effect the current economic situation has had on the actual data that are being collected. The NLSY79 and NLSY97 studies are in their 24th and 14th rounds respectively, and together provide a national representation of different generations that have been tracked through their labor market experience, and beyond into their personal lives. By looking at response rates concurrently with economic factors such as unemployment, the poverty rate, and other factors, we will be able to place response rates within an overall economic setting. Analyzing these variables along with qualitative responses from respondents revealing reasons cited for participation, respondent incentives, employment status, and age could provide valuable insight into how the economy could be affecting response rates, specifically in longitudinal surveys like the NLSY's.

Should First Time Sample Members in Longitudinal Establishment Surveys Receive Special Attention?

Geraldine Mooney, Mathematica Policy Research (gmooney@mathematica-mpr.com); Elisha Smith, Mathematica (esmith@mathematica-mpr.com); Cathie Alderks, SAMHSA (cathie.alderks@samhsa.hhs.gov)

For more than a decade, Mathematica has been conducting an annual census of all known substance abuse treatment facilities through the National Survey of Substance Abuse Treatment Services (N-SSATS) sponsored by the Substance Abuse and Mental Health Services Administration (SAMHSA). Although the annual number of facilities surveyed remains fairly stable at about 17,000, roughly 10 percent of facilities become ineligible annually and are dropped, with roughly the same percentage of new facilities being added. Although the response rate has also remained fairly stable, around 95 percent, achieving that response rate is becoming increasingly difficult. When looking for facilities with the lowest response rates, we discovered that new facilities were much slower and less likely to respond and, throughout the field period, their response rate ranged from 9 to 19 percentage points below that of previously surveyed facilities.

This led us to embed an experiment into the 2010 N-SSATS survey. We customized a cover letter for new facilities that acknowledged their status as a new facility, welcomed them to the survey, and promised a call from us within the next two weeks to answer any questions they might have about the survey or their participation. We then randomized the new 2010 facilities into either the control or experimental group. The only difference in data collection procedures between the two groups was the customized cover letter sent to new facilities and the follow-up telephone calls. This paper discusses the experiment and reports our findings with respect to possible cost savings and changes in procedures.

Nonsampling Errors in the Ohio Family Health Survey's Cell and Landline Telephone Samples.

Bo Lu, Ohio State University (blu@cph.osu.edu); Ronaldo Iachan, ICF Macro (riachan@icfi.com); Timothy Sahr, Ohio State University (Timothy.Sahr@osumc.edu); Seth H. Muzzy, ICF Macro (smuzzy@icfi.com)

The Ohio Family Health Survey (OFHS) is a telephone survey of the health and health insurance status of adults and children in Ohio. The OFHS supplies information of insurance coverage, health status, and health care utilization for several population subgroups in four regions and some metropolitan counties. The latest survey iteration (2010) benefited from the large 2008 survey cycle to include two telephone sample components, a landline sample (LL) and a cell sample. This paper examines the varying response rates and assesses coverage errors in the two samples and across two survey cycles.

Preliminary comparisons of the demographics in the two samples show that:

*35.9% of the cell sample respondents are young (18-35) vs. 9.5% of the LL sample

- *33.7% of the LL sample are old (aged 65+) vs. 12.6% of the cell sample
- *47.8% of the cell sample are male vs. 37.0% of the LL sample
- *76.1% of the LL respondents own their residence versus 63.6% of the cell sample

Response rates were low particularly for the cell sample. We examine whether this phenomenon is a reflection of the socio-political climate of the state in the 2010 health care reform environment. We examine objective survey data and related data that may shed light on these connections.

Refusal Conversion in REACH U.S.: A Targeted Approach.

Lee J Fiorio, NORC @ the University of Chicago (fiorio-lee@norc.org); Kanru Xia, NORC @ the University of Chicago (xia-kanru@norc.org); Ashley Amaya, NORC @ the University of Chicago (amaya-ashley@norc.org)

Respondent refusals characterize all surveys, and they occur for a variety of reasons spanning from lack of interest to mistrust of outsiders. To gain a better understanding of respondent refusals and how to mitigate them, we conducted a refusal conversion letter experiment within the Racial and Ethnic Approaches to Community Health across the U.S. (REACH U.S.) Risk Factor Survey. REACH U.S. is the cornerstone of CDC's efforts to eliminate racial and ethnic health disparities among African Americans, Native Americans, Asian Americans, and Hispanics living in 28 communities throughout the country. The REACH U.S. survey relies on an address-based sampling approach coupled with multiple modes of data collection including telephone, mail, or in-person interviews. The refusal conversion letter experiment was limited to the telephone data collection mode. In particular, we sought to determine whether sending a refusal conversion letter that addressed a respondent's specific concerns and reasons for refusing to participate would increase refusal conversion rates and ultimately interview completion rates among those initially refusing to participate.

The REACH U.S. Phase 3 sample was divided into control and experimental groups of approximately equal size. The control group received the standard REACH U.S. refusal conversion letter that addressed several common reasons that respondents refuse. The experimental group received a refusal conversion letter that addressed a respondent's specific concern or reason for the refusal. After mailing the letters, an interviewer specially trained in refusal conversion techniques telephoned the respondent and attempted to complete the interview. Analyses will compare the two groups on conversion rates, interview completion rates, recontact rates, advancement rates, and overall yield rates.

The Integrated Uses of Surveys, Administrative Records, and Census: The Case of the Census in School and Partnership Research in Informing the 2020 Census.

Zhiwei Zhang, ICF International (zzhang@icfi.com); Christopher Spera, ICF International (cspera@icfi.com); Megan Kindelan, Census Bureau (megan.c.kindelan@census.gov); Peter Miller, Northwestern University (p-miller@northwestern.edu); Ronald Szoc, ICF International (rszoc@icfi.com); Diana Boyd, ICF International (DBoyd@icfi.com); Tasha Boone, Census Bureau (tasha.r.boone@census.gov)

For the 2010 Census, the U.S. Census Bureau's Census in Schools (CIS) program provided educators in all schools (N=118,000) with resources to teach the nation's students about the ways census data can be used in the classroom and delivered the message of the importance of participation to students and their families. Meanwhile, hundreds of thousands of national and local partnerships were formed to facilitate and encourage the individuals' participations in the decennial Census across the nation. This study, supported by the U.S. Census Bureau and conducted by a team at the ICF International, aims to evaluate the effectiveness of the 2010 Census in Schools (CIS) and Partnership programs. One key research question is: Did the CIS program and the activities performed by the partners of the Census Bureau have

any discernable impact on the mail-back response rate? We design and implement an integrated analytic approach with the proposed integrated utilizations of multiple primary and secondary data sources, including (1) a focus group study on informants from K-12 schools fielded by the ICF; (2) a census of the national and local partners working for the 2010 Census administered by the Census Bureau; (3) a representative survey of local partners fielded by the Census Bureau; (4) a focus group study of national partners fielded by the ICF; and (5) the utilization of the planning database of 2010 Census. Variations of response rates and the explanatory covariates across the communities will be especially investigated. Challenges and how research may be responsive to intercensal operational needs and informing the 2020 Census will be discussed.

Dynamics of Consumer Confidence.

Consumer Confidence and Consumer Purchase Behavior During the Economic Crisis of 2008.
Max Kilger, Experian Simmons (maxk@smrb.com); Charles Palit, University of Wisconsin (cdpalit@wisc.edu)

The U.S. economic crisis in fourth quarter 2008 is likely the most severe economic downturn since the Great Depression. The sheer magnitude of the event has changed the consumer landscape in ways not seen in modern times. This paper examines the relationship of a set of consumer confidence measures to consumer purchase behaviors during three contiguous temporal windows spanning the pre-, concurrent and post-crisis periods.

Questions that are addressed include an examination across time of the relationship of three distinct dimensions of consumer confidence – confidence in recent personal economic circumstances, confidence in near-term future personal economic circumstances and confidence in the near-term future of the U.S. economy – to a wide set of specific consumer purchase and product ownership attitudes and actual behaviors including but not limited to the entertainment, travel, automotive and health sectors.

One of the early conclusions of the study is that while much of the media attention on the crisis was focused on events on or around fourth quarter 2008, consumer confidence trends reveal that the American public was perhaps somewhat more aware of the magnitude of the economic downturn in the months preceding the focal point of the crisis than might have previously been thought.

This paper undertakes an investigation of these issues and more utilizing a very large, nationally representative consumer database that collects data on consumer purchases, attitudes and media behavior on a weekly basis. Utilizing this time-series database we will examine the relationship of consumer confidence to indicators of consumer purchase behavior during times of economic crisis along a timeline that brackets one of the most turbulent economic periods in recent history.

Watching for Signs of a Rebound Monitoring Consumer Sentiment 2009 to 2011
Allan Rivlin, Partner, Hart Research Associates, LLC and, Janet L. Streicher, Senior Vice President, Market and Strategic Research, Citibank, NA and. Janet L Streicher, Citibank (janet.streicher@citi.com)

This Paper discusses results from a series of quarterly national consumer polls conducted by Hart Research for Citibank starting in September 2009. We have conducted four surveys to date but the paper will include six by the time of the AAPOR conference in May 2011.

The full length 20-minute telephone RDD surveys with cell-only panel have probed consumer's economic experiences and expectations as the economy has hovered at low levels throughout most of 2010. We have looked closely for signs of a rebound among important groups, defined by region, income, education level, family structure and other factors, but at this point, we can only hope that groups other than the most affluent register an uptick in time for the Annual Meeting. (The next quarterly survey fields just after this abstract deadline.)

Each survey has included 2,000 interviews of the general public, except for the survey in March 2010, that was 2,000 interviews, but 1,400 were of the general public and 600 were of investors.

The paper will report how the Citi Economic Index (a combination of eight questions covering current assessments of the economy and expectations for the future) has varied for important groups. In addition, we will discuss how Americans have adapted to the economic downturn, including changes to their spending, saving and borrowing behaviors. Specific segments of consumers that have experienced a deeper and longer downturn (e.g. mid-westerners and working mothers) as well as those segments that have showed the greatest resiliency (the young and minorities) will be highlighted.

Consumer Confidence: Concurrent and Predictive Validity.

Julie E. Phelan, Langer Research Associates (jphelan@langerresearch.com); Gary Langer, Langer Research Associates (glanger@langerresearch.com)

This paper presents an overview of our examination of the concurrent and predictive validity of the Bloomberg Consumer Comfort Index.

For 25 years, we have assessed consumer sentiment on a weekly basis by asking a random sample of 250 respondents' to rate the national economy, their personal finances and the buying climate. This has resulted in over 325,000 interviews tracking perceptions of current economic conditions. In the past, we have reported on how our weekly Consumer Comfort Index compares to the two prominent monthly surveys of consumer confidence, the Conference Board Consumer Confidence Index and the University of Michigan Consumer Sentiment Index, finding that between 1985 and 2003 the indices tracked each other closely and all significantly correlated with several key economic indicators.

In this paper, we update and extend our examination of the utility of the Bloomberg CCI by assessing whether the index is a leading indicator of several key monthly economic measures, including the Dow Jones Industrial Average, GDP, the unemployment rate and revolving consumer credit, among others. Moreover, taking advantage of the unique weekly format of our index, we examine its concurrent and predictive relationship with weekly initial and continuing unemployment claims. Throughout we discuss the analytical challenges of assessing the true relationship of data that contains a strong time-trend component.

An Analysis of the Relationship Between Public Sentiment and Corporate Performance During A Crisis As Illustrated by the British Petroleum Gulf Coast Crisis.

Charles Palit, Univ. of Wisconsin (cdpalit@wisc.edu); Max Kilger, Experian Simmons (maxk@smrb.com)

One of the key items thought to significantly govern corporate ethical behavior is the reaction of the public to actions and positions taken by a corporation. When the public disapproves of a corporation's actions, the public sometimes expresses its opinions by boycotting the products the corporation produces. This

may entail switching to products produced by competitors or simply not buying the product. Boycotting actions or campaigns of this type are thought to have a harmful effect on the corporation's financial performance.

During the US Gulf Coast crisis, British Petroleum was the subject of a consumer backlash. We are concerned with how the backlash developed and played out across time. What was the relative importance of the various negative events to the strength of the backlash? How effective was the consumer backlash? Was there a differential consumer response to BP that depended upon various key factors such as demographics, environmental attitudes or even strength of consumer loyalty to British Petroleum products? Did the various governmental statements and actions taken during the crisis have an effect on public sentiment towards BP products?

This paper undertakes an investigation of these issues utilizing a large, nationally representative consumer database containing data on consumer purchases, attitudes and media behavior on a weekly basis. Utilizing this database we link various real world events related to the Gulf Coast disaster with week-by-week reports of the consumption of BP products by consumers to assess whether or not there appeared to be an effective effort by American consumers to voice their opinion on the disaster. We examine the effects of a number of factors such as consumer attitudes towards the environment and exposure to mainstream news media as well as developments in the crisis and federal government statements and actions on the reported consumption of British Petroleum products.

Election Polls - Mechanics.

Are Different Approaches to Determining Likely Voters Better Able Handle Early Voters and First-Time Voters?

Jocelyn Kiley, Pew Research Center for the People & the Press (jkiely@pewresearch.org); Leah Melani Christian, Pew Research Center for the People & the Press (Ichristian@pewresearch.org); Michael Dimock, Pew Research Center (mdimock@pewresearch.org)

One of the key challenges in pre-election surveys is determining the likely electorate. Substantial research has shown that people over-report their likelihood to vote, so approaches that ask a single question about intention to vote often result in a more reporting that they will vote than actually turnout. This paper will explore various ways of determining likely voters and compare the profile of likely voters from these various approaches to data about actual voters from the Census Bureau's Current Population Survey and exit polls by the National Election Pool.

In particular, this paper will aim to address two developments that influence pollsters' ability to determine likely voters. First, the proportion of voters casting their ballots before election day has risen rapidly as many states have adopted more expansive early voting laws. A key issue is how to handle the increasing number of respondents who say they have already voted in the weeks leading up to the election, which may also be susceptible to over-reporting. Moreover, the growing number of early voters may have a greater influence on the representativeness of a survey's likely voter sample if their method of determining likely voters is based on a set proportion of voters who are estimated to turnout, particularly if voters who plan to vote on election day are not represented in their accurate proportion. This paper will also explore how well different approaches to determining likely voters can respond when early signs in an election may indicate that more voters who have never voted say they intend to vote. Because of the approaches used by many pollsters, it is often easier for young first-time voters, to pass as likely voters than it is for older first-time voters.

Incumbents, Challengers and Unallocated Votes: An Examination of Late Breaking Voter Decisions in the 2010 Midterm Elections.

Christopher Paul Borick, Muhlenberg College Institute of Public Opinion (cborick@muhlenberg.edu); David G. Wegge, St. Norbert College Survey Center (dave.wegge@snc.edu)

One of the most interesting questions regarding pre-election polls in the United States relates to the way undecided voters break at the end of a campaign. In close elections the division of undecided voters can determine the candidate who prevails in the race. One particular aspect of the allocation of undecided voters involves the presence of an incumbent in a race. Perhaps the most widely cited theory regarding the impact of incumbency is that undecided voters behave differently when an incumbent is in a race. In essence, if the incumbent has not sealed the deal with a voter before the election, it is unlikely that they will be able to gain their support during the final days of a campaign. While this “incumbent rule” is widely used in media accounts of campaigns, a number of studies from recent elections have shown little evidence that incumbents receive a smaller portion of unallocated votes than their challengers. These results stand in contrast with studies between 1988 and 2002 that showed challengers winning a solid majority of votes unallocated in final election polls.

Given these findings it is interesting to consider the outcomes observed in the 2010 midterm elections. In particular, did the large number of incumbent defeats in this cycle correspond with resurgence in the performance of challengers in gathering undecided voters, or did the success of challengers simply reflect a scenario where the final division of undecided voters did not affect the outcomes of the races? In this paper we address these questions through an empirical examination of pre-election polls in 2010 races in which incumbents were on the ballot. The findings will provide insight into the interpretation of pre-election polls in which the percentage of undecided voters may be essential to the determination of the ultimate victor.

Bias in Legislative Generic Ballot Questions: Building a Better House Election Forecast.

Dan Cassino, Fairleigh Dickinson University's PublicMind Poll (dcassino@fdu.edu); Peter Woolley, Fairleigh Dickinson University's PublicMind Poll (woolley@fdu.edu); Krista Jenkins, Fairleigh Dickinson University's PublicMind Poll (kjenkins@fdu.edu)

While predictions about the outcomes of races for high offices, such as governorships, Senate seats and the presidency in the United States are based on polls about the specific candidates in the running, predictions about the composition of the most numerous elected branch, the House of Representatives, are generally based on generic ballot polls. In such polls, respondents are asked which party they will support in the election for the House, rather than being asked which candidate they will support. While there are a number of logistical and theoretical reasons for such an approach – the low salience of House elections, low name recognition of the candidates, the large number of non-competitive seats, the difficulty and cost of matching respondents to candidates – such measures may lead to biased predictions depending on the circumstances of the election, as they did in the prediction of the composition of the House in the 2010 midterm elections.

In the actual election, Republicans won 243 seats, to Democrats' 192 (awaiting some results), a margin which exceeded nearly all of the predictions based on the generic ballot questions. Using a survey experiment embedded in an RDD survey of voters in two states, we show that these predictions were biased in favor of the Democrats, and argue that this bias is the result of candidate selection effects. In states where the Democrats were unable to recruit high quality candidates (as in most states in this election cycle) the generic ballot favored Republicans. In the (very few) states where Republicans were unable to recruit high quality candidates, the generic ballot was biased in favor of the Democrats. We

conclude by suggesting ways to adjust the algorithms used to convert generic ballot measures into electoral outcomes to compensate for these effects.

Nothing Generic About This: Using Generic Ballots to Predict State Legislative Elections.
Andrew Smith, UNH Survey Center (andrew.smith@unh.edu)

The Gallup Poll has used a generic ballot question since the 1950s to predict the national vote in Congressional elections. It has a track record of surprising accuracy, although its 2010 projection caused considerable debate among pollsters and pundits when it overestimated Republican support. But a 6 decade record of mostly accurate projections makes this a useful measure for political scientists in understanding national elections. This paper examines how well a generic ballot question predicts the outcome of state legislative elections.

New Hampshire is an especially good place to test the effectiveness of using generic ballot questions as its entire state House of Representatives and Senate are elected every two years. Furthermore, the size of the two bodies is quite different -- 400 rather anonymous people are elected to the House and 24 much better known people are elected to the Senate. This study looks at the ability of a generic ballot to predict both the party that wins the majority of seats as well as the magnitude of that majority. Data for this study come from final pre-election polls conducted in New Hampshire prior to the 2000, 2002, 2004, 2006, 2008, and 2010 state legislative elections. This time frame is quite interesting as majority control of both chambers switched in 2006 and again in 2010. The generic ballot does quite well in predicting statewide vote in the Granite State and should be considered as a standard predictor in other states.

Methodological Briefs: Data Collection & Post Survey Processing.

Web Survey Live Validations - What Are They Doing?

Scott D Crawford, Survey Sciences Group, LLC (scott@surveysciences.com); Colleen McClain, Survey Sciences Group, LLC (cmccclain@surveysciences.com)

Since Peytchev and Crawford described a typology for web-based survey live-validations in the Social Science Computer Review in 2005, very little has been published documenting specific cases of validations being used, and whether they are serving their intended purpose. Arguably, validations should either improve data quality or reduce respondent burden within a survey (or both), however, the potential is there for a backfire, with validations increasing respondent burden and/or contributing to error.

In this paper, we will examine several web-based surveys conducted over the past several years on social science related topics among various populations, and using paradata collected during the data collection, we will identify what we have found happens when validations are used. We will look at the rate that they are triggered, when they are triggered (early vs. late in a questionnaire), what we know about how the respondent reacted to them, and where possible we will examine if their use had an impact on the survey response. We will also hypothesize as to whether the impact was positive or negative relating to survey quality. We will focus our attention on potential nonresponse as well as measurement errors.

While this paper will not report on any experiments designed to explore web based survey live validation design, it is intended to help identify and prioritize issues for further exploration (possibly in a controlled experiment) during future data collections.

Investigating the General Telephone Interviewer Training Procedure.

Wojciech Jablonski, University of Lodz, Institute of Sociology (wjablonski@uni.lodz.pl)

The aim of this presentation is to outline the results of the methodological study that was carried out among CATI interviewers from October 2009 to August 2010. 12 major Polish research organizations as well as 2 companies in Norway and Iceland participated in the research (Norwegian and Icelandic modules of the project were co-financed by technical assistance funds of the EEA Financial Mechanism and the Norwegian Financial Mechanism within the framework of the Scholarship and Training Fund). The research was based on a standardized self-completion questionnaire, including questions related to various aspects of telephone interviewers' job. Among others, it measured interviewers' opinion on the general training they went through before starting work in a CATI studio. In total, 942 CATI interviewers were surveyed.

The results of two questions will be presented. In the first one the interviewers answered to what extent different elements of the general training (i.e. information about the research company, information about various methodological issues – sampling procedures, types of questions, interviewing techniques – watching and analyzing demonstration interviews, conducting and analyzing training interviews, etc.) are useful for new interviewers. The scale of 1 (redundant element) to 5 (necessary element) was used. The second question was an open-ended one. The interviewers were asked to point out possible changes in training procedure that could be implemented so that the trainees are better prepared for doing their job. Different interviewers' proposals will be presented. They refer to the form of the general training as well as to its content.

Presented results will take into account background variables such as interviewers' age, field of education and work experience.

Obtaining Survey Responses from Two Members of a Health Care Team: An Example of Long-Term Care Organizations.

Melissa Clark, Brown University, Program in Public Health (melissa_clark@brown.edu); Michelle Rogers, Brown University, Program in Public Health (Michelle_Rogers@brown.edu); Vincent Mor, Brown University, Program in Public Health (Vincent_Mor@brown.edu)

To understand organizational functioning, information is often needed from multiple members of the leadership team since no one person may have sufficient knowledge of all aspects of the organization. However, requiring more than one respondent from a single organization increases the complexity of the survey administration process. We administered questionnaires to the Directors of Nursing (DoN) and Administrators (ADMIN) of long-term care (LTC) facilities. A total of 4160 LTC facilities across the U.S. were randomly selected; both the DoN and ADMIN were eligible in 3697 facilities. The DoN and ADMIN were independently mailed a questionnaire. The cover letter included a link and username and password to respond on-line if preferred. Telephone contact attempts were made to non-responders. Questionnaire completion averaged 30-40 minutes and participants were compensated \$35. Overall, 56.5% of eligible respondents participated. Both respondents completed the questionnaire in 41.2% of facilities.

Participation by both individuals was associated with facilities that: were free-standing non-profit; had < 120 beds; had < 10% non-white residents; and were in states with fewer LTC facilities. Among the facilities in which both respondents responded, it took an average of 63.7 days and 8.3 contact attempts for both individuals to participate. We used survival analyses to determine factors associated with number of days and number of attempts for both individuals to participate. No variables were significantly associated with number of days required to receive a response from the second member of the team, given that one individual had participated. Fewer facilities per state, fewer beds per facility, and a lower percentage of non-white residents were associated with the number of attempts needed to obtain the second completion. Creative incentives may be required to obtain adequate response rates among

providers caring for more disadvantaged patients and working in large, for-profit organizations in competitive health care markets.

Interviewer Characteristics Related to Field Interviewer Falsification.

Angel B Griffin, RTI International (agriffin@rti.org); James Pratt, RTI International (jimpratt@rti.org); Sean Shuangquan Liu, RTI International (lius@rti.org); Lee Ellen Coffey, RTI International (coffey@rti.org); Stephanie A. Terrey, RTI International (sterrey@rti.org); Christine Clark, RTI International (chclark@rti.org)

Falsification of data is detrimental to the integrity of any research. When a field interviewer falsifies data, costs increase when investigating possible falsification, reworking fraudulent cases, and hiring replacement staff. Given the negative impact of falsification on research integrity and budgets, can we identify any interviewer characteristics that are directly associated with falsification? Does interviewer age, experience level, or other typically measured performance ratings such as response rates or cost per interview correlate with falsification? Are there differences between interviewers whose work is questionable but not falsified? Using the National Survey on Drug Use and Health (NSDUH) as an example, these questions will be addressed.

The National Survey on Drug Use and Health (NSDUH) is an annual face-to-face, household survey. NSDUH is sponsored by the Substance Abuse and Mental Health Services Administration. First conducted in 1971, this study provides national, state and substate data on substance use and mental health in the civilian, noninstitutionalized population age 12 and older. Approximately 140,000 household screenings and 67,500 NSDUH interviews are completed annually and data are collected in all 50 states plus the District of Columbia by approximately 700 interviewers.

NSDUH implements a complex process to verify the quality and accuracy of each interviewer's work. Sometimes an interviewer's work is verified in person with respondents by another interviewer to determine whether or not the initial interviewer made proper contact. Using 2005 to 2010 non-identifiable archived data, this presentation will examine the correlation of detected falsification with factors such as interviewer experience on the NSDUH (considering both years of service and workload), age, and response rate history. Trends in the number of interviewers found falsifying and their characteristics will also be investigated, as well as any differences between interviewers with questionable work requiring field verification, interviewers who falsified, and those who did not falsify.

A Comparison Between Expert Coders and Computerized Coding Error Detection.

Adam Sage, RTI International (asage@rti.org); Brian Burke, RTI International (bjb@rti.org); Chris Ellis, RTI International (ellis@rti.org); Joe Eyerman, RTI International (eyerman@rti.org); Robert Hughes, RTI International (rhughes@rti.org); Matthew Strobl, RTI International (mstrobl@rti.org)

This methodological brief evaluates the accuracy and benefits of utilizing a newly-developed computer program for analyzing coding error. Using data from the Study of Community Family Life (SCFL), we compare our findings to similar research conducted by Mitchell, et al (2008) on SCFL data. More specifically, we examine "other specify" options and incorrect coding of existing response categories.

The SCFL was part of the Evaluation of Community Healthy Marriage Initiatives, sponsored by the Administration for Children and Families and conducted by RTI International and the Urban Institute. Approximately 4,000 respondents were interviewed in 6 large U.S cities using computer assisted personal interviewing (CAPI).

For this analysis, we have designed a computer program that automatically compares and analyzes text. The first function of our program utilizes datasets and response category information from the questionnaire to determine the appropriateness of the use of the “other specify” option. After creating a list of unacceptable “other specify” terms, the program analyzes the dataset and flags cases that specify an option that already exists within the existing response categories. The second function of our program utilizes transcribed Computer Audio Recorded Interview (CARI) files to determine whether answers were coded correctly. This study examines 2 questions for approximately 2,000 respondents.

The data from our analysis is then compared with data from an earlier study conducted by Mitchell et al. (2008). Their study examined similar research questions, but did so using expert coders. The purpose of this comparison is to determine the accuracy of our computer program compared with expert coders. We then discuss the benefits and practical uses of our computer program for data collection and data quality efforts.

Incorporating Interviewer Characteristics into Item Nonresponse Imputation.

Rebekah Young, The Pennsylvania State University (rly116@psu.edu); Tim Sanders, The Pennsylvania State University Survey Research Center (tas41@psu.edu); Rob Schofield, The Pennsylvania State University Survey Research Center (rschofield@psu.edu); David R. Johnson, The Pennsylvania State University (drj10@psu.edu)

Interviewer characteristics are known to affect item non-response. Incorporating knowledge of these characteristics into post-hoc adjustments for missing data has the potential to help reduce nonresponse bias. In this paper we consider whether, and to what extent, using interviewer characteristics as auxiliary information can improve the fit and utility of multiple imputation models. We use data from the Consumer Health Survey, collected through computer-assisted telephone interviewing (CATI) by interviewers at the Pennsylvania State University Survey Research Center (SRC) between 2007 and 2008 with 5,577 respondents. The interviewer variables to be used include: demographic characteristics, employment history and wages, experience with the survey instrument, and pace of the interview. Many of these factors are known to elicit differential levels of nonresponse, particularly for sensitive questions and for special populations. For telephone surveys, where a small number of interviewers may conduct a large number of interviews, the overall effect from a single interviewer on item nonresponse may be substantial. We examine the potential of interviewer paradata to improve the ability to predict refusals and “Don’t Know” responses to different types of health-related questions. The questions included health topics such as diabetes, blood pressure, heart disease, asthma, and depression, as well as demographic characteristics. We review strategies for incorporating interviewer characteristics into missing data imputation models and assess the usefulness of this approach. Our findings have implications for interviewer training prior to and during studies, as well as suggesting interviewer data that could be collected in future telephone surveys for the purpose of improving models of nonresponse.

Public Perception of Illegal Immigration Policy: Local, State, Regional & National Perspectives.

Political and Social Trends in Arizona, 1992-2010.

Michael J. O’Neil, O’Neil Associates (oneil@oneilresearch.com)

The recent controversy surrounding SB1070 has resulted in a number of perceptions of Arizona throughout the country. Some of these have a basis in fact, many are erroneous. Many of these are reminiscent of an older controversy concerning the state's attempts to adopt a holiday for Martin Luther

King. My paper and presentation will review these controversies as well as political trends in the state for the last 25 years.

Research in Immigration Opinion in Arizona, New Mexico, Texas, and California.

Fred Solop, Northern Arizona University (fred.solop@nau.edu); Nancy Wonders, Northern Arizona University (Nancy.Wonders@nau.edu)

This presentation explores public opinion surrounding the issue of immigration within four states that share a border with Mexico. While these states have similar experiences with undocumented workers coming across the border and impacting local economies, the public opinion in these states is anything but similar.

Analysis of National Data on Immigration Issues from Recent Surveys by Pew.

Mark Lopez (mlopez@pewhispanic.org)

Research in Prince William County, Virginia.

Thomas Guterbock, Univ. of Virginia (tmg1p@virginia.edu)

Research in Prince William County, Virginia, an outer suburb of Washington that adopted a restrictive police enforcement policy in 2008

Survey Data Collection with Smartphones & iPads.

Can Your Smartphone Do This?: A New Methodology For Advancing Digital Ethnography.

Justin T. Bailey, The Nielsen Company (justin.bailey@nielsen.com); E. Nicole Bensky, The Nielsen Company (nicole.bensky@nielsen.com); Michael W. Link, The Nielsen Company (michael.link@nielsen.com); Karen Benezra, The Nielsen Company (karen.benezra@nielsen.com); Hala Makowska, The Nielsen Company (hala.makowska@nielsen.com)

A recent report revealed that the number of smartphone users in the United States has grown 82% in the last 2 years (Intel, 2010). This meteoric rise in the popularity of mobile smartphones also gives rise to a new tool for measuring people in their everyday lives. Smartphones provide portable access to “on-the-go” behaviors in previously unexplored ways. The ability to subtly intrude on the daily lives of people provides access to rich data in a relatively unobtrusive way.

Building on previous work at Nielsen (Lai et al., 2009), we used mobile smartphones to survey 428 South Africans over a 5-week period during the 2010 soccer World Cup. The objective of the research was to assess the impact of the world’s largest sporting event on residents of the host country. Respondents received five questionnaires throughout each day. Questions were directed towards both in-the-moment and in-the-near-past behaviors, as well as general questions about mood and engagement. As a result, panelists provided Nielsen with robust quantitative and qualitative data in which the event’s influence could be measured.

A further advantage of this research tool is the ability to tell a story through digital pictures taken on the smartphone. At the end of each survey, panelists were asked to take a picture of their current activity and given the option of captioning the photo. Using innovative software, we used the images to analyze behavior over the course of the World Cup. In addition, using metatags, we assessed general changes in

comfort levels using the smartphone across time, such as how personalized the pictures became as the participant adjusted to the survey tool.

Lastly, we summarize implications for this research tool and possible future areas of research that could be explored.

“Can you see it now? Good”: Usability Testing of a Mobile Health Application.

Sarah L Cook, RTI International (scook@rti.org); Rita Sembajwe, RTI International (rsembajwe@rti.org); Emily Geisen, RTI International (egeisen@rti.org); Barbara Massoudi, RTI International (bmassoudi@rti.org)

Due to advances in technology and the prevalence of smartphones, researchers are now conducting surveys on handheld mobile devices. However, using handheld devices to complete self-administered surveys has not been consistently evaluated. In particular, few methods of usability testing for these devices have been examined and many of those are not generalizable. In this study, we conduct usability tests on handheld mobile devices to evaluate how well respondents are able to use the target application to input survey information. The purpose of this presentation is to explore the methods used, results of usability testing, and challenges encountered so that survey researchers can improve the quality of conducting surveys on mobile devices.

In this study, we conducted usability testing on an Android smartphone application. The application was designed to keep daily records of asthma and mental health symptoms for patients with those conditions. Prior to launching the study, this application was tested by patients with these conditions to not only assess the usefulness of the application, but also the ease of entering data via mobile phone.

In this presentation, we will discuss the methods used for conducting the usability tests and how well it worked for evaluating the application. We will highlight areas in need of future research in order to improve the success of mobile application usability testing. We will also discuss results specific to usability and our recommendations for future designs of survey questions on handheld mobile devices.

The Smart(Phone) Way to Collect Survey Data.

Carey Stapleton, Service Management Group (cstapleton@smg.com)

As more individuals migrate to smartphones, it is natural that survey researchers will use this new mode of surveying. However, as with any new mode of survey research, the biases and issues introduced by a survey mode must be understood. Smartphones introduce unique challenges to the survey respondent including smaller screen sizes, slower page loading, and a lack of graphics. Each of these issues has the potential to negatively impact data quality and as a result survey researchers need to understand how to mitigate them.

Using customer satisfaction survey data, a test was developed to understand how to minimize bias in order to create the most effective survey for smartphones. Three key questions were answered during this test:

1. What input types work best for survey questions?
2. Should questions be consolidated onto fewer pages?
3. What is the optimal number of survey questions that respondents should be asked?

Four test groups were created to better understand the optimal survey design for smartphones. Each test group has a sample size of over 1,000 responses. The four test groups were compared to each other and a control group of desktop/laptop computer respondents taking the same survey. Some of the key

comparison points include breakoff rates, score differences, inter-item correlations, extreme response styles, and demographic profiles.

Based on results of this test, seven best practices in smartphone survey design were developed and will be discussed.

The Next CAPI Evolution - Completing Web Surveys on Cell-Enabled iPads.

James Dayton, ICF International (jdayton@icfi.com); Heather Driscoll, ICF International (hdriscoll@icfi.com)

Intercept field data collectors working in outdoor environments using electronic devices face a number of challenges traditional in-person data collectors do not, such as conducting interviews in bright sunlight, wind, and/or rain. These interviewers need data collection devices that are lightweight, can be held in one hand or suspended from the neck, are protected from the elements, have outstanding battery lives, are easy to read in various lighting levels, and have large screens able to display an easy-to-read font and graphics.

An issue all field CAPI (Computer Assisted In Person Interviewing) data collectors currently face involves software licensing costs and data security. These stand-alone data collection devices must be synched-up periodically (completed interviews must be transmitted to a central server and any survey/sample updates transmitted back to the laptop or tablet data collection device). In addition, the survey data is at risk if the data collection device is lost, stolen or damaged between 'synchs'. Finally, CAPI software packages generally require expensive survey software development licenses as well as licenses for each deployed data collection device. These costs can be non-trivial, especially if the effort is significant.

Researchers at ICF sought a data collection option that was both innovative and well-suited for conducting interviews in challenging environments without many of the cost and data security implications that can accompany more traditional CAPI implementations. Our solution was web-based surveys on cell-enabled tablets (specifically 3G iPads).

We will measure cost and QA improvements over paper and pencil intercepts, as well as test other possible benefits of the devices, such as using GPS locations combined with date/time stamps to assist with interviewer QA. The results of this pilot study will inform survey researchers on the potential barriers to infield web-based data collection in adverse environments, including network speeds, connectivity issues, environmental challenges and cost.

This Time For Africa: Using Smartphones To Measure World Cup Engagement In South Africa.

E Nicole Bensky, The Nielsen Company (nicole.bensky@nielsen.com); Justin Bailey, The Nielsen Company (justin.bailey@nielsen.com); Michael Link, The Nielsen Company (michael.link@nielsen.com); Karen Benezra, The Nielsen Company (karen.benezra@nielsen.com); Hala Makowska, The Nielsen Company (hala.makowska@nielsen.com)

Given the excitement and buildup of the world's largest sporting event, South Africans had much to pay attention to for the 2010 soccer World Cup held in South Africa. For researchers, even greater was the opportunity to explore different ways to measure the impact of the influx of different kinds of advertising, viewership and sudden attention focused on local South Africans. During the 2010 World Cup Nielsen conducted a study to measure involvement, media consumption and brand purchasing / awareness of local South Africans. Building on previous digital ethnography research (Lai et al., 2009), respondents were equipped with a Blackberry smartphone that included an app with a brief survey. The survey was triggered five times a day for the full 35 days of the World Cup. This survey was administered to 428

South African residents spread across 4 different cities throughout the country. To utilize the mobile nature of this study, respondents were asked questions about what they were currently doing and what they had been doing during the previous few hours. Lastly, they were asked to take a picture of what they were currently focused on. By doing so we were better able to capture in the moment reactions to the respondents' surroundings. The Blackberry smartphone was leveraged and provided as an incentive for the respondents' post-survey completion. Although there were initial concerns about the burden placed on panelists during the 5 weeks, we observed a high level of response (~80% for the overall study). This paper will examine the impact of conducting a long range mobile digital ethnography study on response rates, data quality, and the various milestones within the World Cup. As well, this paper will look at how some of these results can be translated into a similar study done within the US.

Web Survey Methods – Sampling, Participation, Data Quality.

Differential Sampling Based on Historical Individual-Level Data in Online Panels.

Richard H Kelly, Opinionology (rkelly@opinionology.com)

Non-response bias can significantly compromise the representativeness of a sample. In order to combat these problems, researchers tend to mandate marginal quotas to known population proportions. However, the problems associated with quota sampling have been extensively documented when marginal quotas do not fill up at the same time. Differential sampling, whether through propensity scores or model-based response rate adjustments, avoids the problems involved in quota sampling. Differential sampling oversamples subpopulations with lower response rates to ensure survey respondents are balanced according to population proportions. Adjusting for these subpopulations becomes even more important in the online space where individual response rates vary dramatically.

Practitioners who use online panels must develop solutions for coverage limitations, conditioning and respondent attrition, source variability, and non-response bias. We discuss how to more effectively correct non-response bias in online sample to create more accurate estimators from data collected online.

The most common way to adjust for non-response bias involves modeling the response rate for each of the subpopulations and oversampling accordingly. Some real time sources use aggregate level demographic estimates to adjust the flow of respondents into their surveys. However, demographics do not account for a large percentage of the variation in response rates. Panel companies can instead use historical response rate data at the individual level to accurately adjust the flow of respondents and correct for non-response. We discuss a system of implementing individual-level response rate estimation and show the advantages when compared to model-based response rate estimates.

Are you who you say you are? Using a Multisource Cross-validation Methodology for Panel Membership Information.

Kumar Rao, The Nielsen Company (kumarrons@yahoo.com); Tim Dolson, The Nielsen Company (kumarrons@yahoo.com); Mark Kinnucan, The Nielsen Company (kumarrons@yahoo.com)

Over the past few years, market researchers working with data from nonprobability online panels have voiced a number of data quality concerns over issues such as respondent identity, increased satisficing, and possible professionalization. While recognizing the importance of having real, unique, and engaged panelists, panel companies are responding to these issues by introducing a variety of remedial measures such as name-address verification, email address verification, and validation of key demographic information against third-party databases. This study describes one such effort. In this study, for a cross-

section of new members who were recruited online, we validate their identity and self-reported information against multiple data aggregators and identity verification vendors to identify patterns of matches and mismatches. In addition to conducting a cross-vendor comparison of match rates and assessing differences between matched and mis-matched sample portions, we conduct multivariate analysis of match rate at various levels (i.e., vendor, sample groups, etc.). Results reveal a high identity match rate across all vendors with interesting relationships (ex. with panel activity and member age). While match rates for self-reported information were modest in general, the effect was more pronounced for certain demographic groups. This study is likely to have important implications for the data quality aspects of online panel recruitment, participation and retention.

Measuring and Stimulating Respondent Attentiveness in Web Surveys.

Adam Berinsky, Massachusetts Institute of Technology (berinsky@mit.edu); Samantha Luks, YouGov (sam.luks@yougov.com); Doug Rivers, Stanford University and YouGov (doug@yougov.com)

Survey researchers have long known that respondents often fail to pay close attention to surveys. Instead of expending the cognitive effort to interpret the question, mentally search for information, integrate information into a single judgment, and translate the judgment into a response option (Tourangeau & Rasinski, 1998), people take mental short-cuts to minimize the amount of effort to answer a question (Krosnick, 1991, 1999). This tendency is an especially significant concern for internet-based surveys because there is no interviewer present to monitor the quality of answers to survey questions.

Recently, psychologists (e.g., Oppenheimer et al, 2009) have advanced the use of screener questions - questions that are formatted as typical survey questions, but require closely reading instructions to answer correctly - as a way to monitor respondent effort. For example a screener question may at first appear to be asking a respondent about his/her favorite sports teams, but then instruct the respondent to ignore all of the response options and check "none of the above." These questions can serve as a general measure of attentiveness to the survey questions and response instruction.

In our paper, we develop and employ standard screener questions as a method to both gauge and stimulate respondent attentiveness. We then evaluate both the incidence of attentiveness and its effects on the measurement of social and political attitudes.

Display Matters: A Test of Visual Display Options in a Web-Based Survey.

Jennifer Romano Bergstrom, U.S. Census Bureau (jennifer.romano@census.gov); Jennifer M Chen, U.S. Census Bureau (jennifer.chn@gmail.com); Timothy R Gilbert, U.S. Census Bureau (timothy.r.gilbert@census.gov); Matt Jans, U.S. Census Bureau (matthew.e.jans@census.gov)

Surveys are increasingly being conducted online, and it is pertinent to establish clear guidelines for presenting self-administered survey items on a computer screen. Goals should include reducing respondent burden and measurement error. Web survey designers often need to decide how to best present long lists of information and where to place navigation buttons (i.e., next and previous). We evaluated the usability of web-based National Survey of College Graduates (NSCG) questionnaire prototypes. In the first round of testing (n=8), respondents had difficulties proceeding through the survey because the 'Next' button was on the left side of the screen and the 'Previous' button was on the right. In the second round of testing (n=30), four versions of the survey were tested to assess usability of the 'Next' and 'Previous' buttons based on their placement on the screen (right or left side) and additionally, to assess the display format of a long list of occupation options (one-column versus two-columns). Dependent measures included participants' comments from a think aloud protocol, self-reported ratings of

satisfaction with the survey, responses to qualitative debriefing questions, time required to complete the job code item, and eye-tracking data focusing on which button (i.e., 'Next' or 'Previous') and which half of the job code list participants looked at first and more often. Qualitative and quantitative results revealed that participants preferred and performed quicker when the long list of occupation options was displayed in two columns rather than one and when 'Next' was displayed to the right of the 'Previous' button rather than vice versa. Our findings support usability best practice guidelines that recommend eliminating excessive scrolling on Web sites, and following reading conventions (i.e., looking to the right to move forward, as if turning a page in a book).

Saturday, May 14, 2011

8:00 a.m. - 9:30 a.m.

Concurrent Session A

**Addressing the Challenges of Interviewing Youth about Sexual Assault:
Experiences from the National Survey of Youth in Custody.**

Outlier analysis of Youth Reports of Sexual Assault.

David Cantor, Westat (davidcantor@westat.com); John Hartge, Westat (johnhartge@westat.com)

The National Survey of Youth in Custody interviewed juveniles in residential placement about sexual assault while in the facility. This anonymous survey used audio computer-assisted self-interviews (ACASI). Survey measurement error can lead to under- or over-reporting of events. Because of the absence of external validation criteria, it was necessary to use internal measures to assess data quality. For purposes of this evaluation, data quality was assessed using three measures: (1) extreme, inconsistent, or illogical responses (outliers), (2) strength of expected relationships between facility/youth characteristics and reports of assault (construct validity), and (3) interview performance statistics (metadata). The outliers indicate the extent youth reports provide details about the event that are plausible. The measures of construct validity assess the extent youth profiles are theoretically logical. The metadata examine characteristics of the interview and compare response frequencies across multiple respondents, such as the duration of the interview and the frequency of usable responses. This paper describes analyses that assess data quality of youth reports of sexual assault using these three measures by examining the relationship between characteristics of youth and data quality, as well as the relationship between the different measures of quality.

Overall, the results of the analysis found strong evidence of construct validity, as indicated by highly significant relationships between predictors and the reporting of sexual assault. The outlier analysis provided guidance on which interviews were suspect and allowed for examination of results when the outliers were treated in different ways.

Addressing Human Subject Issues for a Survey on Sexual Assault Among Youth in Residential Placement.

Tim Smith, Westat (Timsmith@westat.com)

The National Survey of Youth in Custody collects self-reports of sexual assault from incarcerated youth. This presentation will describe elements of the study design used to meet local and federal requirements for the protection of research participants.

Three issues will be discussed:

- Informing youth about the nature of the survey and their rights as participants. Ethical standards require that information needed to make an informed decision be conveyed to potential research participants. Communicating this information to young people was challenging.
- Mandatory reporting of harm. State law requires that individuals who know or suspect that a youth has been assaulted or is at risk of assault report to local or state authorities. These requirements conflicted with need to maintain the confidentiality of survey responses.

- Emotional harm resulting from survey participation. The survey asked questions necessary to classify events according to definitions conveyed in the Prison Rape Elimination Act. Youth who thought about potentially traumatic events could experience emotional harm.

An important element in the design process was to collect data that could be used to evaluate the design elements affecting human subject concerns. The presentation will provide examples of the type of data collected to achieve this purpose.

Assessment of the effects of Active Consent for a Survey of Youth in Custody.

John Hartge, Westat (johnhartge@westat.com); David Cantor, Westat (davidcantor@westat.com)

One of the major obstacles related to interviewing juvenile offenders is obtaining consent from a parent or guardian. If active parent/guardian consent is required, the response rate is significantly lower than in-loco parentis consent is accepted. Recent research for youth in the general population has found that requiring active parent/guardian consent results in youth with higher risk profiles to be excluded from the survey (e.g., higher obesity rates; higher drug use; more sexual behavior). The purpose of this paper is to examine the effects of requiring active parent/guardian consent for youth in residential facilities. The National Survey of Youth in Custody (NSYC) surveyed a national sample of youth on their experiences related to sexual assault. In approximately two-thirds of the jurisdictions, active parent/guardian consent was required. In the other third, state and facility administrators acted in-loco parentis. The response rate in the in-loco parentis jurisdictions was twice as high as the active parent/guardian consent locations. The rates of reporting sexual assault are compared across these two types of jurisdictions, once controlling for other differences between facilities. This is done through a series of logistic regressions which include the consent type and facility characteristics. The results are consistent with prior research for general population samples --- requiring active parent/guardian consent tends to exclude youth who are more likely to report sexual assault. However, the magnitude of this effect is not as large as found in other studies and it is not as large as one might expect given the large differences in the response rate. The analysis also tests whether the consent type affects relationships when predicting rates of sexual assault. The results of this analysis finds a number strong predictors of sexual assault, but very few seem to vary by consent type.

Procedures to Obtain Active Parental Consent for Youth in Custody.

Andrea Sedlak, Westat (andreasedlak@westat.com); Tim Smith, Westat (timsmith@westat.com)

The National Survey of Youth in Custody (NSYC) interviewed youth in national sample of state-owned/operated residential facilities. For two-thirds of the jurisdictions, active parent/guardian consent was required. The purpose of this paper is to describe the procedures used to obtain consent and to provide the success that these procedures had with respect to obtaining consent. The survey implemented a number of different procedures, depending on the requirements of each particular jurisdiction. These procedures ranged from obtaining written permission through the mail to making telephone calls to parents/guardians. The procedures were split between the project team and the facilities. The paper will describe the extent to which these procedures were successful in getting consent and the major reasons why consent could not be obtained.

Success rates are compared by several different dimensions, including those: 1) for procedures carried out by facility vs. those carried out by the project team, 2) those carried out by telephone vs. mail and 3) those accepting verbal consent vs. prompting over the telephone. Across all of these methods, the primary reason for not obtaining consent was an inability to obtain good contact information from the facility for the parent/guardian. When actually contacted, parents were generally willing to allow their

youth to participate on the survey. Significantly improving consent rates, therefore, depends on developing methods that take advantage of opportunities to approach parents/guardians.

Cell Phone Sampling: Operational Issues.

Integration of Cell Phones at Statistics Canada.

Marco Grenier, Statistics Canada (marco.grenier@statcan.gc.ca)

As observed in the United States and other countries, the number of Canadian households that abandon the traditional telephone land lines for cell phones is constantly growing. This results in declining coverage of the current Random Digit Dialing frame and other methods using the land lines. In addition, it has a negative impact on contact and response rates of telephone surveys. For these reasons, Statistics Canada has started to look into integrating cell phones in the survey development and collection process.

The first step to integrate cell phones in Statistics Canada surveys was to standardize the practices within the organization. New guidelines were developed in order to have the same rules across all surveys. Currently, it is possible to contact respondents over their cell phones under some specific circumstances. However, although permitted, no surveys have started to use cell phone number as sampling unit or as primary contact information.

In 2010, Statistics Canada acquired a file of cell phone numbers from a major mobile telephone provider in order to help incorporate cell phone numbers in surveys. An evaluation of the file was conducted to assess the potential of using it for different aspects of surveys. The evaluation was conducted by many teams for different purposes: telephone first contact, frame building, source of new addresses and pre-tracing.

The presentation will describe the evaluation of the file and will present the recommendations that were made. The planning of the steps for integration of cell phone numbers will also be discussed. This includes the development of a new address based frame as well as a pilot study in the field to study the receptiveness of Canadians to be contacted over their cell phone. This study will have a major impact on the strategy to adopt the use of cell phone numbers in Statistics Canada.

Pre-Call Validation of Random Digit Dialing Cell Phone Numbers: A Field Experiment.

Tanja Kunz, Darmstadt University of Technology (kunz@ifs.tu-darmstadt.de); Marek Fuchs, Darmstadt University of Technology (fuchs@ifs.tu-darmstadt.de)

As the percentage of individuals using solely or mostly cell phones continues to grow, cell phone surveys are increasingly used in survey research. Because in most western countries there is no register where cell phone numbers are listed, the usage of a random digit sample is a common strategy. However, the proportion of invalid numbers is typically higher than in the landline frame. Thus, dialing invalid numbers is extremely time-consuming, resulting in an inefficient data collection and extended fieldwork period.

The present experiment is designed to test mechanisms which can be used to validate random digit cell phone numbers prior to field work: (1) Cell phone number validation services provide verification in real time by performing a Home Location Register (HLR) lookup. (2) Text messaging services can also be used for cell phone number validation by using a text message gateway and sending bulk text messages from the internet to cell phones.

The experiment reported in this paper is designed to help balancing survey costs, interviewer workload and fieldwork efficiency by means of an effective strategy of pre-call validation of randomly generated cell phone numbers. The field experiment applies a between-subjects design (n=25.000) in order to determine the most cost efficient and reliable method. The analysis will focus on two main questions: (1) Is number validation or text message validation a reliable method? (2) Does reduction of the interviewers' workload in consequence of the elimination of invalid numbers justify the additional costs for number validation or text messaging?

Initial findings indicate the efficiency of both number validation and text messaging in improving contact and interview success rate. Furthermore, cell phone number validation seems to be more reliable in identifying invalid numbers than text message delivery reports.

Validity of Questions to Identify Cell-only Households.

Sherman Edwards, Westat (ShermEdwards@westat.com); Michael Brick, Westat (MikeBrick@westat.com); Royce Park, UCLA Center for Health Policy Research (npark@ucla.edu); David Grant, UCLA Center for Health Policy Research (dgrant@ucla.edu)

Many RDD surveys now include samples of numbers assigned to cellular service. It is often useful to identify cell-only households in the course of the interview, whether for screening, for use in weighting or as a classification variable for analysis. There has been little research on the reliability and validity of questions used to classify households by telephone service and use. This paper will compare responses to questions about telephone use in a household screener with a somewhat different set in a subsequent extended interview. These questions were part of the 2009 California Health Interview Survey (CHIS). CHIS is a biennial survey of California households, with one adult selected at random from each household screened. One child and one adolescent associated with that adult may also be sampled. The adult interview is approximately 35 minutes long. CHIS 2009 included both landline and cell RDD samples. The screening interview for cell numbers included a question on the presence of a landline in the household ("Does your household have a regular telephone at home?") and another question on the relative use of landline and cell by "you and the people that live with you." Near the end of the adult extended interview, similar questions were asked about the presence of a landline ("Is there a regular or landline telephone in your household?") and about relative use. About 20 percent of those classified as cell-only from the screener reported having a landline in the extended interview.

The paper will explore the differences between responses to the screener and extended interview, including comparing statewide estimates of cell-only households from each source with national estimates from the NHIS. It will also characterize those who change their responses between the screener and extended interview in an attempt to understand the reasons for the change.

The Next Generation: Using Cell Phones to Survey Households with Children.

Kristie M. Hannah, ICF Macro International (khannah@icfi.com); Randal ZuWallack, ICF Macro International (rzuwallack@icfi.com); William Robb, ICF Macro International (wrobb@icfi.com)

Random digit dialing (RDD) surveys are becoming less representative of the general population as more and more households discontinue the use of their landlines in favor of cell phones. Several recent studies have detailed the implications of this trend on surveys of adults. Our research focuses on the cell-only implications for surveys of children. Many young adults are cell-only and many of them have young children. Currently, 22 percent of all adults are cell only but 31 percent of adults living with children under the age of five are cell only. In a recent survey, we found that a higher percentage of adults surveyed via cell phone reported having children under 18 than those surveyed on a landline. Using data from the

National Health Interview Survey (NHIS), we examine characteristics and trends of children in cell only households. We also discuss implications of surveying children via cell phone, such as random selection of a child within a household, and corresponding weighting factors.

The Telephone Point of Purchase Survey Cell Phone Hit Rate Test.

Aniekan Okon, Bureau of the Census (aniekan.a.okon@census.gov); James Arthur, Bureau of the Census (james.arthur@census.gov)

The Telephone Point of Purchase Survey (TPOPS) is a Random Digit Dialing Computer-Assisted Telephone Interview (CATI) that collects information on what people buy and where they buy it from selected households. The information collected in the TPOPS, along with information from the Consumer Expenditure Surveys, which are also conducted by the Census Bureau, is used by the Bureau of Labor Statistics (BLS) to update the Consumer Price Index (CPI).

Currently, the (TPOPS) has coverage and representative demographic issues because cell phone numbers are excluded from its sample frame. Not including households that only have cell phones may leave a gap in the outlet and expenditure data, because these households may purchase items and services differently than households with landline phone numbers purchase them. To resolve this problem, BLS wants to include a frame of cell phone numbers in the TPOPS sample. Before including the cell phone frame in the TPOPS, a test will be conducted to determine the number of cell phone numbers that need to be called to result in a productive interview. The test will also be used to evaluate how often cell phone numbers move out of the geographic area for which they were selected.

The authors will discuss the main goals of the test, but also will look at the response rates and completion rates for both the cell phone frame and the landline frame and offer comparisons

Cultural Competence in Survey Research – What is it?

Increasing Cultural Sensitivity as a Means of Improving Cross-Cultural Surveys: Methods Utilized in the California Health Interview Survey (CHIS) 2001-2011.

Elaine G Zahnd, Public Health Institute (ezahnd@phi.org); Sue Holtby, Public Health Institute (sholtby@cruzio.com); David Grant, UCLA Center for Health Policy Research (dgrant@ucla.edu)

Designed as a broad public health surveillance system, the California Health Interview Survey (CHIS) is the largest population-based state health survey in the United States (45,000-55,000). RDD telephone interviews are conducted biennially with a randomly selected adult, adolescent, and child in households scientifically sampled from every California county.

A major CHIS goal has been to provide health information on the state's diverse racial/ethnic groups. California has one of the most diverse populations in the nation, and a significant percentage of CHIS respondents have limited educational backgrounds, are immigrants with lower acculturation levels, or speak a language other than English as their primary language. Procedures to increase cultural sensitivity and cross-cultural transparency have been employed to improve survey question functioning and to achieve high participation. Techniques include: 1) Multicultural Technical Advisory Committee; 2) English language simplification; 3) Cultural adaptation; 4) Translation in Spanish, Mandarin Chinese, Cantonese Chinese, Korean and Vietnamese; 5) Asian subgroup oversampling; 6) Cross-Cultural Cognitive Testing and Behavioral Coding; and 7) Cultural training for interviewers.

The Multicultural TAC advises on languages, topics, and measures, such as helping select discrimination measures. The English language simplification review goal is to reduce or eliminate linguistic complexity. Many Californians are bilingual speakers, although English is not their primary language. Bilingual speakers are offered the option of responding in English or in one of the five translated languages. The major challenge for the translators is to produce culturally and linguistically appropriate translations that adhere to the question intent. CHIS cultural review has taken different forms, currently blending into the translation process. English simplification and cultural review not only improve the reliability of responses among those who are interviewed in English, but also increase comprehension among respondents interviewed in one of the other languages. CHIS benefits from the multiple methods utilized to increase cultural sensitivity.

Breaking Barriers for Nonstandard Language Speakers in Survey Research.

Julie Ingels, Mathematica Policy Research (jingels@mathematica-mpr.com); Erin Panzarella, Mathematica Policy Research (epanzarella@mathematica-mpr.com)

According to the 2000 Census, 19.7% of the United States population 5 years and older speak a language other than English at home, and 8.1% of the population speak English less than "very well." Further, 4.7% of the population 5 years and older live in linguistically isolated households, which means that no member of the household 14 years and older speaks only English or speaks a non-English language and speaks English "very well" (United States Census Bureau).

Groups with a small sample presence, such as speakers of non-standard languages (i.e. languages other than English and Spanish) increasingly have an important role to play in American social policy. As a result, survey researchers need to work toward breaking language barriers so that non-standard language speakers can respond to surveys and have a voice in public policy. Breaking these language barriers will reduce bias by facilitating higher response rates among individuals who cannot complete a survey in English or Spanish, and will also increase the generalizability of the results of the survey. Given these facts, we urgently need to identify ways to circumvent language barriers to participation in survey research. Unfortunately, funding tends to be a serious constraint on this effort: Very few studies have the available funds to translate a survey into multiple languages, train interviewers and researchers on cultural sensitivity, and produce contact materials in multiple languages, especially when the number of sample members in each language group tends to be very small.

We will share some practical steps we developed to help increase participation in telephone surveys among non-standard language speakers. In our presentation, we will discuss successes we have had working with a private translating service. We will also identify areas where further work is needed.

Cultural Sensitivity and Questionnaire Design.

Janet A Harkness, University of Nebraska-Lincoln (jharkness2@unl.edu)

Questionnaires in various implementations are the main means of collecting data in surveys; other data collection tools include collecting biomarker specimens and results on performance tests such as grip strength measurements.

The questions asked and the response options offered are the implementation stage of a research design that begins with a definition of research goals in relation to a given concept, followed by identification of relevant constructs to be measured via indicators that have been identified as suitable for this purpose. Survey questions therefore can be seen as the linguistic vehicles for indicators. Biomarker specimens and measurements of performance can be viewed similarly: they provide information (data) relevant for the research goals and analysis.

Many comparative surveys (cross-cultural or cross-national, possibly multipopulation) are based on designs which try to standardize the measurements undertaken by standardizing the questions asked across populations. In order for this "ask the same question" approach to produce comparable data, the different instruments, often in different languages, need to produce results which can be compared. In basic design terms, this means that the questions asked of each population and the indicators they implement must tap specified concepts in the ways envisaged in the research design

This paper discusses challenges for questionnaire design and quality in comparative settings. It focuses first on comparability across different cultures with regard to the selection of conceptual dimensions, indicators, and question wording for source questionnaires. (A source questionnaire is the basis of all other versions in other languages.) The second part deals with measurement issues raised by the need to translate a source questionnaire into different languages. The last section points to lessons to be learned from comparative research for the design and testing of instruments to be used in multicultural societies and in countries with pronounced cultural and linguistic minority populations.

Cultural Awareness and Sensitivity: From Design through Dissemination

Beth-Ellen Pennell and Yuchieh Lin, University of Michigan and Margaret Brill, Kalamazoo College.

Beth-Ellen Pennell, Institute for Social Research, University of Michigan

(bpennell@isr.umich.edu)

Knowledge of local context is critical to understanding the behaviors and attitudes that surveys seek to measure. This presentation will focus on combining quantitative and qualitative methods to gain insight into local context and culture in cross-cultural survey research. Cultural sensitivity in this sense depends not only on the relevance and meaning of the research questions and instrument for given local contexts but also intersects with the approaches and methods used throughout the survey lifecycle from overall design through data collection to data dissemination. Examples include such fundamental issues as how respondents interpret the research setting and their role in the research, to issues of courtesy and courtesy bias, interview privacy, matching (or not matching) interviewers and respondents on key demographic characteristics, among many other considerations. Quantitative methods such as survey research emphasize objectivity and standardization of procedures and strategies. Such an approach may, however, conflict with appropriate and necessary adaptation to local cultural norms. This presentation explores the middle ground between these two extremes. It provides an overview of techniques and practice primarily drawn from the mixed-methods literature. These will include an overview of semi-structured methods, participatory techniques, and methods triangulation to increase appropriate cultural and context-sensitive survey research methods. The strengths and weaknesses of these techniques will be considered for the cross-cultural survey context, followed by suggestions for a set of practical recommendations in order to gain a more informed and balanced view of local context which enables us to effectively implement culturally-sensitive survey methodologies.

Measurement Error and Nonresponse Bias: Current Practices for Addressing Language Minority Issues in Pretesting and Data Collection.

Brad Edwards, Westat (EdwardB1@westat.com)

Cultural sensitivity can have a variety of meanings, but here we examine the sensitivity of survey methods to detect cultural differences. Culture and language are inextricably linked, and in this paper we focus on methods for addressing different languages. Cultural and language minority issues have been a source of societal conflict in North America at least since colonial times. In the U.S., English became so dominant that it now functions in effect as a prerequisite for entrance into mainstream society. However, language minorities persist, and have grown in numbers with each major wave of immigration. Spanish is currently the second most commonly spoken language in the U.S. by a large margin.

Current survey methods that address language minority issues range from total exclusion (e.g, a monolingual questionnaire and non-English speakers are ineligible) to full inclusion (via translated instruments and outreach materials). Full inclusion is rarely attempted except for Spanish-speakers, and full exclusion has become increasingly rare for general population studies. Recent guidelines at the U.S. Census Bureau encourage pretesting in all languages that will be used in the final data collection instruments. Other survey methods for addressing language minority issues in interviewer-administered surveys include the use of interpreters, bilingual interviewers who “translate on the fly” or with a separate script, and proxies who speak English. (See Harkness et al, 2010)

This paper will review current practice in the context of measurement error and nonresponse bias. Recent research has investigated measurement error in interpreted interviews compared to interviews conducted by bilingual interviewers using translated instruments; measurement error in comparative surveys measured by behavior coding of audio recordings; and nonresponse bias that can be attributed to exclusion of language minorities. Findings are presented from new research that uses pretesting to predict measurement error for specific questions in English instruments and Spanish translations.

Improving Response Rates in Health & Medical Surveys.

Are You Still Home? Recontacting Respondents for Follow-up Surveys.

Kathleen S. O'Connor, Centers for Disease Control and Prevention (koconnor1@cdc.gov); Alicia M. Frasier, NORC at the University of Chicago (Frasier-Alicia@norc.org); Shannon TenBroeck, NORC at the University of Chicago (TenBroeck-Shannon@norc.org); Stephen Blumberg, Centers for Disease Control and Prevention (sblumberg@cdc.gov)

Surveys of low prevalence population groups can be difficult and expensive to conduct. One cost saving strategy is to recontact those identified in an initial survey as having a characteristic of interest. The 2009-2010 National Survey of Children with Special Health Care Needs (NS-CSHCN), a large population based telephone survey, provided a unique opportunity to identify approximately 6,000 school-aged children who were reported as ever having an autism spectrum disorder, intellectual disability, and/or other developmental delay diagnosis, for a targeted follow-up interview. A nationally representative sample of this size of children with such rare conditions would be very expensive to identify any other way.

At the end of the NS-CSHCN, respondents were asked to provide names, addresses, and alternate telephone numbers and type (cell, landline, or work) for a follow-up survey about the health and health care of the sampled child. Data collection is ongoing, but to date, almost 89% of respondents gave or confirmed their current mailing address. Respondents were much more reluctant to give the name or initials of their child (64%) than their own name or initials (94%). Approximately 21% provided an alternate telephone number, of which 86% were reportedly a cell number.

We will explore the benefits and drawbacks of collecting recontact information, and whether item response propensity during the initial interview predicts the likelihood of participation rates and nonresponse bias to the follow-up interview. In doing so, we will examine how respondent and child demographic characteristics are related to whether recontact information is provided and whether more complete recontact information results in greater likelihood of participation.

Does Customizing Medicare Surveys with Health Plan Names Justify Its Cost? A Randomized Experiment.

Amelia Haviland, RAND Corporation (haviland@rand.org); Marc N. Elliott, RAND Corporation (elliott@rand.org); Nate Orr, RAND Corporation (orr@rand.org); Carol Edwards, RAND Corporation (carola@rand.org)

The Affordable Care Act and its emphasis on patient-centered care have re-emphasized the importance of understanding the role of health insurance in patient healthcare experiences. There is evidence that many, including Medicare beneficiaries, may not recognize either the name of their health insurance plan or that Medicare Advantage (MA) plans are a (managed care) form of Medicare insurance. What beneficiaries understand or recognize about their Medicare coverage is central to a survey experiment conducted in 2010. The Medicare CAHPS survey is administered to 295,000 MA beneficiaries in 343 MA plans each year, with each survey customized with the beneficiary's MA plan name, resulting in many relatively small print runs.

The purposes of this experiment were to assess whether the use of a less expensive "generic" survey form, as opposed to a survey form customized with MA plan names, would harm beneficiary engagement with the survey – lowering overall response rates or item completion rates, or disrupt trending by changing how beneficiaries answered evaluative items. The experiment randomized 5,000 beneficiaries across MA plans to receive the generic version of the survey. Results show that mail response rates for MA beneficiaries who received a generic survey did not decline; in fact, they were slightly higher (48.5%) than for those receiving a customized survey (47.9%) - failure to recognize plan names might have lowered response rates on some customized surveys. Response rates during telephone follow-up were also similar (11.7% vs. 11.4%) as were rates of item missingness. When we compared mean evaluations of plans and healthcare on survey measures from generic and customized surveys we observed no significant differences. Thus use of "your plan," as opposed to plan-specific customized surveys, may reduce survey costs with no loss to response rates or data quality.

Increasing Response Rates In A Survey Of Physicians And Nurses.

Jeanette Y Ziegenfuss, Mayo Clinic (ziegenfuss.jeanette@mayo.edu); Jon Tilburt, Mayo Clinic (tilburt.jon@mayo.edu); Timothy J Beebe, Mayo Clinic (beebe.timothy@mayo.edu)

Surveys are an important tool for assessing physician and nursing professionals' practice patterns and guideline adherence. The Patient Protection and Affordable Care Act of 2010, emphasizing both patient and provider sentiments and practices, will accelerate the use of surveys as a means to establish estimates for health policy guiding provider practice behavior. In order to accurately meet these needs, it is important to obtain quality survey data, consisting of low item and unit nonresponse. One strategy that has been shown to increase response rates in physician surveys is the use of first class stamps or priority mail (VanGeest et al., 2007). In a survey of a representative sample of 1,600 nurses and physicians in Minnesota, we randomized individuals to one of two envelope conditions: a padded 8.5x11 inch envelope or a similarly sized priority mail envelope. All mailings included a laser pen incentive, a cover letter signed by a team of physician and nurse researchers and a survey booklet. The survey itself assessed healthcare professionals' perspectives on pandemic H1N1 (Swine Flu). After the first mailing, the response rate was 53.9% and did not differ across envelope condition (55.6% in priority mail envelopes compared to 52.2% in padded envelopes, $p=0.18$). Overall, nurses were more likely to respond (59.6% compared to 48.3% of physicians, $p < 0.001$). The envelope used did not impact response tendency for physicians, but did for nurses. 63.5% of nurses randomized to the priority mail envelope responded compared to 55.8% of those randomized to the padded envelope ($p=0.03$). There was no difference in item nonresponse across the two envelope conditions. While priority envelopes were shown to be more effective in a population of nurses, future analysis including an estimation of cost per complete will be undertaken to determine the most effective strategy in a population of physicians.

Decision Factors Related to Physician Participation in Surveys.

Paul C Beatty, National Center for Health Statistics (pbb5@cdc.gov); David Woodwell, National Center for Health Statistics (dwoodwell@cdc.gov)

Obtaining physician cooperation in surveys is difficult because they operate under considerable time constraints and typically receive many survey requests. At the same time, the need for data from physicians is increasing, and greater guidance is needed on strategies to maximize their participation. Literature on physician respondents, anecdotal experience, and some studies conducted at the National Center for Health Statistics suggest that when physicians consider survey requests, they make judgments about legitimacy, benefit, and burden. Judgments about the legitimacy of the survey request are heavily influenced by communication about sponsorship, and are critical initial determinants of cooperation. Benefits to physician respondents can include both token incentives, and larger payments that attempt to provide realistic compensation for physician time; both have varying degrees of effectiveness, and we explore under which circumstances they are likely to have the greatest effect.

Whereas judgments about legitimacy and benefit can presumably be made quickly and early, judgments about burden may be more multi-faceted and can overrule legitimacy and benefit judgments at later points of the response process. Questionnaire length is often considered to be a proxy for burden, but is actually only one component—burden also includes question complexity and appropriateness of mode given response tasks. A study conducted at NCHS also suggests that increasing instrument length may reduce burden in some ways (e.g., by providing clearer and more thorough response options), producing some higher item response rates even though overall participation decreases. We conclude by considering the overall design elements that can affect the willingness of physicians to participate. Although in some respects similar determinants are involved in response decisions in general population surveys, physician circumstances may lead them to have different sensitivities to legitimacy, benefit, and burden. Taking these sensitivities into account may help optimize physician participation.

Drilling Down to Examine a Dentist Survey Response Rate: Lessons for Surveys of Professionals.

Colleen K Porter, University of Florida (cporter@dental.ufl.edu); Henrietta L. Logan, University of Florida (hlogan@dental.ufl.edu); Frank A Catalanotto, University of Florida (fcatalanotto@dental.ufl.edu); Heather G Steingraber, University of Florida (hgs@phhp.ufl.edu)

In August to December 2010, we conducted a mixed-mode survey of 1,201 Florida dentists who treat children, starting with a FedEx delivery, followed by an email reminder, postcard, email invitation to complete via the web, a second full mailing, and final email reminder. The final response rate was about 75%, with almost two-thirds of the completed surveys being received back within 7 days of initial delivery (a pattern that provides cost savings by eliminating the need for follow-ups). Factors contributing to the high response rate include the following: (1) a letter of support from the Florida Dental Association, which boasts membership of 90% of the state's dentists, (2) a sample that had gone through screening steps to improve probability that the dentist met the criteria for the study, (3) implementation by the university which is the state's primary provider of continuing education for practicing dentists and had the state's only dental school for years, (4) a \$10 non-contingent advance cash incentive included with the initial mailing, (5) a topic of great interest to many dentists (Medicaid policies for dental providers), (6) a questionnaire constructed on evidence-based principles of visual design for self-administered surveys, (7) initial delivery via 2d-day FedEx rather than mail, (8) a strategy of starting with mail and offering web later (to avoid response-reducing issues with too much choice that have been noted in recent years). To understand the contribution of each of these factors, we examined the literature for previous methodological studies: For example, studies had demonstrated that sending via FedEx rather than first-

class mail significantly increases response rates among professionals, that advance incentives are more effective than promised incentives, etc. This amalgam of survey practices was more effective at optimizing response than any of the individual components.

Questionnaire Design: Improving Data Quality.

Outside the Answer Boxes: Messages from Respondents.

Fran Featherston, National Science Foundation (ffeather@nsf.gov); Kristin Stettler, U.S. Census Bureau (kristin.j.stettler@census.gov)

The Business R&D and Innovation Survey, conducted jointly by the U.S. Census Bureau and the National Science Foundation (NSF), is mailed annually to 40,000 businesses. In 2008, a redesigned version of the questionnaire increased the number of pages for the long form from 18 to 56. The short form increased from 7 to 28 pages. We received 1,467 paper questionnaire responses and 2,446 web responses for the 2008 long form. As part of the evaluation of data quality produced by the redesigned questionnaire, we coded the extraneous information that respondents wrote on the paper survey forms. That is, we coded any information written outside the answer boxes. These messages from respondents included such things as strikeouts of categories, notations that questions were not applicable, and strikeouts of entire questions or pages of questions.

This paper examines the usefulness of this coding of paper responses for evaluating the design of a new questionnaire. We provide examples of messages from respondents concerning (1) data quality issues, (2) questionnaire design flaws, (3) data availability issues, and (4) interpretation of the content of the question. We also provide examples of data analyses that can be used to augment the information gleaned from respondent messages to provide a more rigorous tool for evaluating survey design.

Lessons Learned from the 2009 NHTS in Travel Behavior Measurement.

Adella Santos, Federal Highway Administration (Adella.Santos@dot.gov)

For four decades the National Household Travel Survey (NHTS) has provided a wealth of transportation information about the American Public's travel behavior. As one of the largest household travel survey in the world, NHTS has many uses, including congestion management, safety planning travel demand modeling, air quality analysis, and special population studies. In most large national surveys, the conduct of this survey presented methodological challenges in design, measurement, processing and estimation. This paper presents some of the lessons learned in the most recent iteration of the study (2009). The paper will cover changes in questionnaire design to increase data accuracy to capture travel related measurements such as trips rates, distance, mode and purpose of transportation. It will illustrate how placement of key travel indicators changed responses rates in comparison to the 2001 survey. It will focus on travel indicators of walking and biking for trip rates and distance and purpose of trips to show item response rate differences. We illustrate how question placement can have important impact on incidence estimates and data quality.

Instability Overlooked: Evidence of the Importance of Household Roster Data Collection and Matching Over Time.

Nola du Toit, NORC at the University of Chicago (dutoit-nola@norc.org); Cathy Haggerty, NORC at the University of Chicago (haggerty-cathy@norc.org)

Many studies on instability examine changes in relationships and household composition by using measures such as “are you married?” or “how many children in this household?” and then comparing the answers across time. Conclusions are drawn from any changes noted in respondents’ answers. However, these measures are one-dimensional; they do not capture the types of changes that may have occurred between waves. For example, respondents may be married at both waves, but to different people. This misconception calls for a more nuanced examination of households and relationships of members at each point in time.

The Making Connections Survey is a longitudinal study of residents in low-income neighborhoods in ten US cities. Baseline data collection started in 2002 and, as of this year, six of the original ten sites have completed a third interview. The survey includes a household roster that collects demographic information about each member of the household, as well as their relationship to the respondent. Household members on the roster are then matched across waves. This makes the data particularly robust in that it is possible to tell changes in the household composition and relationships on an individual basis, instead of relying on singular measures. That is, we are able to discern whether the respondent is still married to the same person, or if the children living in the household are the same children from the first wave.

This study examines the discrepancies in results between the single measures (e.g. “are you married?”; “how many children in the household?”) and the detailed roster-matching of household members across time. This study is important because it will illustrate, first, the importance of more detailed data collection on household members, and second, considering the fluid nature of poor families, how much instability is overlooked by the usual measures.

Who Missed the Skips?: Empirical Results from a Self-Administered Survey.

Jeremy Redford, Education Statistics Services Institute, Macrosys, LLC (ctr_jredford@air.org); Sarah Hastedt, National Center for Education Statistics (Sarah.Hastedt@ed.gov)

Self-administered questionnaire designers are always mindful of the impact of skip patterns, since incorrectly followed skip patterns can impact the quality of survey data. Understanding the characteristics of respondents likely to miss a skip pattern and the types of skips they miss can help survey researchers improve questionnaire design. As the National Household Education Surveys Program (NHES) moves from an RDD survey to a self-administered mail survey, questionnaire designers take special interest in the effectiveness of navigational instructions on the paper forms. Research on self-administered questionnaires has shown skip instructions can impact whether respondents correctly navigate the questionnaire, but the literature has been less conclusive on whether one type of instruction is superior to another. The 2009 NHES pilot study provides data on which skip patterns were missed and under what circumstances. This paper utilizes data from the Parent and Family Involvement in Education (PFI) survey of the NHES pilot. Approximately 750 respondents for the PFI survey completed questionnaires with different types of skip instructions: 1) a skip box that told the respondent what to do based on how they responded to previous questions, 2) brackets and arrows with “Go To” instructions, and 3) parenthetical instructions that told respondents who should answer the question. There are twenty-eight total skips analyzed among the three categories. This paper will explore whether demographic factors were correlated in the 2009 NHES pilot test to missing skip patterns in general or to specific skip instructions. The paper will analyze which skip patterns were followed incorrectly and general demographic characteristics of the respondents who missed skips including: education level, income, and age. In addition, the different skip instructions will be compared against one another to investigate the effectiveness of each skip type.

Technical Aspects of the Construction, Coverage, Limitations and Future of the DSF.

A Summary of Delivery Sequence File Coverage Research at the U.S. Census Bureau.

Timothy L Kennel, U.S. Census Bureau (tkennel@survey.umd.edu)

The Delivery Sequence File is the largest source of addresses for many Address Based Sampling Frames. The coverage of the Delivery Sequence File can have a large impact on the quality of Address Based Samples. The U.S. Census Bureau has performed a number of studies which illuminate the coverage of the Delivery Sequence File as well as several Address Based Sampling Frames constructed from the Delivery Sequence File. In this report, we summarize some of the key findings of this research. We provide estimated coverage rates for city-style addresses on the Delivery Sequence File by various geographies and lag times for new construction to appear on the DSF. Some preliminary research comparing Census 2010 address canvassing results to the Delivery Sequence File will be presented.

Constructing the Address-Based Frame: Present and Future.

Phillip L Faulstich, Valassis (plfaulst@valassis.com); Ned English, NORC (english-ned@norc.uchicago.edu); Ashley Amaya, NORC (Amaya-Ashley@norc.org); Michael Latterner, NORC (Latterner-Michael@norc.org)

Address-based (ABS) samples based on the United States Postal Service (USPS) Computerized Delivery Sequence (CDS) are becoming increasingly popular alternatives to landline-RDD and traditional listing. As ABS emerges as a main-stream sampling frame, researchers have encountered a number of new challenges and questions that come with the emergence of any novel methodology. Such questions include; understanding CDS data, coverage quality, response-rate calculations, and weighting techniques.

This paper takes a step-back from the above substantive questions to discuss the address frame itself. We seek to form a better foundation on how the frame is constructed and what changes may be anticipated in the future so researchers may formulate sampling, weighting, and analysis techniques that can most efficiently employ the frame and answer their questions.

First, we discuss the foundation of the address frame: how the USPS maintains its database file; how updates are performed in the postal data chain; and what types of addresses may be included and excluded. This section will also discuss the number and types of changes that appear and how quickly they are integrated. We will also focus more closely on areas generally perceived to be under-covered such as rural areas and unique addresses.

Second, we discuss the future of the address frame. The USPS is undergoing several changes and releasing additional data that will enhance the coverage and usability of the frame. We focus on three of these changes: the conversion of all addresses to city-style addresses (i.e. 123 Main St); supplying unit numbers to drop deliveries; and providing addresses that are currently not delivered to (i.e. new construction and unoccupied rural units).

Using Ancillary Information to Facilitate Address-Based Sampling in Rare Populations.

Ashley Amaya, NORC at University of Chicago (amaya-ashley@norc.org); Ned English, NORC at the University of Chicago (english-ned@norc.org)

Address-based sampling through the USPS delivery sequence file (DSF) has been presented as an alternative to random-digit dial (RDD) in recent years. The DSF, when purchased through a vendor, contains more than just address information which may prove useful for sampling. Some of this

information is supplied by the DSF itself – vacancy status, college housing, seasonal addresses, etc. Other non-address auxiliary data can be matched to the address by vendors from market research databases. These include items such as race/ethnicity, income, or other household characteristics.

Our paper examines the utility of the auxiliary non-address information on the DSF for developing sample designs. Auxiliary data could potentially improve household identification in addition to providing stratification information and/or target rare populations.

At question is what flags are available, what is the source of these flags (DSF, market research, etc.), how do the different categories of flags perform, and what trade-offs may exist between efficiency and coverage. Using a variety of NORC-funded research projects, this paper will discuss pre-existing flags, such as drop delivery, seasonal, and college indicators, and how they can be used to maximize sampling efficiencies via stratification. First, we will also describe the kind of demographic information that can be appended to the file to enrich it, both at the address and tract level. Second, we will evaluate how accurate flags based on race/ethnicity and others are and quantify the efficiencies gained (based on coverage, rates, and timelines). Third, we will explore the uses of the flags for mailing, specialty language interviewer assignment in CATI, and calling rules. Lastly, we will examine the efficiency gains as a trade-off with statistical efficiency (design effect and weighting variability).

An Evaluation of a Training Protocol Designed to Supplement the Coverage of an Address-Based Sampling Frame.

Joe McMichael, RTI International (mcmichael@rti.org); Allison C McKamey, RTI International (mckamey@rti.org); Bonnie E Shook-Sa, RTI International (bshooksa@rti.org); Stephanie J Stolzenberg, RTI International (sjstolz@rti.org); Katherine B Morton, RTI International (bowman@rti.org); Vincent G Iannacchione, RTI International (vince@rti.org)

Studies have shown that address-based sampling (ABS) frames for in-person surveys suffer from undercoverage of the household population, particularly in rural areas. To bridge this coverage gap, RTI International developed the Check for Housing Units Missed, or CHUM, (McMichael et al 2008) frame supplementation procedure for in-person surveys that use ABS as a sampling frame. The CHUM procedure enables field interviewers to locate dwelling units not included on the ABS list. When done correctly, an ABS frame supplemented with the CHUM can provide near-complete coverage of the household population. Incorrect implementation of the CHUM procedure can lead to undercoverage and/or frame multiplicities. Successfully training field interviewers to implement a frame supplementation procedure is essential to any study that utilizes an ABS sampling frame.

Correct implementation of the CHUM procedure requires field staff who are knowledgeable about the types of situations they will encounter when performing the procedure. To that end, RTI developed a training protocol for the CHUM procedure and evaluated it through an experimental set of field exercises designed to provide information about whether field staff could perform the procedure correctly; what procedures were the most challenging or difficult to perform in the field; and determine future improvements to the training materials and methods based on these findings. Field interviewers' performance was evaluated on multiple domains including which component of the CHUM procedure was being performed, the difficulty of the CHUM interval, and whether performance improved after the field interviewers gained experience with the procedure. This paper describes the training protocol, the experimental design developed to evaluate field staff performance, and an analysis of the data collected from the field exercises including logistic regression modeling of the factors that contribute to whether the CHUM procedure was implemented correctly.

The Challenge of Polling in Congressional Races: A View from 2010.

The Challenge of Polling Congressional Races: A View from 2010.

Jim Lee, Voter Survey Service/Susquehanna Polling & Research (james@susquehannapolling.com); Don Levy, Siena Research Institute (dlevy@siena.edu); Patrick Murray, Monmouth University Polling Institute (pdmurray@monmouth.edu); Seth Rosenthal, Merriman River Group (seth@merrimanriver.com); Mike Mokrzycki, Independent consultant (mike@mikemokr.com)

A great deal of attention, both in the media and in the polling community, is given to statewide and national polling. We rarely take a wide-ranging look at the unique challenges of public polling in Congressional districts. The 2010 election, where more than 60 House seats changed party, witnessed a large number of Congressional race polls. At least 30 different polling organizations issued House election polls during the final two weeks of the campaign.

U.S. House races are considered among the most difficult elections to poll. According to one report, the average error in assessing the final victory margin in 2010 Congressional District polling was more than seven percentage points. Issues with polling in House races include the lack of credible population parameters at the district level, difficulty isolating geographic boundaries for sample frames leading to either over- or under-coverage, and problems identifying district-level dynamics that are distinct from the national context in which the election is taking place. Furthermore, House race polls are covered most extensively by local media outlets who rarely have the knowledge base or resources to assess whether poll results come from a reliable source.

This roundtable discussion will examine some of the unique challenges posed by polling at the Congressional District level. The panel includes representatives of both independent and partisan polling organizations who were active in U.S. House polling this past election.

Web Survey Questionnaire Design.

Up Means Good: The Impact of Screen Position on Evaluative Ratings in Web Surveys.

Roger Tourangeau, JPSM (rtourang@survey.umd.edu); Frederick Conrad, Survey Research Center, University of Michigan (FConrad@umich.edu); Mick Couper, Institute for Social Research, University of Michigan (MCouper@umich.edu)

This paper presents results suggesting that respondents use the position of the item on screen as a cue in evaluating the item. People may expect positive things to be higher up on the screen than bad things. We dub this expectation the “Up means good” heuristic. We tested this notion in a series of experiments embedded in four web surveys. The first and third experiments testing this heuristic show that respondents are faster to answer when the positive end of the scale is at the top than when the negative end is; the first experiment also shows that the order of the scale points makes no difference when the scale points are arrayed horizontally. The second, third, and fourth experiments shows that items are rated more favorably when they are positioned near the top of the screen than when they appear lower down on the screen. These experiments vary in how they manipulated the position of the items on screen. Taken together, the studies present further evidence that respondents may rely on incidental visual cues in making judgments as they complete web surveys. Many survey judgments are known to be context-sensitive. The screen position may also be a contextual variable for evaluative judgments in web surveys.

Investigating The Impact Of Scale Height, Width, And Spacing On Responses In Web Surveys.
Hanyu Sun, Joint Program in Survey Methodology (hsun@survey.umd.edu)

The impact of visual features of questions on survey responses may be more marked in web surveys than in self-administered paper questionnaires, since web surveys often include richer visual material. Existing research on web surveys shows that respondents in web surveys may use visual cues to help interpret the questions or answer options. For example, respondents may make inferences about the conceptual range covered by a response option based on its height or width. Options that take up more space on the screen may be seen as representing a broader conceptual range than those encompassing less space. As a result, we predicted that respondents would be more likely to select a response option when it appears wider on the screen. We tested this prediction in an experiment embedded in a web survey. The experiment presented the response options for a set of questions either vertically or horizontally; in addition, the response options were spaced evenly, with gaps between two of the options, or with one of the options presented on multiple lines — that is, some response options took up more space on the screen than others. This presentation presents the findings from the results from this experiment as well as the results from five similar web experiments that examined the effects of height, width, and spacing on responses in web surveys. In addition, practical guidance on the design of response scales in Web surveys based on the results from these surveys will be provided.

Visual Design and Usability of Filter and Follow-up Questions in a Web Survey.
Carl M Ramirez, US Government Accountability Office (ramirezc@gao.gov)

How to best present a series of filter items and the sets of follow-up items conditional on each filter has been a recurring questionnaire design challenge in self-administered surveys, and to a lesser extent, interviewer-administered surveys. Two approaches are commonly used: Ask all the filter questions first, then instruct respondents to complete one set of conditional questions for each filter question they previously endorsed (“grouped” structure). The second approach is to ask one filter question at a time, and use branching instructions and visual cues to immediately answer or skip over the corresponding follow-up items (“sequential” structure). Each navigational approach may have implications for visual design, respondent conditioning, perceptions of burden, and satisficing behavior. Using CATI-like navigation programming, web surveys can employ a hybrid approach: Filter questions using hyperlinks that, when the response condition is met, immediately open popup subforms with follow-up questions, and upon sequential completion of that filter’s set of follow-ups, return the respondent to the list of filter questions. One advantage of this hybrid approach is that respondents see a short, uncluttered list of filter questions, without skip patterns, as in the grouped approach, but also can be sequentially presented with applicable follow-up questions immediately, avoiding the need to recall answers to earlier filter questions. The implementation of this approach in one web survey interface, however, produced patterns of item nonresponse and suboptimal navigational behavior that reinforce several textbook principles of questionnaire visual design and usability. The response effects are described, and possible remedies suggested.

An Investigation of the Impact of Stylistic Features on Web Survey Responses.
Rebecca Medway, Joint Program in Survey Methodology, University of Maryland (rmedway@survey.umd.edu); Roger Tourangeau, Joint Program in Survey Methodology, University of Maryland (rtourangeau@survey.umd.edu)

Lack of respondent engagement is a concern in web surveys. In an attempt to combat this problem, survey practitioners often try to design web questionnaires in ways that they believe respondents will find

visually appealing. However, prior research has found that, when answering survey questions, respondents may draw on features that researchers intended to be purely stylistic. For example, when the appearance of response option labels varies across the options, respondents may perceive the options to be further apart conceptually than when they are presented in a similar fashion to one another. This may in turn affect the survey responses that respondents provide.

This presentation will discuss the results of an experiment included in a web survey with more than 2,000 respondents. The respondents were randomly assigned to one of four experimental conditions that varied the format of the response options: one group got standard labels for each option, the second got labels that increased in boldness from left to right, the third got labels that decreased in boldness from left to right, and the fourth group got labels with emoticons. In addition, respondents were randomly assigned to one of two response option orders, and to either a fully-labeled or endpoint-only-labeled scale. The results of this experiment indicate that these three factors may influence survey responses, suggesting that practitioners should use caution when adding stylistic features to web questionnaires.

Saturday, May 15, 2011

10:00 a.m. - 11:30 a.m.

Concurrent Session B

AAPOR's Transparency Initiative – Progress & Prospects.

AAPOR's Transparency Initiative - Progress and Prospects.

Peter V Miller, Northwestern university (p-miller@northwestern.edu); Reg Baker, Market Strategies; Tim Johnson, University of Illinois - Chicago

The AAPOR Transparency Initiative, an effort to encourage openness in the practice and publication of survey research, has been joined by scores of supporting research organizations. Over the past year, the Initiative's steering committee has been working to put mechanisms of research disclosure into practice. This panel will discuss the issues faced by the committee, its consultation with supporters, the current status of the Initiative and plans for continuing effort.

Advanced Statistical Methods for Exploring Interviewer Effects in Behavior Coded Data and Survey Pretest Data.

Interviewer Effects in a Small-Scale Experimental Research Survey of Demographic Questions.

Elizabeth Nichols, U.S. Census Bureau (elizabeth.may.nichols@census.gov); Jennifer Hunter Childs, U.S. Census Bureau (jennifer.hunter.childs@census.gov); Rolando Rodriguez, U.S. Census Bureau (rolando.a.rodriguez@census.gov)

In November 2006, the U.S. Census Bureau fielded a RDD CATI survey called the Questionnaire Design and Experimental Research Survey which had two experimental panels. A total of 1870 interviews were completed, approximately evenly split between the two panels. Ten different interviewers were assigned to the each of the two panels. Because of a lack of resources, interviewers did not rotate through the panels. In an attempt to balance the panels, staff used amount of interviewing experience to make the two interviewing groups approximately equivalent.

The primary objective of this experiment was to test different questionnaire designs for conducting coverage measurement research. Additionally, researchers embedded an experiment focusing on how to confirm age, given a valid date of birth. We analyzed respondent burden, as measured by interview length, as one of the comparisons between the two experimental panels. This paper shows how results of regression models predicting interview length change when accounting for the small number of interviewers through fixed effects as compared with adding a random interviewer effect. This paper demonstrates the importance of controlling for interviewer effects, even in a relatively simple, small-scale experiment.

Using Random Effect Models, Intraclass Correlations, and Design Effects to Explore Interviewer Effects on Respondent Behavior and Psychological States in Dyadic Data.

Matt Jans, U.S. Census Bureau, Center for Survey Measurement (matthew.e.jans@census.gov)

Coding behaviors of survey interviewers and respondents at each conversational turn can be a helpful way to understand the interviewer-respondent dynamic. However, the complexity of these dyadic data (e.g., coded observations nested within each interview-respondent pairs, where each code is not statistically independent, but determined partially by previous codes) makes it difficult to determine the effect that interviewers have on respondent behavior. Based on coded interviewer and respondent turns from 185 telephone interviews conducted by the Reuters/University of Michigan Surveys of Consumers, this analysis treats respondent codes as nested observations within interviewers by using random interviewer effects models, the resulting intraclass correlations (i.e., rho-ints), and design effects of interviewers (i.e., deff-ints). These statistics are taken as exploratory indicators of “interviewer effects” on coded respondent behavior and rated psychological states, specifically cognitive difficulty and affect. Results show moderate deff-ints (ranging 2.14 - 2.81) for several of the indicators evaluated. These include amount of overspeech (e.g., intentional or accidental interruptions), number of total utterances (e.g., length of the exchange), request for clarification or repeat, perceived difficulty ratings, proportion of utterances with respondent laughter, respondent laughter, rated affect valence (positive or negative), rated affect intensity, rated cognitive difficulty, and conversation management (e.g., otherwise-uncodable behavior that moves the conversation forward). The approach and results are presented in the context of social psychological theory about interviewer-respondent interactions, with specific focus on the use of advanced statistical tools (e.g., random intercept models) for measuring interviewer effects.

The Effect of General Interviewer Experience on Item Non-Response: A Verbal Behavior Study.
Ipek Bilgen, University of Nebraska-Lincoln (bilgenipek@yahoo.com); Robert F. Belli, University of Nebraska-Lincoln (rbelli2@unlnotes.unl.edu)

During survey data collection, the number of missing items has been observed to differ between experienced and inexperienced interviewers (Cleary et al., 1981; Kennickell, 2002; Singer et al., 1983). However, the results regarding the direction of this effect are equivocal and the potential competing mechanisms are yet to be explored. In this study, we aim to disentangle the mediating mechanisms behind the relationship between interviewer experience and item nonresponse via investigating the change in interviewer verbal behaviors due to experience. In addition, as the interviewing technique—whether standardized or flexible calendar—has been found to have a critical impact on response errors (Belli et al., 2007), we examine the association between interviewer experience and item nonresponse in both standardized and calendar interviews. The current study uses a nationwide sub-sample of participants from the 2001 Panel Study of Income Dynamics (PSID) who were interviewed via telephone retrospectively about their life course events. Respondents were randomly assigned to either Event History Calendar or Standardized Conventional Interviewing methods. In our analyses, we measure general interviewer experience via years of experience gained during interviewer’s lifetime and focus on respondent’s final don’t know (DK) responses as the item nonresponse measure. The general hypotheses are that the direction of the interviewer experience and item nonresponse association will vary depending on how interviewer and respondent verbal behaviors change in different interviewing methods as interviewers gain experience, and that the impact of interviewer experience will be stronger in calendar interviews than in standardized interviews due to the flexible nature of the calendar interviews.

Exploring the Reliability of Behavior Coding Data.
Jennifer Hunter Childs, U.S. Census Bureau (jennifer.hunter.childs@census.gov); Nathan Jurgenson, U.S. Census Bureau (nathan.jurgenson@census.gov)

Behavior coding examines interviewer and respondent interactions, often with the intent of evaluating the quality of survey questions. In practice, several coders listen to these interactions and code prescribed behaviors. Often the coders are asked to code the same cases, to be used as a measure of reliability of

the coding. If the reliability is poor, then the results of the coding are subject to coder effects and may be less meaningful. The kappa statistic is often used to measure reliability, as a conservative measure of reliability that corrects for chance agreement. This paper explores different ways of measuring reliability using the kappa statistic and what can be learned by examining reliability of the codes in different ways. For example, we will examine reliability scores by speaker (e.g., interviewer and respondent), by question (e.g., age, race, home ownership), by code (e.g., exact question reading, major change in question wording, adequate respondent answer, request for clarification) and by type of question (e.g., select one, multiple choice, use of a flashcard). Examining reliability in these ways will allow us to recommend a strategy for reporting reliability in future behavior coding studies as well as learn more about the behavior coding method itself and characteristics of questions that can cause problems.

Cell Phone Sampling & Weighting.

Propensity Model for Weighting Dual Frame Telephone Samples.

Robert Benford, GfK Custom Research North America (robert.benford@gfk.com); Trevor Tompson, Associated Press (ttompson@ap.org); Julian Baim, GfK MRI (julian.baim@gfkmri.com); John Lien, GfK CRNA (john.lien@gfk.com); Lancey Heyman, GfK MRI (lancey.heyman@gfkmri.com)

Propensity models are a potential solution to compensate for non-coverage and non-response, problems that threaten telephone surveys even when frames do cover the entire population of inference and where response tendencies by frame may lead to differential response or potential bias by frame. Dual-frame RDD (landline and cell) surveys improve coverage of the entire adult population and possibly improve non-response when dual phone service households are included regardless of sample frame. In this paper we model cell only versus dual service persons as a function of demographic and attitudinal variables found in Mediamark Research & Intelligence (GfK-MRI). This reference source provides a robust set of potential covariates for modeling purposes. In the paper we develop the model, compare the application of this model to our current first stage pre-weight practice over two AP-GfK polls, and apply the model to the landline only sample. Further, the paper addresses the bias and variance components that result from the different weighting approaches.

Dual Frame Weighting and Estimation: Considerations when Incorporating Cell Phone Sample within a Telephone Survey.

Elizabeth Ormson, NORC at the University of Chicago (ormson-elizabeth@norc.org); Kennon R. Copeland, NORC at the University of Chicago (copeland-kennon@norc.org); Stephen J. Blumberg, National Center for Health Statistics (swb5@cdc.gov); Kirk M. Wolter, NORC at the University of Chicago (wolter-kirk@norc.org); N. Ganesh, NORC at the University of Chicago (ganesh-nada@norc.org)

Telephone surveys have traditionally utilized random-digit-dial (RDD) samples from the landline telephone household population. Given the growth in wireless-only households (representing 24.5% of all households at the end of 2009, Blumberg and Luke, 2010), telephone survey designers are exploring new strategies in sample design and weighting to avoid resultant coverage bias in survey estimates. One obvious solution is to include a cell-phone sample within the survey. While incorporating a cell-phone sample is relatively easy to implement through a dual-frame design, the greater costs to complete interviews with cell phone sample often lead to undersampling of the cell-phone frame. A challenge is how to create dual frame weights to appropriately avoid coverage bias while minimizing variance increases due to disparate sampling rates between landline and cell samples.

The National Survey of Children with Special Health Care Needs (NS-CSHCN) is designed to produce estimates at the state level. A landline RDD sample design was utilized during the first four quarters (Q3/2009 – Q2/2010) of data collection. The RDD sample was then supplemented with a nationally representative cell phone sample of households (Q3/2010 – Q4/2010) to provide coverage of the growing wireless-only population. This design resulted not only in different sampling rates for landline and cell frames, but also varying differences in sampling rates across states. Single frame weights for the landline sample can be derived at the state level, but dual frame weights for the combined landline plus cell phone sample in this design need to consider the geographic levels at which weights can be derived.

This paper focuses on alternative weighting approaches for such a dual frame landline and cell-phone sample. Discussion will not only address the issues of coverage bias and estimate variance, but also the issues of level of estimation and appropriate population controls, with applications to the NS-CSHCN.

An Evaluation of Popular Weighting Approaches in Dual Frame RDD Surveys.
Courtney Kennedy, Abt SRBI (kennedy@srbi.com)

There is substantial variation in how survey statisticians weight data from overlapping landline and cell phone designs. According to the 2010 AAPOR Cell Phone Survey Task Force Report, "...there is no consensus regarding how RDD cell phone samples should be weighted, especially when combining them with RDD landline samples." This state of affairs can make it difficult for survey sponsors and other consumers to determine whether the weighing protocol implemented for a given dual frame RDD survey was appropriate. We bring empirical research to bear on this issue by evaluating several popular weighting protocols. The data come from five overlapping landline and cell RDD surveys of the general population conducted in 2010. We compute four separate weights for each study according to popular methods currently being used in the industry. These methods include the 0.5 compositing factor, calculating the compositing factor based on the effective sample sizes, adjusting for the probabilities of selection in each frame, and estimating the dual user response rate in each sample. We compare the weighting approaches with respect to their impact on the reliability of survey estimates (design effect) as well as their computational demand.

Stratification of Cell Phones: Implications for Research.

David Dutwin, Social Science Research Solutions (ddutwin@icrsurvey.com); Kathleen Call, Shadac (callx001@umn.edu); Donna McAlpine, University of Minnesota (mcalp004@umn.edu); Tim Beebe, The Mayo Clinic (Beebe.Timothy@mayo.edu); Robyn Rapoport, SSRS (rrapoport@ssrs.com)

Telephone studies have commonly utilized stratification for two main purposes. The first is to ensure adequate statistical power of populations of interest who cluster by geography, telephone exchange, or listed telephone information. But as well, stratification is then often utilized in non-response weighting adjustments. Since little is generally known of non-responding households (their race, age, etc), researchers are often limited to useful geographic definitions of households with which to make nonresponse weighting adjustments, and stratification is often one of the better variables of choice. With the increase in dual-frame (landline and cell phone) telephone research, however, researchers are confronted with a second frame (the cell phone frame) in which there is no useful listed data and no database readily available to map exchanges to specific geographies. And, even if such a database existed, it could be of dubious value given the mobility of cell phone owners and their cell phone numbers. For the 2009 Minnesota Health Access Survey and the 2010 Minnesota Drug and Alcohol Needs Assessment Survey, however, cell phones were stratified by cellular switch point. Research questions

particular to this paper are first, what was the “hit rate” of actually reaching a person by strata who, in fact, lived in that strata, as opposed to some other strata or another state entirely? Second, how does this compare to the hit rate of similar landline strata? Furthermore, does response vary significantly by the strata or by specific switch points, and if so, what are the implications and efficacy of utilizing strata or specific switch points for nonresponse adjustments in the weighting? And finally, what difference in survey estimates does such nonresponse weighting lead to in cell phone samples, comparatively to similar adjustments in landline samples?

Issues in Multi-Mode Data Collection.

How Do Respondents React When Asked to Self-Report Their Behavior?

Marla D. Cralley, Arbitron Inc. (marla.cralley@arbitron.com); Kelly M. Dixon, Arbitron Inc. (kelly.dixon@arbitron.com)

Often self-report survey instruments such as time-use or diet diaries are used to measure specific respondent behaviors. Concerns arise around the accuracy of the reports. Do respondents accurately report their behavior, do they embellish their reports in an attempt to satisfy the researcher, and/or do they actually alter their behavior during the survey period?

Arbitron PPM(TM), a system to passively collect Radio and Television media use over time among an on-going panel of respondents, has been implemented in 43 top U.S. metros, replacing the traditional paper Radio and Television self-report diaries previously used in the markets. During the Summer of 2008, Arbitron asked a test sample of PPM(TM) panelists to keep a one-week Radio diary while continuing to wear their meters.

This paper will present the results of the study, comparisons of Radio listening passively collected through the PPM(TM) technology with respondent self-reports for the same time period. Patterns of differences will be compared and contrasted by demographic characteristics. In additions, passively collected listening information from diary keeping panelists will be trended across several weeks to explore differences in actual behavior correlated with the diary keeping task.

Who Follows The Rules? Within Household-Selection Procedures In Web And Mail Surveys Of The General Population.

Kristen Olson, University of Nebraska-Lincoln (kolson5@unl.edu); Jolene Smyth, University of Nebraska-Lincoln (jsmyth2@unl.edu); Mathew Stange, University of Nebraska-Lincoln (mathew.stange@huskers.unl.edu); Lauren Walton, University of Nebraska-Lincoln (lauren.walton8008@gmail.com)

Household surveys are increasingly moving from interviewer-administered modes to self-administered modes for data collection. With this move to self-administered questionnaires comes the problem of how to select a respondent within a household. Recent research has shown that many households do not appropriately follow within household selection procedures in mail surveys (Battaglia, et al., 2008). Yet this question has not been evaluated in web surveys of the general population or in mixed mode mail and web surveys. In this paper, we examine accuracy of within household selection using an oldest adult/youngest adult method in self-administered surveys. The frame for this study comes from a telephone survey conducted with Nebraska residents in which the oldest adult/youngest adult method is used to select the initial respondent. One year later, these telephone participants were randomly assigned to be followed up via a mail survey, a web survey, or one of two mixed mode surveys (mail with web

follow-up, web with mail follow-up) using identical oldest adult/youngest adult household selection methods (n=566, AAPOR RR1 46%). Additionally, the sex of the previous respondent was provided to help facilitate identical selection of respondents. Yet surprisingly high rates of discrepancies occur when comparing the two sets of respondents. For example, 6.5% of respondents to the follow-up study differ on sex and 13.3% differ by at least three years of age. This paper will examine characteristics of people who appear to have followed the selection procedures appropriately compared to not following them; we also examine whether discrepancies in respondent selection vary across survey mode. We also examine predictors for discrepancies based on one criterion but not another (e.g., matching on sex but not on age). Preliminary analyses indicate that, as expected, mismatches are more likely to occur in larger households.

Attitudes Toward A-CASI And IVR To Measure Elder Mistreatment Among Older Adults: Results From A Population-Based Survey Experiment.

Scott Beach, University of Pittsburgh, University Center for Social & Urban Research (scottb@pitt.edu); Richard Schulz, University Center for Social and Urban Research, Univ. of Pittsburgh (schulz@pitt.edu)

It is difficult to derive sound population-based estimates of elder mistreatment. New survey technologies like audio computer-assisted self-interviewing (A-CASI) and interactive voice response (IVR) have been shown to increase reports of sensitive behaviors relative to human interviewers, but have not been used with older adults to measure elder mistreatment. Our group recently reported results from a randomized survey experiment (n=903) that showed (1) older adults (60+) were able to use survey technologies like A-CASI and IVR, and (2) these privacy-enhancing techniques resulted in higher prevalence rates of financial exploitation and psychological mistreatment than interviewer-based CAPI and CATI. This paper provides analyses of three follow-up attitudinal questions administered to 370 older adults who had agreed to use A-CASI (n=182) or IVR (n=188): (1) perceived ease of use of the technology, (2) perceived privacy of the technology versus a human interviewer, and (3) preference for the technology versus a human interviewer. Analyses focus on differences in perception between A-CASI and IVR.

Sociodemographic variables (gender, age, race, education, marital status, household composition), physical disability, cognitive function, depression, social network and social support, and previous use of e-mail are explored as potential covariates. Preliminary results show that, while perceived ease of use of the two technologies was equal, A-CASI was more likely than IVR to be perceived as increasing privacy, and more likely to be chosen versus a human interviewer. In addition, African American elders found the technologies more difficult to use than Whites, and were also less likely to see privacy advantages and have preferences for A-CASI versus human interviewers. Results have implications for the use of survey technology to measure sensitive topics among older adults.

Comparing Response Propensities between a Web Mail Multi-Mode Survey and a Mail Survey.

Zeynep Suzer-Gurtekin, ISR - University of Michigan (tsuzer@umich.edu); Mahmoud Elkasabi, ISR - University of Michigan (mkasabi@umich.edu)

The Survey of Consumer Attitudes (SCA) is a monthly national RDD landline telephone panel survey conducted by the Survey Research Center at the University of Michigan. Since 2009 SCA has been investigating efficient and effective protocols for mail only, and multi-mode web and mail survey data collection in order to improve coverage and operation costs. These alternative modes are a response to increasing undercoverage, nonresponse, and operation costs in RDD landline telephone surveys (Curtin, Presser, and Singer 2005). Current thinking among some survey methodologists and practitioners is that there is an underlying attitude toward mode preference that subjects use to decide to participate in a survey. When multiple mode options are available, there is an expectation that a subject's propensity to

respond will increase. Despite this expectation empirical evidence to date is mixed (Couper and Miller 2008). This paper presents findings from the SCA web-mail survey tests. Among the findings, the web-mail multiple mode choice had response rates that were less than the mail only mode, which had survey response rates of 68.8% with a \$5 prepaid incentive, or 54.5% without incentive. We explore significant predictors of response propensities in the web-mail survey examining data on both respondents and nonrespondents available from a prior wave of interviewing.

References:

Couper, Mick, and Peter Miller. 2008. "Web Survey Methods." *Public Opinion Quarterly* 72(5) 831–835.
Curtin, Richard, Stanley Presser, and Eleanor Singer. 2005. "Changes in Telephone Survey Nonresponse over the Past Quarter Century." *Public Opinion Quarterly* 69:87-98.

Methodological Briefs: Improving Questionnaires.

A Question Order Effect for Opinions about Tax Increases.

John Tarnai, Social & Economic Sciences Research Center (tarnai@wsu.edu); Tim Faiella, Social & Economic Sciences Research Center (tfaiella@wsu.edu)

A variety of kinds of question order effects are possible, as described by Schuman and Presser (1981). We investigated one type of order effect (Part-Part Contrast Effects) using questions about willingness to have taxes increase by some percentage to help balance a state budget. The questions were included in a statewide survey about economic conditions in the state. A total of 500 households participated in the study conducted in 2009. The questions varied the amount of tax increase between 5% and 10% and randomized the order of the questions in the survey. The results clearly demonstrate statistically significant differences in the percent of respondents willing to endorse tax increases depending on the order in which the questions are answered. A greater percentage of respondents are willing to endorse the 10% tax increase if this question is asked first, than if it is asked after the 5% tax increase question. Similarly, a greater percentage of respondents are willing to endorse the 5% tax increase if this question is asked first, than if it is asked after the 10% tax increase question. However, the order effect is not as great with the 5% question, than it is with the 10% question. The paper discusses the implications of these results for designing survey questions about public opinion regarding tax increases.

How the Order of Response Options in a Running Tally Can Affect Online Survey Estimates.

Tom Wells, The Nielsen Company (twells_17@comcast.net); Mario Callegaro, Google (callegaro@google.com); Charles DiSogra, Knowledge Networks (cdisogra@knowledge networks.com)

In the design of online surveys, running tallies or constant sums are used to help respondents sum up the allocation of amounts so that the total sums to 100%.

We hypothesized that for time allocation, the order of the presentation of the time categories could make a difference in the distribution of reported time spent. We expected primacy effects, with the first-presented time category having a higher allocation of time than the later-presented options.

Two independent experiments were conducted, one on a sample of Google online advertisers, and one on a general population adult sample from KnowledgePanel®.

In the Google experiment, advertisers were asked to estimate percent of time to complete an online task consisting of 3 subtasks. The order of the subtasks was randomized. Consistent with our hypothesis, we

found primacy effects -- the subtask presented in position 1 obtained a higher mean allocated time than when presented in later positions.

In the KnoweldgePanel experiment, respondents were asked to provide running tallies of the percentage of television they typically watch during the morning, afternoon, and evening (separately for weekdays and weekends). As before, the order of the subtasks was rotated. Primacy effects were detected again, however differences by position were small and not statistically significant. Because time spent watching TV is a regular activity, viewing patterns are more likely to be encoded or ingrained in memory, and more likely to be reported reliably, with responses less susceptible to order effects.

In the second experiment, we also tested the restriction that running tallies must sum to 100%. Roughly 25% of respondents were prompted for out-of-range sums on the first tally (weekday) and 5% were prompted on the second tally (weekend). This drop-off in prompting indicates that respondents quickly learned what is required to complete a given task.

Effects of Cue Lists and Question Format on Survey Response.

Lisa Lee, NORC (lee-lisa@norc.org); Parvati Krishnamurty, NORC (krishnamurty-parvati@norc.org); Rene Bautista, NORC (bautista-rene@norc.org)

A common survey task is to ask respondents to recall instances that fit into a particular category. Survey questions often include cue lists—examples of items from the category being asked about. For example, in the National Longitudinal Survey of Youth (NLSY) Round 9 questionnaire, respondents are asked: Do you have any credit cards, such as Visa, American Express, or credit cards for specific stores, such as department stores or gas stations? Cue lists are thought to reduce survey measurement error by providing information for the respondent to interpret what the question is about, what types of responses should be included, and providing retrieval cues for the relevant information. However, the literature is divided on whether cues facilitate or inhibit response, whether categories (“fruit”) or instances (“apple”) are better, and how the effects of cues may be mediated by the question format. A web experiment was conducted in which respondents were asked questions about the convenience foods and the healthy foods they had eaten in the last seven days. The experiment varied the number of cues that respondents were presented (zero, two, or four cues) and the type of cues presented (category vs. instance). In addition, respondents were presented with one of three question formats: yes/no (Did you eat convenience foods...), frequency (How many times...), and enumeration (What convenience foods...). The findings indicate that although the effect of number of cues was slight, both type of cue and question format influenced responses. The results demonstrate how the contents of a cue list and question format can influence survey estimates and suggest implications for how survey questions should be designed.

Designing Questions for Web Surveys: Effects of Check-List, Check-All, and Stand-Alone Response Formats on Survey Reports and Data Quality.

Jennifer Dykema, University of Wisconsin Survey Center (dykema@ssc.wisc.edu); Nora Cate Schaeffer, University of Wisconsin Survey Center (schaeffe@ssc.wisc.edu); Jeremy Beach, University of Wisconsin Survey Center (jpbeach@wisc.edu); Vicki Lein, University of Wisconsin Survey Center (vlein@ssc.wisc.edu); Brendan Day, University of Wisconsin Survey Center (bday@ssc.wisc.edu)

For self-administered instruments, questions asking if respondents engaged in a series of different behaviors can be formatted using:

- Check-all response formats in which respondents are asked to mark all of the categories that apply under the assumption that marked categories are equivalent to “yes” and unmarked categories are equivalent to “no;”
- Check-lists response formats in which questions are presented in a grid or matrix and “yes” or “no” are offered as explicit choices; or
- Stand-alone response formats in which each question is presented separately and with its own set of “yes” and “no” categories.

For web surveys, the small body of research comparing check-all to check-list formats shows that check-lists are associated with respondents endorsing a larger number of the individual questions (providing higher counts of the behaviors) and taking longer to answer the questions (possibly reflecting deeper processing of the items). However, other research on web surveys demonstrates that long questions and grid/matrix formats may encourage breakoffs. Thus, a plausible alternative to either the check-all or check-list format is to format items as stand-alone questions. To the best of our knowledge no one has compared all three of these response formats simultaneously and in a web survey, and that is the focus of the current experiment. We examined the three response formats for several series of questions about news consumption and social activities in a web survey of college students (RR1 18.3%). We examine how the formats differ with regard to the following outcomes: breakoff rates, endorsement of the individual items, completion times, primacy effects, and criterion validity.

Exploring Visual Design Effects In Web Survey Questions. An Experimental Study Using Animated Answer Scales.

Matthias Emde, Darmstadt University of Technology (emde@ifs.tu-darmstadt.de); Marek Fuchs, Darmstadt University of Technology (fuchs@ifs.tu-darmstadt.de)

The rapid development of the internet and internet data collection methods offers a multiplicity of research opportunities. Even though web surveys became a commonly used approach, research in this field still has to cope with coverage, non-response and self-selection issues. Beyond these challenges in sampling design, web surveys still mimic paper questionnaires with respect to their layout and appearance. Even though the internet offers various graphic and multimedia design features, very little is known about the influence of animated web surveys especially for the question-answer-process. Research has focused on designs reducing non-response while the influence of visual presentation of questions itself is still vague. In a survey among the university journal readers (N=1,000) we conducted an experiment exploring the effects of implementing visually animated smiley questions. Respondents were randomly assigned to one out of three versions: The first version presented a fixed design, the smileys in version two changed their color and size when the cursor hovered over while in version three the smileys zoomed out and a text answer category was displayed. In addition every respondent had to answer a control item referring to the same content as the smiley questions on a rating scale.

Results will help understanding the direct influence of visual design elements on the question-answer-process. The use of animated scales and in particular of smiley scales is often considered critically since answers are generated on a more emotional affective basis than a rational cognitive one. By varying the accentuation of cognitive and affective elements in combining smileys and text elements we will elaborate the direct influence on the question-answer-process and on data quality.

Results of a Survey Experiment on Frequency Reporting: Religious Service Attendance from the 2010 ANES Panel Recontact Survey.

Matthew DeBell, Stanford University (deb主@stanford.edu); Lucila Figueroa, Stanford University (lucilaf@stanford.edu)

Surveys regularly measure the frequency with which respondents engage in particular activities, such as attending religious services, exercising, or watching or reading the news. Accuracy depends on the wording and format of the questions. This report assesses responses to two sets of questions on the frequency of religious service attendance that were asked on the ANES 2010 Panel Recontact Survey, which was administered on the Internet. Respondents were randomly assigned either to receive a “traditional” question based on standard GSS and ANES practice, which asked respondents to choose from a list of seven frequency categories, or to receive an “optimized” version of the question that asked respondents to report the actual number of times they attended services during a reference period of their own choosing. The paper will compare rates of item nonresponse, response distributions, elapsed time to complete the questions, and predictive validity for political outcome variables for the two question designs. The paper will recommend the preferred version of the question to use when accuracy is a priority and will assess the appropriateness and cost-benefit tradeoffs of an alternative question format when administration time is prioritized above accuracy.

Asking Sensitive Questions: Do They Affect Participation In Follow-Up Surveys?

Ines Schaurer, GESIS – Leibniz Institute for the Social Sciences (ines.schaurer@gesis.org); Bella Struminskaya, GESIS - Leibniz Institute for the Social Sciences (bella.struminskaya@gesis.org); Lars Kaczmirek, GESIS - Leibniz Institute for the Social Sciences (lars.kaczmirek@gesis.org); Wolfgang Bandilla, GESIS - Leibniz Institute for the Social Sciences (wolfgang.bandilla@gesis.org)

It is a well-known finding that sensitive survey questions such as the income question show an increased item nonresponse compared to other questions. Less is known about how and whether such questions affect the decision process of respondents to participate in follow-up surveys.

The analysis of a previous survey of the general population in Germany where respondents were asked to take part in a follow-up survey shows that 47% of those who answer the income question are willing to participate in a follow-up survey. In contrast, this figure plunges to only 22% for those who decline to provide income information (total n=818). There are different possible explanations for this finding. Firstly, those respondents with a lower extent of trust might have a higher probability for item nonresponse and likewise a higher probability to reject the follow-up survey participation. Secondly, asking the sensitive question itself could influence the later decision (especially for those people who are not totally convinced).

The current study uses an experimental design to disentangle these two possible processes and explanations. The random split experiment was conducted during a telephone survey which included a request for participation in online follow-up surveys. Half of the respondents received a question about their net household income, while the other half was not asked any income information. At the end of the interview all respondents were asked whether they are willing to join an online panel. The study tests the causal relationship of sensitive questions and subsequent survey participation. We also discuss the connection between item nonresponse on sensitive questions and the decision to participate in subsequent surveys. Finally, we give recommendations concerning the tradeoff of asking such questions in the first interview compared to later follow-up surveys.

Re-examining Classic Measures in Public Opinion Surveys.

Is Life Getting Better Or Worse? A Look at the Order Effect on the "Ladder of Life" Question in Public Opinion Surveys.

Wendy Wang, Pew Research Center (wwang@pewresearch.org); Rich Morin, Pew Research Center (RMorin@PewResearch.org)

Introduced by Hadley Cantril and the Gallup Organization in the mid-1960s, the "ladder of life" battery of questions has become a classic survey question measuring quality of life. Respondents were asked to give a numerical rating to their present quality of life, on a scale of zero to ten. Then using the same scale, they are asked to rate what their life had been like five years earlier, and what they expect it to be like five years from now. The long trend since the 1960s shows some fluctuations in ratings in different time periods, although the overall trend has been relatively flat.

Using recent data from a May 2010 Pew survey and previous Gallup and Pew surveys dating back to 1964, we explore the possible order effect on the "ladder of life" question. In a random split-form design included in the May survey, half the sample was asked the ladder series immediately after an opening question of how satisfied they were with the way things were going in the country. The other half of the sample was asked the ladder question after they were asked what their day had been like. Both questions have been used in the past to introduce the "ladder of life" question. The effect was large: a 17 percentage-point difference in how people saw their present lives, depending on which opening question they had heard.

We look at how this effect varies by demographic characteristics such as gender, age, race and how long this order effect may "linger" in the questions that followed. In addition, we will also compare the context of the ladder of life question in the past surveys. In doing so, we hope to draw attention of survey researchers to the context of the questions when designing the questionnaire and comparing trends.

The Measurement of Political Ideology.

**Jon D Miller, International Center for the Advancement of Scientific Literacy
(jondmiller@umich.edu)**

For most of the last 50 years, political scientists and public attitude analysts have accepted Converse's (1964) conclusion that American voters are non-ideological in their political thinking. In the last decade, an increasing number of political scientists have found a growing political polarization in the United States, characterized by a strong alliance of religious and social conservatives on the right and a resurgent left-center coalition that was successful in 2006 and 2008 (but not 2010).

The purpose of this paper is to examine various measures of political ideology that have been used in these analyses. Specifically, the paper will examine several self-reported categorizations (the seven-point scale favored by the American National Election Study, the three-point typology used by the CBS-New York Times surveys, the five-point scales used by the ABC-Washington Post surveys, and zero-to-10 scales used by Miller and others). The analysis will examine the substantive portrait provided by each kind of measure and the relationship of each of these measures to partisanship, participation, and selected demographic and attitudinal variables.

The paper will contrast these self-reported measures of ideology with a constructed scale of political ideology. The constructed scale uses a combination of reports of the level of issue salience and the direction and strength of attitudes on a set of national political issues. The resulting scale reflects a distribution of political ideology weighted by issue salience.

Using national survey data from the 2004 and 2008 presidential elections, the analysis concludes that there is substantial substantive agreement between self-reported and constructed measures of political ideology, but there are some important differences. The primary differences reflect the relative size of the center or non-ideological group and the distribution of strongly held ideological positions.

An Injured Party?: A Comparison of Political Party Response Formats in Party Identification.

Samantha L. Schwartz, ICF International (sschwartz@icfi.com); Frances M. Barlas, ICF

International (fbarlas@icfi.com); Randall K. Thomas, ICF International (rthomas@icfi.com); Regina A. Corso, Harris Interactive (regcorso@aol.com); Ronald Z. Szoc, ICF International (rszoc@icfi.com)

Burden and Klofstad (2005, Political Psychology) compared political party identification (PID) using “feel that you are” in the item stem with “think of yourself as” in the item stem and found significant differences in the proportions of people identifying themselves as Republicans. Neely (2007, Political Psychology) partially replicated this effect. We sought to explore this effect in more detail in a web-based survey experiment by developing a number of different item stem variants (including ‘think’ versus ‘feel’ along with ‘identify’ and ‘attached to’). One set of respondents were randomly assigned to a branching option (with a follow-up item to determine strength of identification), while another set of respondents were assigned a single response question (with 7 graded response categories). We compared categorical identification and dimensional measurement (with the categorical then strength branching format converted into a graded scale) with a number of political opinion measures, including presidential evaluation, political ideology, along with approval of the president’s military, economic, and foreign policy decisions, and attitudes toward a number of social and political issues (including abortion, immigration, etc.). We found substantial differences in PID resulting from nature of response format (branching versus single response) rather than the nature of the item stem presented – a single response scale took less time to complete and led to more extreme responses than a branching format (raising identification as Democrat and Republican by as much as 5%). In addition, single response PID showed a higher correlation with political opinion items. We discuss the theoretical and applied implications for using political party ID for analysis and weighting and how differences in mode of interview could influence results.

Revisiting Open-ended Questions.

Floyd Jackson Fowler, Center for Survey Research, UMass Boston (floyd.fowler@umb.edu); Carol Cosenza, Center for Survey Research (carol.cosenza@umb.edu)

In the early days of survey research, open-ended questions, questions that respondents answered in their own words, were dominant. Researchers felt they were the only way to find out what respondents really had to say, and such questions also were thought to make interviews more engaging for respondents because they could feel that their voices had been heard.

Over the years, open-ended questions have almost disappeared. The appeal of avoiding coding, the ease of having interviewers and respondents enter data directly, and the neat, orderly character of responses to fixed-choice questions all drove out open-ended questions. However, in the course of trying to develop good measures of what people know about certain medical decisions, we had occasion to revisit them.

The reason we started playing with open-ended questions was we had some concerns about fixed-response questions as measures of knowledge. One problem with fixed-response fact-based questions is that their difficulty depends as much on how easy it is to rule out the wrong answers as on knowledge of the right answer. In addition, if there are four answers offered, a person could choose the correct answer 1 time in 4 by chance, making it harder to get a good reading on how many respondents actually know the correct answer. Asking questions in open-ended form potentially obviates both of those problems. So, we conducted a set of 8 split-ballot studies, comparing open and fixed-response versions of questions designed to measure the same bit of knowledge. We compared the results with respect to missing data and estimates of whether or not the respondents got the “correct” answer. This proposed presentation will report the interesting results of those experiments.

Surveying Ethnic Populations.

Testing a Privacy-Protected Method with Language Minorities for Collecting Data on Immigration Status.

Young Chun, NORC at the University of Chicago (YChun2@gmail.com); Fritz Scheuren, NORC at the University of Chicago (Chun-Young@norc.org)

Neither decennial census nor government surveys ask direct questions about immigration status because such sensitive questions are deemed to be self-incriminating for some respondents. By offering absolute anonymity protection, the Grouped Answers Method (GAM) we test in non-government-conducted data collection is designed to (1) increase foreign-born respondents' willingness to participate in in-person interviews that ask about immigration status and (2) minimize measurement errors arising from perceived threat. Unlike the indirect methods currently used, the GAM method directly estimates the foreign-born in all legal immigration statuses; the only indirect estimate is that of the number of unauthorized residents. The GAM method offers the possibility to directly obtain reliable, valid data on legal immigration statuses, which have been sought for years. In a project sponsored by the National Science Foundation, we administer GAM in cognitive tests with two of six major language groups, covering 98 percent of the American population. The GAM has built on and improved previous anonymous survey techniques – such as randomized response, aggregated response, and the item-count method. We report findings that appraise the usability and limitations of the GAM by investigating culturally sensitive aspects of a survey of the foreign-born, including recent arrivals, unauthorized residents, recent immigrants and those with limited English proficiency. We discuss cognitive tests devoted to Hispanic and Asian populations.

Cleaving the Past Behind: A Comparison of Response Formats in the Measurement of Ethnic and Racial Origins.

Frances M. Barlas, ICF International (fbarlas@icfi.com); Randall K. Thomas, ICF International (randall.k.thomas@gmail.com); Wm. Bryan Higgins, ICF International (bhiggins@icfi.com)

Ethnicity is difficult to measure in complex, multi-ethnic societies. Classification distortions arise when using mutually exclusive but overlapping categories due to mixed ancestries and multiple social identities. In a web-based survey, we compared categorical classification with different measures of ethnicity by ancestral origins. Respondents first completed a traditional categorical measure of race/ethnicity (with a separate Hispanic item) and then asked about their and their parents' birth state/country. We randomly assigned one of 3 response format with the first asking about the birthplace of their grandparents. Two other response formats asked "Thinking about your own ancestry, about what percentage of your ancestors came from the following groups or locations?" and presented 7 alternatives (Spanish-speaking or Portuguese-speaking nation, Africa, Asia/South Pacific Island, Europe, Native American/Eskimo, Middle East, South Asia, plus 'another place'). One response format had fixed categories (None, Less than 5%, 5% to 24%, 25% to 49%, 50% to 74%, 75% to 94%, 95% or more) while the other format used a constant sum task (0-100) to distribute ancestry proportions. While all three measures were correlated with cultural identification, the relationships between ethnic and racial identification, ancestry, and cultural identification were more complex than first expected.

Behavioral and Attitudinal Differences between Hispanics in the U.S. from Whom Data are Gathered in Spanish and Hispanics in the U.S. from Whom Data are Gathered in English.

Paul J Lavrakas, Independent Consultant (pjlavrakas@hughes.net); Trevor N Tompson, The Associated Press (TTompson@ap.org)

A national survey of 1,521 Hispanics living in the United States was conducted by NORC University of Chicago for a consortium of researchers from The Associated Press, Univision, The Nielsen Company,

The Bill and Melinda Gates Foundation, and Stanford University. Data were gathered from one head of house from a nationally representative sample of Hispanic households in the United States from March 11, 2010, through June 2, 2010. The designated sample of Hispanic households was provided by The Nielsen Company and taken from their national enumeration survey of Hispanic households in the U.S. Data were gathered via mail, the Internet, or the telephone. The majority of questionnaires were completed via mail (82%); 15% were completed via telephone, and 4% via the Internet. The AAPOR RR3 was 44%. In terms of language, 53% of the completions were via the English questionnaire and 47% were via the Spanish questionnaire. In terms of nativity, 37% of the respondents were born in the United States, 36% were born in Mexico, and the remaining 27% were born in another country. The survey questionnaire gathered data about Hispanics' experiences with and aspirations for their own educational attainment and that of their children, their ethnic-identity, matters pertaining to household financial well-being and health, media usage, political attitudes, and a variety of demographic and other background characteristics. Our paper will report findings about the group of Hispanics who provided data in Spanish versus those who provided it in English related to (a) their behavioral, attitudinal, and psychographic characteristics and (b) the effort it took during the field period to gain a completion from each group of Hispanic respondents. Linked to the implications of our findings, we address the serious mistake that often is made when surveys weight the Hispanic respondents in their samples.

Framing a Path to Citizenship: The Implications of Question Wording on Responses by Ethnicity.
Marisa A. Enriquez, University of Arizona (enriquem@email.arizona.edu); Kate Kenski, Dept. of Communication, University of Arizona (kkenski@email.arizona.edu)

Research on question framing has indicated that the way in which questionnaire items are phrased can lead respondents to select some answers over others (Goetz, 2008; Schuman & Presser, 1977; Wahl, Svensson, & Hyden, 2010). Framing can seemingly affect people's expressions of attitudes on surveys. Previous research has also indicated that Hispanics hold different attitudes than non-Hispanics on several public policy positions (Kenski & Tisinger, 2006). The current study examines the extent to which question wording about immigration affects opinions and the extent to which those potential effects vary by ethnicity. Using data from the 2008 National Annenberg Election Survey (n=6,073), we investigated how word choice describing those who were not in the country legally affected people's opinions toward providing a "path to citizenship" for those who agreed to return to their home country for a period of time and pay substantial fines. Respondents to the survey were randomly assigned to one of three framing conditions: "illegal aliens," "illegal immigrants," and "undocumented immigrants." Contrary to expectations, the framing of the questions about providing a path to citizenship did not produce significantly different results when the overall sample was examined. Differences appeared, however, when the data were disaggregated by ethnicity. Specifically, Hispanic respondents were significantly more affected by question wording than were non-Hispanics. The implications of these findings are discussed.

The Story behind the 2010 Mid-Term Election.

The Voters' Voice in 2010.

Gary Langer, Langer Research Associates (glanger@langerresearch.com); Julie E. Phelan, Langer Research Associates (jphelan@langerresearch.com)

The 2010 midterm elections produced the largest shift in seats in the U.S. House of Representatives in postwar history – a dramatic rebuke to the in-power Democratic Party and a stunning change from voter preferences just two years earlier, when Barack Obama led the Democrats to a sweeping victory.

What changed?

In this paper we examine economic and attitudinal data leading up to the 2010 midterms, and exit poll data from the election itself, to draw a portrait of the electorate and its motivations. We add data from previous pre-election and exit polls, in 2008 and earlier, to place the 2010 election in historical context.

Elements we examine include the influence of economic conditions on public attitudes, comparative support for the Tea Party political movement and the two major political parties, changes in voter turnout and in voter preferences among key demographic and attitudinal groups, and the issue preferences and priorities of 2010 voters. We also discuss differences between voters and nonvoters in 2010, with an eye toward the presidential election of 2012.

A Pox on all Their Houses: The American Electorate in the 2010 Election.

**Allan L McCutcheon, Univ. of Nebraska-Lincoln, Survey Research and Methodology
(amccutcheon1@unl.edu)**

The results of the 2010 National Election Pool (NEP) exit polls indicate that the national electorate held a favorable view of neither political party, and the 'Tea Party' movement appeared to fair no better, with only 41% of the exit poll respondents expressing a positive view of the movement. Unlike nearly all analyses that are currently available, this paper examines these divisions among the electorate using a multivariate analysis of the 2010 NEP exit poll data. Specifically, the paper focuses on a simultaneous examination of the differential impact of demographic characteristics and issue preferences among the exit poll respondents that appear to have lead to a highly fractured electorate.

Preliminary analyses indicate that respondents' personal economic issues and concerns played a more important role in their evaluations of the parties and movement, than did the their attitudes toward other national issues (e.g., healthcare, tax cuts, same sex marriage). Following an analysis using a multinomial logistic regression analysis of the national data sample, a multi-level model analysis of voter characteristics and issue preferences will be used with the state-level samples of votes to examine possible structural (i.e., state-level) factors that may influence the differences of support for the two parties and the Tea Party movement (see e.g., Gelman 2008).

Gelman, A. (2008) *Red State, Blue State, Rich State, Poor State: Why Americans Vote the Way They Do*. Princeton, NJ: Princeton University Press.

Who Are Those Tea Partiers And Who Supports Them?

Thomas Lamatsch, Monmouth University Polling Institute (thomas.lamatsch@monmouth.edu)

Both Republican and Democratic voters tend to be more moderate than party activists for these established parties. How does this work for the tea party movement? Political commentators from all sides still have a hard time describing the tea party activists. Descriptions range from "Libertarians" to "Social Conservatives" to "A protest movement", or "Bunch of angry people from the square states". It is even harder to describe their general support base, those people favoring and potentially voting for tea party candidates but not engaged to the level of attending public rallies.

That leads to the question: "Who are the tea party supporters in the general electorate how does a voter that intends to vote for a tea party candidate in November look?"

The paper will strive to answer three questions:

- Is there an average tea party supporter, if yes how can he or she be described?

- Are positive views of the tea party translating into voting for its candidates?
- As a group, are tea party supporters more or less cohesive than supporters of the Democrats or the Republicans?

The paper will use original data from pre-election surveys in several congressional districts in New Jersey, Delaware, and Pennsylvania conducted by the Monmouth University Polling Institute. The data includes districts where tea party candidates have defeated the establishment Republican candidate in the primary (such as NJ 06 or Delaware) as well as districts with candidates who are "tea party endorsed".

The Character of the Tea Party Movement.

Emily Ekins emilyekins@ucla.edu Ekins, University of California, Los Angeles (emilyekins@ucla.edu)

The recently emergent Tea Party movement has demonstrated a unique character, with its participants espousing views that have been largely absent from recent political discourse. Instead of debating contemporary political issues, such as taxes, gay marriage, and abortion, Tea Party activists have reintroduced past debates about individuals' locus of control, the role of government intervention in the economy, and the origins of American prosperity. Despite the limited government rhetoric of most Tea Partiers, many favor entitlement programs such as Medicare and Social Security and oppose outsourcing, liberalizing immigration, and the construction of the Islamic Cultural Center in New York City. In this paper, I seek to explain how these seemingly contradictory beliefs coexist in the movement. To do this, I use a combination of surveys including an analysis of Tea Party rally signs, a 700-person survey at the Virginia Tea party Convention, a collection of Internet surveys administered with the University of Virginia's YourMorals.org project, corroborated with media polls' raw data to glean the true character of the Tea Party. I find there are two primary groups—one leaning libertarian and the other conservative. These groups are statistically distinct from one another on social and cultural issues and similar on economic issues and abstract notions of liberty and socialism. Ultimately my analyses suggest that some of the Tea Party has origins in libertarian philosophical views responding to expanding governmental scope following the financial crisis, but other parts are driven by the perception of cultural threat.

The Role of Social Networks Over Time in Party Identification and Political Participation.

Darby Steiger, The Gallup Organization (darby_miller_steiger@gallup.com); Bob Torongo, Gallup (bob_torongo@gallup.com); Kiley Nemeckay, Gallup (kiley_nemeckay@gallup.com); Samuel Arbesman, Harvard University, Department of Health Care Policy (arbesman@gmail.com); Alistair James O'Malley, Harvard University, Department of Health Care Policy (omalley@hcp.med.harvard.edu)

Studies of social networks have been on the rise in over the past 2 decades to examine linkages between one's social circle and health outcomes, political participation, civic engagement and many other outcomes. Of particular interest is role that social networks can have on party identification and propensity to vote. Specifically, as a person's social network changes over time, does the changing nature of their friends' political persuasions impact their own political leanings?

This longitudinal study, in conjunction with the Christakis Lab at Harvard University, aims to study how social networks change over time and how this impacts a variety of behavioral and social outcomes. Using the national Gallup Panel, the study used dual "content neutral" name generators to produce the social network. Between April 2009 and June 2010, panelists were surveyed three times to define their network and report on outcome variables such as life satisfaction, political participation, health outcomes, and civic engagement.

Examining all waves of data, this paper examines the role of network size, intensity, geographic proximity, and homogeneity on voting behavior and political preferences. The paper also explores how shifts in the party identification of members of one's social network may influence changes in the ego's political proclivities and propensity to vote. Finally, the paper also examines the impact of the network on the panelist's actual voting behavior in the 2010 midterm elections.

Results of this research will further inform how voters are influenced and how opinion researchers can utilize social network analyses to add nuance to election forecasting.

Saturday, May 14, 2011

1:15 p.m. - 2:15 p.m.

Poster Session 3.

Response Consistency to Similar Questions asked in one Study Instrument Administered in a Five Country Trial.

Annette M Green, RTI International (agreen@rti.org); Tyler D Hartwell, RTI International (thartwell1@nc.rr.com); Willo Pequegnat, NIMH (wpequegn@mail.nih.gov); Godfrey B Woelk, RTI International (gwoelk@rti.org)

This study focused on prevalence of common bacterial and viral STDs and risk behaviors among individuals at high risk for HIV in five low- and middle-income countries (China, India, Peru, Russia, and Zimbabwe). Consistency of responses to very similar questions asked in different sections of the study instrument is the focus of this paper. These questions dealt with conversations about safe sex and alcohol use.

Participants were asked early in the questionnaire if they had spoken to anyone in the past three months about AIDS, other sexually transmitted diseases, or condoms. In a latter section of the questionnaire the participants were asked if they had spoken to, up to five partners, about the same subjects in the last three months. We found 15% of those participants who said they had not talked to anyone in the last three months about the subjects of AIDs, STDs, or condoms did admit to talking to their partner(s) about these subjects. Forty-three percent of those who said they did talk to someone about the subjects did not confirm their partner(s) were among those they had spoken to concerning these issues.

Study participants were asked if they used alcohol before sex with each partner. This was limited to the last 10 sexual encounters. At the end of the questionnaire subjects were again asked about alcohol use. Here we found 36% said they had not used alcohol during the last 10 encounters with a partner but admitted to alcohol use in the last 30 days. Of those who admitted alcohol use in the past 30 days, 70% admitted alcohol use prior to sex.

In this paper, demographic characteristics and prevalence of STDs are some of the factors used to research similarities and differences in responses to these 'twice-asked' study questions both overall and by country.

Category Selection Probing in Online Access Panels.

Dorothee Behr, GESIS - Leibniz Institute for the Social Sciences (dorothee.behr@gesis.org); Lars Kaczmirek, GESIS - Leibniz Institute for the Social Sciences (lars.kaczmirek@gesis.org); Michael Braun, GESIS - Leibniz Institute for the Social Sciences (michael.braun@gesis.org); Wolfgang Bandilla, GESIS - Leibniz Institute for the Social Sciences (wolfgang.bandilla@gesis.org)

Cognitive interviewing is a well-established method for evaluating and improving a questionnaire prior to fielding. However, its present implementation brings with it some shortcomings: small sample sizes, considerable time and cost investment, and the possibility of interviewer effects. In this study, we test non-probability online access panels as a means to conduct cognitive interviewing, in particular category selection probing, in order to avoid the above mentioned shortcomings. The study reports on the results

of 1,007 respondents from Germany. In order to identify implementation features that lead to a high number of meaningful answers, we explore the effects of (1) different panels, (2) different probing variants, and (3) different numbers of preceding probes on the answer quality. The overall results suggest that category selection probing can indeed be used in web surveys. Using data from two access panels – a community panel where members can actively get involved, e.g. by creating their own polls, and a ‘conventional’ panel where answering surveys is the members’ only activity – we find that high community involvement does not increase the likelihood to answer probes or produce longer statements. Testing three probing variants that differ in wording and provided context, we find that presenting the context of the probe (i.e., the probed item and the respondent’s answer) produces a higher number of meaningful answers. Finally, the likelihood to answer a probe decreases with the number of preceding probes. However, the word count of those who eventually do answer the probes does not decline with an increasing number of preceding probes.

Using VoIP and a Local Cellular Phone to Contact Hard-to-Reach Respondents for the National Birth Defects Prevention Study.

Rebecca G Devlin, Abt SRBI (r.devlin@srbi.com); Jennita Reefhuis, Centers for Disease Control and Prevention (nzc5@cdc.gov); Sarah Tinker, Centers for Disease Control and Prevention (zzu9@cdc.gov)

Technological advances in telephony, including cellular phones and caller ID, have created numerous challenges in conducting telephone surveys, not the least of which is the ability of respondents to screen out unknown numbers. However, these same advances may offer ways to overcome barriers to contact. The proposed research explores the effect of two such technologies: Voice over Internet Protocol (VoIP) and cellular technology to allow the display of a local area code on call screening devices, on the contact rates in the National Birth Defects Prevention Study (NBDPS).

Since 2009, Abt Associates and Abt SRBI have conducted telephone interviews for the Atlanta site of the NBDPS. Starting in 1997, and sponsored by the Centers for Disease Control and Prevention (CDC), the NBDPS is the largest population-based study investigating the causes of birth defects. Interviews include questions about the respondent’s pregnancy and the three months before she became pregnant, pregnancy history, diet and health and medication use. Interviews are conducted from an Abt SRBI telephone center in Hadley, Massachusetts. Cases are obtained from a birth defects surveillance system; controls (mothers of infants without birth defects) are randomly selected from vital records.

Because most respondents were located in the Atlanta, GA metropolitan area, voice over Internet protocol (VoIP) was obtained in the Massachusetts telephone center so that a local Atlanta area code would appear on caller ID when making study calls. Additionally, a cell phone with a local Atlanta area code was purchased for calling hard-to-reach respondents. The proposed research explores whether using VoIP and the cellular phone increased the contact rate. Results may inform other telephone surveys that have respondents in a specific region that are dialed from a remote location. Discussion of applications to other telephone surveys will include using an experimental design.

An Evaluation Of The Collection And Production Of Medical Condition Estimates From The Medical Expenditure Panel Survey.

Frances M. Chevarley, AHRQ/DHHS (fran.chevarley@ahrq.hhs.gov)

The prevalence and costs of medical conditions are important to public health. This paper analyzes condition estimates from the Medical Expenditure Panel Survey (MEPS), a nationally representative panel survey studying health care use, access, expenditures, source of payment, insurance coverage,

and quality of care. Each year a new panel begins and each panel has 5 rounds of data collection over 2 ½ years that covers a two-year period.

There are alternative ways to produce condition estimates based on MEPS data. The first is based on responses to condition-specific questions in the household CAPI instrument. Condition estimates associated with health events such as doctor visits (treated prevalence) can also be made. In addition, estimates based on conditions that bothered the person that occurred between the rounds or since the last interview can be made. The respondent along with other family members who may be present during the interview report the information.

Starting with Panel 12 (2007 Panel) in an attempt to reduce respondent burden and to improve the reporting of conditions, the enumeration of priority conditions was moved to the first round of data collection and earlier in the CAPI instrument. The questions are then asked in subsequent rounds in certain situations. Several other changes were made at this time in the data collection procedures/tabulations/editing. In addition, there was a change in the sample design in 2007. This paper focuses on the effects of changes to the priority condition-specific questions of the MEPS CAPI instrument.

Evaluating the Meaning of Vague Quantifier Terms in Questionnaires.

John Tarnai, Social & Economic Sciences Research Center (tarnai@wsu.edu); Danna Lee Moore, Social & Economic Sciences Research Center (moored@wsu.edu); Marion Schultz, Social & Economic Sciences Research Center (schultzm@wsu.edu)

Many questionnaires use vague response terms, such as “most,” “some,” “a few” and survey results are analyzed as if these terms have the same meaning for most people. This assumption has rarely been evaluated to assess its accuracy. In an attempt to evaluate the extent of the meaning these terms have for people, we investigated this by first asking household respondents to answer several questions that used these vague quantifiers. These questions were then followed by companion questions that asked for a numerical percentage comparable to the vague quantifier questions. The questions were included in a statewide telephone survey about economic conditions in the state. A total of 500 households participated in the study conducted in 2009. The results demonstrated several things. First, that there is considerable variation in the meaning of “most,” “some,” “a few” at least in terms of the range of percents which apply to each term. Second, when people are asked to give a percent response they don’t distribute their responses evenly among the full range of percents possible. Instead, people tend to provide answers that are grouped into categories of 10%, 20%, 30%, etc. We present the results of this study along with recommendations for improving questions with vague quantifiers. The paper discusses the implications of these results for designing survey questions that use vague quantifier terms.

An Acoustic Analysis Of Respondents’ Expressed Conviction: Testing A Refinement Of Telephone Surveys In Predicting Electoral Results.

Anne-Marie Lesperance, University of Montreal (anne-marie.lesperance@umontreal.ca); Victor J. Boucher, University of Montreal (victor.boucher@umontreal.ca)

Surveys by phone involve the recording of subjects’ spoken responses, which are usually consigned by an interviewer in terms of their literal meaning (“yes”, “never”, politicians name, etc.). However, oral responses also contain acoustic information on speakers’ attitudes and affect, for instance, in expressing a degree of conviction. This information is lost and yet may serve to predict behaviors. In a pilot study we found that several measurable acoustic variables of speakers’ oral responses show categorical changes in expressing a strong conviction. Our present study applies these findings in analyzing a corpus of

recordings from a phone survey on voters' intentions. The corpus consists of approximately 300 interviews realized by Leger Marketing. We have two objectives. First, we aim to determine if acoustic aspects of responses denoting conviction on a choice of candidates maintain categorical characteristics as in our pilot study. Second, we want to determine if classification of responses taking into account acoustic variables of conviction enhances the prediction of electoral results compared to responses that are consigned strictly in terms of literal meaning. Our expectation, based on preliminary analyses, is that acoustic analyses will serve to reclassify answers and that this can reduce error margins and improve the rate in predicting electoral results. The acoustic analyses focus on two prominent candidates, where voter intentions were split. The present project will lead to define cut-offs in acoustic variables that can be automatically detected.

Tracking American Community Survey Mail Response During the 2010 Census.

Stephanie Baumgardner, U.S. Census Bureau (stephanie.k.baumgardner@census.gov); David Raglin, U.S. Census Bureau (david.a.raglin@census.gov)

Due to the increased awareness surrounding the 2010 Census activities, we hypothesized that response to the American Community Survey (ACS), a continuous survey conducted every month by the Census Bureau, would be affected during early 2010. The 2010 Census publicity and outreach effort utilized a comprehensive marketing strategy known as the 2010 Census Integrated Communications campaign. This campaign employed many techniques to raise awareness of and encourage participation in the 2010 Census including paid media, public relations, promotions and partnerships. These techniques were tailored to eight major segmentation groups of the U.S. population.

This paper describes the effects on ACS mail response during the 2010 Census collection period. We compared the 2010 ACS mail check-in rates to the 2009 rates to provide an assessment of the impact that the 2010 Census publicity may have had on public awareness and cooperation and also to provide feedback on changes in the ACS workloads due to the 2010 Census. We also examined mail check-in rates by each of the eight segmentation groups to determine if some groups benefitted more from the advertising used to promote the 2010 Census. The differences in the 2010 and 2009 rates at the national level as well for each of the segmentation groups show a clear trend during the first half of 2010.

In addition to tracking the 2010 ACS mail response, we studied ACS mail form completeness as another measure of respondent cooperation. Differences in the completeness of the 2009 and 2010 ACS mail forms were considered at the national level and for each of the eight segmentation groups. As with the ACS mail check-in rates, a common theme can be seen that indicates the willingness of the American public to fully participate in the ACS, even during the 2010 Census.

Recruiting and Hiring: The Impacts of Specialized Skill Sets.

McKinlay Jeannis, RTI International (mjeannis@rti.org); Carmen Lopez, RTI International (clopez@rti.org)

Recently the criteria for recruiting and hiring telephone data collection staff have become more specialized. Previous standards for staffing and recruiting consisted of customer service or call center experience and an education level of at least a high school diploma. Over the past several years, projects have requested a specific set of skills particular to the study to be taken into account when recruiting and staffing. These specialized skills have ranged from, helpdesk support for web surveys, phone records abstraction and data entry, targeted tracing of specific hard to contact sub-populations, and institutional contacting all of which required higher educational levels beyond a high school diploma. Analyzing data from a recent study's criteria in recruiting and hiring, we intend to present the impacts of attrition and

performance in correlation to specialized skill sets such as records abstraction and education. This data could help target those skills and requirements that will enable call centers to recruit and hire data collection staff that will yield better performance and higher retention rates.

Developing an Internet Response Mode for the American Community Survey (ACS).

Jennifer Guarino Tancreto, U.S. Census Bureau (jennifer.guarino.tancreto@census.gov); Mary Davis, U.S. Census Bureau (Mary.C.Davis@census.gov); Mary Frances Zelenak, U.S. Census Bureau (Mary.Frances.Zelenak@census.gov)

In 2000, the Census Bureau tested an Internet response option for the ACS. Researchers found that offering the Internet as a response option during mail data collection actually decreased the overall response rate, and that very few respondents completed the questionnaire on the Internet (Griffin et al., 2001). Since 2000, technological advances have been instrumental in moving towards a paperless society. For surveys, the benefits of using the Internet, such as speed of responses and automated quality checks, have inspired survey organizations to experiment with Internet data collection. The Census Bureau has recently developed an Internet response option for ACS that will undergo field testing in April 2011. The main purpose of the field test is to determine the impact of various strategies for notifying sampled addresses about the Internet response option on response rates. The Internet instrument was designed to present stimuli consistent with other data collection modes used in the ACS to minimize mode effects, while taking advantage of the technology to the extent possible to improve data quality. This poster will focus on the development of the Internet instrument, including the guiding principles behind the design as well as modifications resulting from expert review and usability testing. We will also highlight the paradata we will collect to help evaluate the design of the instrument.

Environmental Disaster? Global Warming And Subjective Wellbeing: Some Gain And Loss Estimates From Survey Research.

Jonathan Kelley, University of Nevada, Reno and University of Melbourne (kelley@international-survey.org)

Scientists and environmental groups have raised alarms about many consequences of global warming. But the impact of temperature change on the general public has not been estimated empirically. I provide preliminary estimates of the magnitude of the effect on subjective wellbeing (SWB). DATA: The climate change module in the International Survey Center's May, 2009 omnibus survey (N=2295), a representative national sample of US adults. METHODS: Regression estimates of temperature's impact on SWB, made commensurable through dollar equivalents estimated from the impact of income on SWB. QUESTIONS: "How do you feel about your life – how satisfied are you with – The weather here in the winter? The summer weather? Your life as a whole?" each answered on an 11 point "Delighted-Terrible" scale, scored from 0 to 100. RESULTS: People are about 5.5 points more satisfied with their summer weather than their winter weather. Regression analysis shows that this translates into .55 of a subjective wellbeing point. Summers are about 18° Celsius warmer than winters in the Northern Hemisphere so 1° translates into $.55 / 18 = .03$ life satisfaction points. If global warming raises temperatures by about 2° Celsius over the next century, as expected, that would increase wellbeing by .06 points. It takes about \$20,000 in income to increase life satisfaction by one point. So global warming is worth $.06 * 20000 = \$1,200$ per person per year, equivalent to a 2% or 3% increase in GNP. I also present estimates based on alternative methods incorporating locality-based temperature information. There is still much uncertainty in these estimates, and global warming may induce other changes (storms, rainfall patterns, etc.), the costs and benefits of which need to be calculated for an overall estimate of global warming's effects on subjective wellbeing in the general public.

Exploring New Data Capture Technology and Financial Software for the Consumer Expenditure Survey.

Jeanette Duddy Davis, Bureau of Labor Statistics (Davis_Jeanette@bls.gov)

The Consumer Expenditure Survey (CE) program consists of two surveys, the Quarterly Interview Survey and the Diary Survey, that provide information on the buying habits of American consumers, including data on their expenditures, income, and demographic characteristics. The survey data are collected for the Bureau of Labor Statistics by the U.S. Census Bureau. In 2009, the Consumer Expenditure (CE) Survey Program began the initial phases of a major survey redesign to improve data quality through a verifiable reduction in measurement error. One potential source of measurement error, recall bias, may be addressed through the use of mobile data capture devices and/or financial software, in place of or in supplement to the survey's current CAPI interview and paper diary. For example, financial software might help respondents remember online purchases and automatic payments that they might easily forget to report on the surveys. CE, in collaboration with Westat, examined currently available mobile data capture devices and financial software packages to identify features and applications that have potential to improve the CE data collection process. Our findings may help to reduce measurement error in the CE program and other household surveys.

Assessing Survey Accuracy across Multiple Domains.

LinChiat Chang, LinChiat Chang Consulting, LLC (contact@linchiat.com); Jon A Krosnick, Stanford University (krosnick@stanford.edu)

The prominence of surveys is based on the assumption that survey results are accurate. In contrast to the plethora of papers and books addressing survey errors resulting from coverage, sampling, nonresponse, and measurement issues, no comprehensive review exists of the evidence on survey accuracy. In writing this paper, we sought to generate a thorough review of decades of empirical research testing the accuracy of survey estimates against objective benchmarks not derived from self-reports.

Our review spans a wide spectrum of content areas, including crime and deviance, demographics, economic indicators, healthcare, labor force statistics, market research, philanthropy, politics, psychology, substance abuse, media, and much more. We show how survey measurement has been evaluated using a variety of methods - comparing each respondent's self-report (e.g. hospital stays, tax evasion, home ownership, medical screenings and diagnosis) with objective individual records of the same phenomena; correlating aggregate distributions of respondents' reports with distributions of the same phenomena derived from secondary data; comparing one-time aggregate survey estimates (e.g. height, weight, healthcare costs, absenteeism, cigarette smoking) with available benchmarks; and comparing trends over time between longitudinal aggregate survey estimates (e.g. NCVS crime rates, NES voter turnout rates) and available benchmarks. Using this literature, we also attempt to identify the conditions under which surveys provide more accurate estimates and the conditions that undermine survey accuracy.

2010 Census Confidentiality Notification Experiment.

Gianna Dusch, U.S. Census Bureau (gianna.s.dusch@census.gov)

The purpose of the Confidentiality Notification Experiment is to determine the effect of confidentiality-related messages, written in language that respondents can easily understand, on housing unit mail return rates and form completeness to the mail questionnaire. These messages are embedded in the cover letter that accompanies the initial census questionnaire mailing. The effect of two messages are of interest to this study: an administrative records use message, which has not previously appeared in

decennial production cover letters, and an alternative to the “standard” statistical purposes statement that has previously been and is currently used in production.

The intent of these messages is to convey to respondents, in clear and understandable language, certain aspects of the concept “statistical purposes.” Namely, these census data are used in aggregate form and the Census Bureau receives data from other federal agencies that it links to respondents’ individual data. The Census Bureau performs this linkage in order to carry out administrative programs that enhance census data. These messages were developed with the intent that they should not, in and of themselves, depress unit or item response. This study was developed to assess the effect of those messages.

The results from this research will be used to develop statements for potential use in the 2020 Census. Results will also inform future policy decisions at the Census Bureau where issues of communicating data protection and confidentiality to respondents are concerned. They will also provide general information on the topic of informed consent and administrative records use topics to interested parties, such as the Federal Committee on Statistical Methods’ subcommittee on administrative records.

Tracking a City’s Recovery: New Orleans Five Years after the Storm.

Sarah Cho, Kaiser Family Foundation (SarahC@kff.org); Bianca DiJulio, Kaiser Family Foundation (BiancaD@kff.org)

In 2006, one year after Hurricane Katrina pounded the city of New Orleans and water breached its overburdened levees, the Kaiser Family Foundation conducted the first of what has become a series of three major surveys of Orleans Parish residents. In 2010, the latest installment of the series was conducted during a new, unfolding economic and environmental disaster: the oil spill of unprecedented size in the Gulf waters off the Louisiana coast. At the five year anniversary of Katrina’s deadly arrival we explore the following questions: How do New Orleans’ residents feel now that the storm and flooding are half a decade behind them? How satisfied are they with the rebuilding process and what challenges and concerns animate them, both Katrina-related and otherwise? What does this new Gulf Coast environmental disaster mean to them thus far?

Using data collected in 2010, 2008 and 2006, we provide a thorough look at how residents view the rebuilding process, the extent to which they see New Orleans as having healed from Katrina, and the extent to which they expect to be impacted by the Gulf Coast oil spill. We assess residents’ feelings about: pressing local policy debates; the extent to which they are facing financial, employment and health challenges; and paint a picture of a city which remains divided by race and income, even as race relations seem to be on a positive trajectory. Finally, we also touch on the ways that the methods used for sampling and data collection changed as the situation on the ground evolved post-disaster, from an address based sample and face-to-face interviews in 2006 through mixed mode data collection in 2008 to a more traditional RDD landline and cell phone survey in 2010.

A Stimulus Package for Stubborn Sample Members: Boosting Web Survey Participation through Strategic Study Communication & Telephone Contacts.

Jeffrey Franklin, RTI International (jwf@rti.org); Nicole Tate, RTI International (ntate@rti.org); Joe Simpson, RTI International (jbs@rti.org); Nadia Paoli, RTI International (npaoli@rti.org)

Electronic technology has become ubiquitous in today’s society and interpersonal communication. We have adapted our data collection approach to meet the needs of our sample members and to take advantage of our technological capabilities: providing a web interview, communicating via HTML-formatted e-mail, incorporating text message reminders and introducing the study to sample members via

YouTube. Since a web interview option was introduced on our studies in 2003 as an alternative to a telephone interview, the proportion of respondents choosing the web option has increased from 38 percent to 81 percent of completed interviews, due in part to a customized communication approach that emphasizes to sample members the ease and convenience of completing the survey independently on the Web.

Conducting nationally-representative, large-scale survey research presents unique challenges that call for innovations in data collection methods. As data collection mode preferences have shifted, ways of reaching sample members have needed to change as well. We will include in this presentation an examination of results of varied communication types in large-scale, multi-mode studies conducted with young adults who were surveyed between 2003 and 2009. In addition, we elaborate on how our communication strategy has evolved to accommodate and encourage web completion. We will focus on how we tailor our written, electronic and telephonic messages to sample members for cross-sectional and longitudinal studies, the systems we avail to sample members to assist with web completion, and our implementation strategy for when we communicate with our sample.

The RWJF Health Care Consumer Confidence Index: Public Perceptions of Health Care and the Health Reform Debate.

Peter Graven, SHADAC (grave165@umn.edu); Sarah Gollust, University of Minnesota (sgollust@umn.edu)

Public opinion played an important role in the divisive battle over the Patient Protection and Accountable Care Act (PPACA), both driving the demand for action and responding to legislative developments. These bidirectional effects can also be seen in measures of public perception. The Robert Wood Johnson Foundation's Health Care Consumer Confidence Index (RWJF Index) was designed to measure consumer confidence in response to ongoing changes in the health care system. The RWJF index includes two components: (1) a measure of recent health care use and costs of care ("experiences") and (2) a measure of concerns about future health care costs ("expectations"). The RWJF Index was created from 9 survey items added to the University of Michigan's Surveys of Consumers, a monthly telephone survey of 500 adults in the contiguous United States. In this paper, we show that citizens' "expectations" were more significantly associated with key health reform political events than were their "experiences." Beginning with a base of 100 in April and May 2009, the future cost index decreased during the politically tumultuous August Congressional recess to 91.2, with comparably little change in the recent measure. In the following month, when the Senate bill was introduced, the future index increased to 105.0. Overall, groups with plausibly greater self-interest in the health care system (e.g. lower income, less healthy) had lower confidence while those with more generous coverage (e.g. over 65 years old, privately insured) had higher confidence and were significantly less likely to believe that health reform was an important goal. The RWJF Index is another measure to assess health reform impacts as implementation continues in 2011. We discuss the contribution of these findings to the literature on policy feedback effects, suggesting that measurements of opinions, expectations, and experiences offer insights into the impact of policymaking on public perceptions.

Media Coverage of Polling Methodology in 2010 Multi-Candidate General Elections.

Benjamin Duffey, University of Michigan (biduffey@umich.edu)

In multi-candidate, state-level elections, voters rely largely on the media to convey poll results and inform their voting strategy. This trend highlights the need for poll coverage to contain at least a basic discussion of methodological quality.

I analyze polls and poll coverage from six multi-candidate races in 2010 (AK, FL, CO, ME, MN, and RI). After establishing measures of poll accuracy and methodological quality, I model the relationship between poll accuracy and methods quality in these six races. I compare the results of the model to a content analysis of poll coverage about these races in order to gauge the media's success in conveying methodological quality. Finally, I use the results of this comparison to discuss the potential effects of the under-reporting of methodological considerations.

The Gulf Oil Spill and Support for Offshore Drilling in Coastal Louisiana.

Robert Kirby Goidel, Louisiana State University (kgoidel@lsu.edu)

Even amidst the largest natural disaster in American history, Louisiana coastal residents continued to show strong support for offshore drilling. The reason is simple: Residents perceived the spill as a technical failure that could be fixed with better engineering. In this paper, we explore attitudes toward the Gulf Oil spill among Louisiana coastal residents and how well these attitudes correlate with local news coverage of the disaster. The results are based on two surveys of coastal residents during the oil spill and a content analysis of local newspaper coverage. The results reveal that respondent perceptions were reflected in local news coverage. As a result, the spill was seen narrowly as a technical problem and the fault of a single corporation, BP, rather than a larger environmental, policy, or systemic problem.

Racial Differences in Nonresponse Patterns between Landline and Cellular Telephone Surveys.

Sunghee Lee, University of Michigan (sungheel@isr.umich.edu)

With the substantial increase in the rate of cell-phone only use in the general population, it is well known that the rate is much higher among Hispanics than non-Hispanic whites. To decrease potential under-coverage problems coming from such disproportional patterns of cell-phone only usage, surveys have begun to sample cellular phone frames. While this addresses the coverage problem, the level of Hispanics participation compared to others may differ between landline and cellular phone surveys contributing to nonresponse bias.

This study examines racial differences in nonresponse between landline and cellular phone samples, focusing on Hispanics and non-Hispanic whites. The data come from the 2009 California Health Interview Survey which sampled both landline and cellular frames. We first examine the phone usage of Hispanics in California and whether their usage resembles that reported by Hispanics in National Health Interview Survey. We then examine the race composition using sampling weights and nonresponse adjustment weights. This is an indirect measure of nonresponse bias due to differential response by race. The comparisons on the sampling- and nonresponse-weighted racial compositions will be done between the landline and cellular phone samples within race. These are expected to provide information about whether the nonresponse pattern is different between the two samples. In other words, this will allow us to examine whether, for example, Hispanics' nonresponse pattern on a landline telephone survey is the same as that on a cellular phone survey. Additionally, if differences in response by race are discovered, the same analysis will be applied to other survey variables to examine how differing nonresponse patterns affect survey estimates.

Same Family, Different Planet: Comparing the Views of Teens and Parents on Their Relationship and Potential Substance Use.

Roger Vaughan, Columbia University (roger.vaughan@columbia.edu); Emily C. Feinstein, National Center on Addiction and Substance Abuse, Columbia University (efeinstein@casacolumbia.org); Jordon Peugh, Knowledge Networks (jpeugh@knowledgegenetworks.com); Jeffery Shand-Lubbers, Knowledge Networks (jlubbers@knowledgegenetworks.com)

Research with dyads of parents and teens provides a way to compare differences and similarities in attitudes, opinions, and experiences within the same households. Capturing and comparing these two perspectives can be helpful for planning policy interventions and communications related to helping both parents and teens deal with substance use issues.

As part of a research series on teens and parents, the National Center on Addiction and Substance Abuse at Columbia University (CASA) conducted surveys of parents and teens using KnowledgePanel, Knowledge Networks' (KN) national probability-based web panel that is recruited through RDD and ABS. Data were collected on 456 parent/teen dyads.

Examining these dyads, we uncover important differences between teen and parent perceptions about their relationship and the availability to and use of drugs and alcohol by teens. We find significant within pair differences on each of eight variables asked of both teens and parents. Looking at these data in total can in some cases obscure the degree of difference between parents and teens in the same household. While 7% of parents and 11% of teens overall report that half or more of the teen's friends drink alcohol, looking within pairs 30% disagree on how many of the teen's friends drink. One of the largest differences between parents and teens was perceptions of teen's likelihood to try drugs in the future, with 61% of parent/teen pairs in disagreement. Generally, parents are more likely than teens to expect that the teen will try drugs in the future.

The full paper will present information on KN methodology, survey development, sampling methodology, and comparison to a parallel RDD survey. Presented results will identify demographic factors associated with greater or lesser differences among pairs and explore associations between parent/teen disconnect and risk factors for substance use.

Data Collection: Sample Members Responses, Influences, and Perception of Data Security.
McKinlay Jeannis, RTI International (mjeannis@rti.org); Tamara L Terry, RTI International (tterry@rti.org); Dawn Thomas-Banks, RTI International (dbanks@rti.org)

As identity theft and data security breaches occur in the United States within government and private organizations both public opinion and sample member's willingness to provide personal identifiable information are impacted. The Bureau of Justice Statistics (BJS) has reported that Identity theft within households has increased by 23% from 2005 to 2007. In the article "Identity Theft Reported by Households, 2007- Statistical Tables", the BJS reports that in 2007 alone, 7.9 million households which accounts for 6.6% of all households in the United States have been a victim to one or more types of identity theft. Public opinion is commonly formed by personal experience or events that are reported in the media. One of the most headlined security breaches that occurred recently in 2006 was a stolen U.S. Department of Veterans Affairs (VA) laptop that contained the names, social security numbers, and date of births of 26 million plus veterans. With identity theft rates increasing and events of security breaches being reported in the media, we perceive that research study participants are conscious and skeptical on providing personal identifying information.

In our presentation we intend to demonstrate how identity theft and data security breaches in the United States may have influenced sample member's willingness to provide personal identifiable information on a cross-sectional project that collected data in 2004 and 2008. We will analyze sample members' responses to questions during web and telephone data collection modes that attempt to obtain personal information such as social security number, age, sex, income, employment status and race.

Utilizing Facebook Application for Disaster Relief: Social Network Analysis of American Red Cross Cause Joiners.

Jennie W. Lai, The Nielsen Company (Jennie.Lai@Nielsen.com); John Skvoretz, University of South Florida (jskvoretz@usf.edu)

With the exponential growth of Facebook users worldwide, survey researchers have strong interests to understand how these users are leveraging social media to connect individuals and share information with each other. This study explores the phenomenal trend of utilizing a Facebook application called Causes to help users organize into online communities for a specific cause and mobilize their resources for disaster relief during the Haiti earthquake disaster. Two separate samples of 100 joiners each from the American Red Cross (ARC) Cause on Facebook were randomly selected before and after the Haiti earthquake disaster to examine the differences of the composition (i.e., attributes) and structure (i.e., relational ties) of each social network. The social network analysis performed for this research study intends to fill the gap of historical research literature on recruitment to activism and support provision following a disaster in the digital age of the 21st century. The results of this study show how understanding the membership size of online communities, salient identity for the cause through organizational affiliations, interpersonal ties among the joiners, density of the network as well as gender diversity can be crucial recruitment factors to leverage for disaster relief efforts. The Internet has both expanded the geographic boundaries of disaster communities and helped expedite the communication between its users to mobilize the necessary resources for supporting the victims of a disaster. The findings unveil implications for how survey researchers should further explore the beneficial partnership between disaster relief organizations and online social networks in mobilizing their resources for a speedy response to disasters.

Redesigned Meter's Effect on Improving Recruitment and Compliance.

Arianne Buckley, Arbitron Inc. (arianne.buckley@arbitron.com); Darin Harm, Arbitron (darin.harm@arbitron.com)

Arbitron has developed an electronic meter that can automatically detect audio exposure to encoded radio signals. We ask panelists to wear their meter everyday from rise to retire in order to measure media exposure. Arbitron conducts ongoing research to improve the sample performance of our panels.

Former panelists have expressed the desire for a smaller, sleeker meter. Arbitron has responded to these comments by improving the design of the meter, while also improving its functionality. Among the many changes made, the redesigned meter has a more modern look and is less bulky to carry. Further, the redesigned meter transmits panelists' media data back to Arbitron via an internal cellular system, which eliminates the need for panelists to connect equipment to a telephone line to transmit data.

Prior to releasing the redesigned meter into our current ratings panel, Arbitron tested the impact of the new meter using separate sample in two markets. The purpose of the tests was to ensure that the redesigned meter performed properly and to determine the impact on panelist recruitment and compliance.

For this presentation, overall sample performance comparisons—including recruitment and compliance rates—between test panelists using the redesigned meter and current panelists using the current meter will be presented. The analysis will first look to see if the new meter design and functionality negatively or positively affects recruitment rates. Next, the analysis will examine compliance among the test panelists. After joining the panel, will the improved look and functionality of the meter improve compliance?

Ideology and Empathic Profiles.

Zeljka Buturovic, Zogby International (zeljka@zogby.com); Leann Atkinson, Zogby International (leann@zogby.com); Joe Mazloom, Zogby International (joe@zogby.com)

The maintenance of large and diverse ideological coalitions is necessary for obtaining political power in a democracy, but presents a problem because significant number of people has limited interest and information about policies and voting. We propose that one of the ways in which ideological groups are developed and maintained are through shared emphatic profiles. According to this view, people who belong to different ideological groups have a different pattern of sympathy towards various groups, such as foreigners, animals, criminals or drug addicts. These general affinities then work as a foundation on which more sophisticated policy views, such as attitudes toward animals rights or capital punishment are later built. The result of our survey of 2067 likely voters gives preliminary support for our view: we found that, when sympathy is measured on a ten-point scale, liberals and progressives appear more sympathetic toward animals and foreigners than are conservatives and libertarians. Conversely, though not to the same extent, conservatives are more sympathetic toward American soldiers and babies than are progressives and liberals. Liberals were also significantly more sympathetic than conservatives and libertarians toward criminals, drug addicts, and the homeless. The relatively stronger affinity of liberals toward animals is corroborated by the result of another survey of 3,766 adults, where we found a correlation between one's belief that "humans are animals" and his ideology measured on a nine-point scale to be about 0.40. Generally speaking, the variability in sympathy is greater among conservatives than among liberals, which could explain a variety of policy differences among the groups, from attitudes towards international law to the popularity of idea of the 'expanding moral circle' among liberal intellectuals.

Accounting for Nonresponse, Interviewer Effects and Subpopulation Membership Under Responsive Designs.

John L. Eltinge, Bureau of Labor Statistics (Eltinge.John@bls.gov)

Groves and Heeringa (2006) and subsequent publications have led a number of survey organizations to consider the use of responsive designs for allocation of data-collection resources. Practical applications of these methods to specific surveys involve a wide range of issues, including targeting of subpopulations in multipurpose surveys; accounting for partial interviewer compliance with recommended interventions; and adjustment of weights and variance estimators to account for nonresponse, interviewer and other effects. This paper outlines several possible approaches to these issues, and considers strengths and weaknesses of each. The primary approaches differ in the ways in which they: (1) condition on certain outcomes, e.g., the number of respondents in specified subpopulation; (2) use models to describe unit nonresponse and interviewer effects; and (3) use randomization approaches to characterize the impact of design features like differential assignment of incentives or other nonrespondent-conversion efforts. Some previously developed "representativeness indicators" are interpreted within the framework defined by (1)-(3). Potential applications to BLS household and establishment surveys illustrate the main ideas considered in the paper.

Key words: Conditional inference, incomplete data, Interviewer compliance, Misclassification effects, Model-based approach, Randomization-based approach, Representativeness indicator, Resource allocation

Differences In Primary Care Physicians' And Psychiatrists' Treatment Preferences For Difficult-To-Treat Disorders: Results Of A National Survey.

Kenneth A. Rasinski, The University of Chicago, Department of Medicine (krasinsk@uchicago.edu); John D. Yoon, University of Chicago, Department of Medicine (jyoon1@medicine.bsd.uchicago.edu); Farr A. Curlin, University of Chicago, Department of Medicine (fcurlin@medicine.bsd.uchicago.edu)

There is a current debate about the role and effectiveness of social relationships and spirituality in medicine. In addition, certain disorders are known to be difficult to treat and some have even been labeled “medically unexplained symptoms”. Nonetheless the disorders are troublesome to patients and are treated by both Primary Care Physicians (PCPs) and Psychiatrists. We conducted a national mail survey of PCPs (n=896 rr=63%) to ascertain attitudes about the role of patient choice in seven difficult to treat conditions (alcoholism obesity, nicotine dependence, depression, anxiety, chronic back pain and fibromyalgia) and to assess attitudes about relational/spiritual and medical treatment preferences. A small national mail survey of psychiatrists was conducted concurrently (n=312, rr=64%).

For each condition a significantly greater percentage of PCPs attributed patient choices as contributing to the disorder than did psychiatrists. The relationships between attributions of patient choice on preferred treatment type – greater attention to relationships and spirituality and taking the appropriate medication, for depression, anxiety, and fibromyalgia/chronic fatigue syndrome were examined. A scale ($\alpha=.84$) was created by summing over responses to the 7 patient choice questions. Multiple regression analyses indicated that when compared to PCPs, psychiatrists believed patients with these disorders would benefit more from paying attention to relationships and spiritual life. For PCPs there was a significant positive relationship between seeing patient choices as responsible for the disorder and believing patients would benefit from attention to relationships and spiritual life. For Psychiatrists this relationship was significantly smaller.

For both depression and anxiety beliefs that patient choices were responsible for the disorder were associated with beliefs that medicine would not benefit the patient. For fibromyalgia/chronic fatigue syndrome no relationship between patient choices and the benefits of medication was found, but psychiatrists indicated that medication would be beneficial to a greater degree than did PCPs.

How Do We Reduce The Impact Of Respondent Fatigue In A Household Expenditure Diary?
Jenny Lynch, Statistics Canada (joncaslynch@videotron.ca)

Respondent fatigue in a diary survey is a problem previously documented and observed in most diary expenditure surveys. Given the large burden imposed on the respondent, a decrease in daily reported expenses is generally observed during the recording period.

The Canadian Survey of Household Spending was recently redesigned and each household now completes a one hour interview followed by a two-week diary. In order to reduce respondent burden and fatigue, respondents were allowed to provide receipts instead of transcribing all information in the diary. Due to budget considerations, the follow-up during the recording period was done by means of a telephone contact instead of a personal visit as is used in most expenditure surveys.

The poster will show the impact on respondent fatigue of some new procedures implemented during the redesign. We will mainly focus on the receipts, the telephone follow-up and the point at which this follow-up contact should be made during the recording period. In addition, the gain obtained from a follow-up on forgotten items at the time of diary pick up will be presented.

Demographic Assessment of Mixed-mode Sample Frames and Panel Recruitment Methods.
Ryan McKinney, Arbitron (ryan.mckinney@arbitron.com)

The continued decline of landline telephone use in the United States has caused survey researchers to explore the use of additional sampling frames in order to maintain a statistically valid sample of respondents. To address the issue of landline undercoverage, Arbitron began supplementing their sample frame with an address based cell phone frame which is constructed by the use of an address frame screener questionnaire. More recently, Arbitron has gone beyond using only a cell phone supplement to also including unlisted landline records and in person recruitment for households that do not respond to an address based questionnaire.

In total, the two sample frames result in four distinct recruitment methods providing robust sample area coverage. Most importantly, the expanded use of the address based sample frame to supplement the traditional landline frame has resulted in a higher level of demographic representation to match the sampled area's universe estimates. This presentation will discuss the advantages offered by combining the two frames and recruitment methods. The analysis will cover assessments of respondent demographics that can be reached by the different modes as well as the success of each in reaching traditionally difficult age/race groups.

Unauthorized Immigration and Language Policy: Ohio Voter Opinions.
Vincent J Palozzi, Miami University OH (palozzvj@muohio.edu)

Various concerns related to unauthorized immigration such as birthright citizenship and immigrants' access to public schooling, along with concomitant language policy issues including workplace English and the language rights of criminal suspects, are among the most contentious challenges facing the U.S. today. With historic elections in both 2008 and 2010 that signaled the American public's desire for solutions, the opinions of U.S. voters on these issues should be a priority for our government representatives and lawmakers. However, little is known about the attitudes of registered voters toward many of these immigration and language policy concerns. Compared to other public policy topics, few surveys have been conducted, and most of those have been solely quantitative in nature and focusing on a limited set of issues. Likewise, the scant number of related qualitative studies have generally utilized focus groups for their data. In contrast, this paper reports on a mixed-methodology study of over 400 randomly selected Ohio voters that investigates their opinions on ten language and ten unauthorized

immigration policy issues, the underlying concerns related to their opinions, the socio-demographic variables that may predict them, and the degree of relationship between the two policy areas in the minds of voters. Based on the author's prior work, the analysis includes quantitative measures of correlation and ANOVA, in addition to a discourse analysis of respondents' remarks to confirm, clarify, and expand upon the quantitative results in order to provide a more unified picture of voters' attitudes toward these policy areas. Furthermore, the author seeks to confirm the role that gender, ideology, political party, and presidential choice play in voters' attitudes, as well as the role that economic considerations, cultural values, and a sense of fairness play in their opinions.

Identifying Common Verbatim Errors through use of Field Observations.

Christine Clark, RTI International (chclark@rti.org)

Standardized survey interviewing dictates that interviewers read questions verbatim. Reading questions verbatim ensures that every respondent is presented survey items in the exact same way. Training interviewers on the importance of reading all questions verbatim is one important way to ensure that standardized interviewing procedures are followed during each interview. Knowing which questions interviewers are most likely to misread allows project staff to improve training programs and reduce verbatim reading errors.

The National Survey on Drug Use and Health (NSDUH) provides national, state and substate data on substance use and mental health in the civilian, noninstitutionalized population age 12 and older. Data are collected on a quarterly basis each year, with approximately 140,000 household screenings and 67,500 interviews completed annually. Since 2001, the NSDUH has used field observations as a data quality tool to assess how closely field interviewers (FIs) adhere to project protocols and to subsequently improve training and field procedures. Each quarter NSDUH project staff observe FIs conducting in-person interviews and note any errors committed. Part of the field observation process includes checking if FIs are reading questions verbatim.

This paper reviews the innovative methods and tools developed on the NSDUH to capture and report verbatim errors and exact questions that are observed not being read verbatim. We provide a summary of the types of questions with the most verbatim errors observed on the NSDUH and offer possible explanations for these errors. The types of questions with high verbatim errors observed on the NSDUH may be indicative of common verbatim errors committed on other surveys as well. Being aware of the types of questions most likely to not be read verbatim can help survey managers improve training methods. This presentation will also benefit survey managers seeking to implement or improve upon field observations on their project(s).

A Multivariate Study of Respondent Engagement.

Steven H. Gittelman, Mktg, Inc. (Steve@mktginc.com)

A study by the University of North Carolina found that the predilection to take surveys was genetically influenced. In a controlled experiment, the behavior of one identical twin was a good predictor of the behavior of their sibling, while fraternal twins were not.

In survey research, the engagement level of respondents is a driver of non-response error, as well as a key influence on the quality of responses. Those who are less engaged are less likely to participate in the survey process and often fail to provide quality responses, thus driving measurement error.

In this analysis, the metadata from 1100 diverse studies, emanating from a variety of online sources, as well as covering an abundance of topics, was examined to determine the driving factors of respondent engagement.

There has been considerable discussion as to the influence of structural changes within a questionnaire that might improve respondent engagement. Instruments replete with grids, that drag on over forty minutes, beg the question as to the quality of responses that we are eliciting.

Genetics is an influencer that we, as practitioners, may never have given much thought. However, there are other variables that drive respondents to embrace the survey process that still remain unquantified. In this work, we find that the source of the sample has great influence, only exceeded by the affinity that the respondent have toward the subject matter under study. The impact of these two variables is so dominating, it swamps out structural changes within the questionnaire as a driving factor.

Here we treat a large number of possible variables and present our effort to create a predictive model of engagement. It would be useful to know the impact that various drivers of engagement have on our studies and how variables that we control might create a more engaged respondent.

Using Paradata to Manage Responsive Collection Design.

Francois Laflamme, Statistics Canada (francois.laflamme@statcan.gc.ca)

With the implementation of Responsive Collection Design (RCD), survey managers need better active management tools to assess, monitor and manage the data collection survey process. The main idea of RCD is to constantly assess the data collection process using the most recent paradata information available and adapt data collection strategies to make the most efficient use of the remaining resources. The purpose of the active management in the RCD context is to provide timely and relevant data on survey progress and performance, so that decisions about when and how to adapt collection strategies can be made and problems with collection can be identified and corrected earlier. Paradata (or process data) sources play a key role in the planning, development and implementation of active management for RCD. This paper describes the plans, tools (including key indicators) and strategy used to actively manage the two RCD pilot surveys at Statistics Canada.

Economic Impact Analysis in Indigenous Fishing Communities: A Political Ecology Perspective.

Robynne A. Locke, ICF Macro International (rlocke@icfi.com)

Political ecology explores the conflict over ecological distribution and its effect on natural resources, livelihood, and local environments (Escobar, 2008). In particular, this framework recognizes that conflict over ecological distribution often arises from conditions of unequal power. This is especially true with indigenous communities, whose voices and participation regarding the development of their land and resources have been historically marginalized (Escobar, 2008, p. 3, 15). This history of conflict has presented challenges to conducting research with indigenous communities resistant to outside intervention; the very communities whose participation is necessary for effective natural resource policy (Escobar, 2008). This paper will present ICF Macro's recent experiences addressing this challenge through the implementation of community-based sampling in indigenous Alaskan and Hawaiian fishing communities. The objective of this research is not only to evaluate the effectiveness of this approach in increasing access to traditionally closed social networks, but also to explore the ethical dimensions of a more participatory approach to survey research through the inclusion of indigenous stakeholders.

An Investigation of the Quality of the United States Postal Service Delivery Sequence File for Use as a Sampling Frame for a Commercial Building Study.

An Liu, NORC at the university of Chicago (anliu1101@gmail.com); Susan Hinkins, NORC at the university of Chicago (Susan-Hinkins@norc.org); Kanru Xia, NORC at the university of Chicago (kanru-xia@norc.org)

Surveys have begun using the U.S. Postal Service (USPS) Delivery Sequence File (DSF) as a sampling frame due to the cost of traditional in-field listing. While analysis of the quality of the DSF for household surveys have been well-investigated, how well business addresses are covered by the DSF file is an area for further investigation.

NORC conducted a test in 2006 to check the suitability of using USPS DSF file as the frame for a commercial building survey. Eckman, Colicchia, and O'Muircheartaigh (2007) found in general the USPS DSF doesn't contain all the addresses it was intended to. Under-coverage could bias the survey estimates if buildings which are not covered are very different from those which are on the frame. The 2007 Commercial Building Energy Consumption Study used a list purchased from an USPS address vendor to enhance the 2003 area frame, which consists of buildings recorded from the field. A frame was created from the USPS list by restricting to business addresses and removing duplicate listings that could be identified. The USPS frame and the 2003 area frame were both used for sample selection. Because the 2003 area frame listing does not always include street addresses, it is difficult to determine from the frame listing alone whether or not the same building is included on both frames. For a random sample of the selected buildings for the 2007 CBECS, field work determined whether or not the selected building was included in both frames. We will use this information to investigate the likelihood of under-coverage and properties of the buildings not covered by the USPS listing. In addition, we will compare the frame listings in order to determine if there are any areas where the USPS frame contains fewer listings than 2003 area frame, potentially an indication of under-coverage issues.

Governing in Difficult Times: Analysis of Responses from the 2009 Dayton Citizen Perception Survey.

T. David Jones, Wright State University- Center for Urban and Public Affairs (david.jones@wright.edu); Brittany D. Sumler, Wright State University-Center for Urban and Public Affairs (sumler.2@wright.edu)

Many US cities are dealing with the strain of declining revenues in the wake of the recent recession. Nowhere are these declining revenues more devastating to local government than in America's rust belt-cities that once relied heavily on the automotive industry as a primary source of revenue.

Since 1997, Wright State University's Center for Urban and Public Affairs has conducted a biennial telephone survey on behalf of the City of Dayton to assess satisfaction with city services and support for new policy initiatives (this survey has been conducted since 1979 by another firm). Survey results have been used by the City of Dayton to develop new programs, as well as to determine programs and services that can be cut in times of deficit. The 2009 survey served an additional purpose- to assess citizen views on the refined role of government given recent cuts to the city budget.

The sampling technique used for this survey is unique in that it focuses on sub-samples of priority boards. For more than 30 years, Dayton has used priority boards as a mechanism of government, allowing Dayton's 65 neighborhoods to have a voice in City government. Currently, the city has seven priority boards that serve as "mini city halls", acting as the official voice for their neighborhoods. Elected officials regularly solicit input from these Priority Boards when making policy decisions at the city level. By sampling for the Citizen Perception Survey at the priority board level and weighting data to the city level,

city officials are able to obtain data that is statistically sound at a micro-level, allowing for sound decision making.

This paper will explore public response to the role of government, as well as changes in survey methodology and sampling that have been necessitated by the flight of residents from the City.

The Effects of Emphasizing an Incentive in a Mailed Survey.

Vrinda Nair, Arbitron (vrinda.nair@arbitron.com); Yelena Pens, Arbitron (yelena.pens@arbitron.com)

Arbitron Radio Ratings recruits all household members (aged 12+) for a one-week diary survey of radio listening. Once the person who answered the phone recruitment call ("the consentor") agrees to participate in the radio ratings, the household is mailed a package containing a diary, a small cash gift and, in select households, a bonus insert letting each person know they will receive a \$10 promised incentive (PI) for returning a completed diary. While each person in a PI eligible household receives a bonus insert, there is some concern that respondent may forget they are eligible if the bonus insert is discarded.

In Fall 2010, we tested the effect of adding a sticker to the cover of each diary sent to PI eligible households to inform and remind the diary keeper about the \$10 promised incentive. The purpose of this test was to assess the difference in return and usable diary rates when a sticker was added to the cover of each diary in the promised incentive household. In this test we randomly assigned half of the households that were eligible for a promised incentive to receive their diaries with the reminder sticker while the other half of the households served as a control and received the standard treatment.

In this presentation, we will discuss the impact of emphasizing the incentive on the survey instrument. First, will focus on diary return rates and whether the sticker had an impact on response. Second, we will explore whether or not the sticker affected diary usability (completion of a sufficient number of days and the required demographic data), as this is an indicator of data quality. We hope to provide guidance on whether placing emphasis on an incentive payment has a positive or negative impact on survey response and data quality.

Individuals' Experiences With Public Opinion Data: Fostering Political Engagement Through Multi-Sensory Presentation.

Ann E. Williams, Georgia State University (annwilliams@gsu.edu)

In an increasingly interactive new media environment, the public encounters new opportunities to engage with news. The ways in which public opinion data is collected and disseminated to news audiences is also evolving as technologies advance. This paper employs public opinion data from a randomly selected, representative national sample (N ~ 2200) to address research questions about how individuals' experiences with public opinion data can foster political engagement. The analyses reflect an overarching theme: namely, encountering public opinion data through multi-sensory presentation modes associates with increased levels of political efficacy, trust in the media, trust of data, and other measures of political engagement. The study also highlights significant differences across modes of presentation that involve auditory, visual, interactive, and participatory functions. Data presentation that involves multiple traditional forms of auditory and visual presentation are rated among the most easily understood forms of data presentation, whereas highly interactive, web-based maps are perceived to be significantly less understandable. While all of these different forms of experience with data correlate positively with measures of democratic engagement, experiences that are interactive and participatory share the

strongest connection to such outcomes. This raises normative questions about how best to make data accessible and meaningful to the public – questions that are approached in this paper and that warrant further exploration in future studies.

Using Community Information and Survey Methodology for Bias Reduction to Enhance the Quality of the Air Force Community Assessment Survey.

Zhiwei Zhang, ICF International (zzhang@icfi.com); Christopher Spera, ICFI International (cspera@icfi.com); Jyothsna Prabhakaran, ICF International (jprabhakaran@icfi.com); Jennifer Harvey, ICF International (jharvey@icfi.com)

The 2011 US Air Force Community Assessment Survey is the latest part of a series of surveys of this kind that continues to be a critical tool for senior leaders in determining the strengths and needs of Air Force communities and informing community action planning at the installation, MAJCOM and Air Force levels. As our Airmen and their families continue to rise to a multitude of global challenges, it is important to persevere efforts to promote resilience and ensure the robust participation in this survey. This study will use external public community information to explore the possible relationship between the macro level contextual factors (i.e., substance use and health behaviors at the continental level) and the aggregated survey assessment results. In addition, we will apply the advanced survey methodological techniques to examine the special potential possible noncoverage and/or nonresponse issues in general that may impact the web survey participations.

The Influence Of The Direction Of Likert-Type Scales In Web Surveys On Response Behavior In Different Respondent Groups.

Florian Keusch, Vienna University of Economics and Business, Austria (florian.keusch@wu.ac.at)

Attitude measurement in web surveys mostly relies on requiring respondents to indicate their agreement or disagreement with each of several items under the same Likert-type scale in a grid format. As it is known that respondents do not only attend to the words that convey the questions but also to the visual language of a questionnaire (i.e., format and shape of response scales, verbal and numerical labels of scale points, spacing, positioning, and order of response options) it is essential to understand how this effects the response process. Although there is no conclusive evidence about the influence of the direction of extreme point labeling (e.g., Belson, 1966; Friedman et al., 1993; Salzberger & Koller, 2010; Weng & Cheng, 2000), applying the “near means related” heuristic (Tourangeau et al., 2004; 2007) to horizontal Likert-type scales would suggest that the proximity between the item and the positive anchor of the scale in a agree-to-disagree format would lead to different results than a reversed scale (disagree-to-agree format).

This study aims to bring forward how the direction of Likert-type scales in grid formats influences the response behavior of different respondent groups. In two independent web surveys with online panel members and professionals respondents were assigned to one of four treatment groups. The direction of Likert-type scales as well as the use of numerical labeling of scale points was experimentally varied in a full-factorial 2 (agree-to-disagree vs. disagree-to-agree) x 2 (with numerical labels vs. without numerical labels) design. The influence of scale presentation was measured on different indicators of data quality (response latency, item omission, non-differentiation in grids, response sets).

Saturday, May 14, 2011

1:20 p.m. - 1:30 p.m.

Demonstration Session 3

Demonstration 3A.

Obtaining High Rates of Parental Consent: Additional Evidence for the Field Enrollment Specialist Approach.

Elisha Smith, Mathematica Policy Research (esmith@mathematica-mpr.com); Susan Sprachman, Mathematica Policy Research (ssprachman@mathematica-mpr.com)

Although there is an extensive literature on the benefits of passive versus active consent, there is less information about what types of strategies are most cost-effective and result in higher rates of participation. Our paper extends previous work by examining the differences in cooperation and cost-effectiveness between two years of data from the Universal Preschool Child Outcomes Study (UPCOS).

Mathematica Policy Research first conducted UPCOS in 2007 and it has included child assessments, interviews of parents and teachers, and classroom observations. In the second round of UPCOS (UPCOS2), child care providers gathered consent forms and field staff conducted in-person follow-up at each site. In 2010, Mathematica conducted the fourth round of UPCOS (UPCOS4) and implemented the field enrollment specialist (FES) approach for the first time in the UPCOS study. FES are field staff hired explicitly for the purpose of enrolling sites in the study and gathering consent forms. They perform several specialized tasks to increase consent rates, including attending parent-teacher meetings to describe the study, visiting schools to collect consent forms, and explaining the study to staff at participating schools.

The Head Start Family and Child Experiences Survey (FACES) was the first Mathematica study to implement the FES approach and found the procedure significantly increased cooperation. In additional support of our findings from FACES, preliminary comparisons show the FES strategy on UPCOS has led to increased rates of participation and the reception by teachers and programs has been extremely positive.

This paper will compare the costs of preparation for fielding and conducting data collection between UPCOS2 and UPCOS4. Some variables we will examine include cooperation rates, number of sample releases, and hours spent enrolling providers. The findings from this paper imply the FES strategy can produce cost savings on local projects as well as projects that require extensive travel.

Saturday 1:35 p.m. - 1:45 p.m.

Demonstration 3B.

School Data Collection and Parental Permission Slips: Integrating Project Research Staff in Schools to Increase Return Rates.

Becky Durocher, RTI International (durocher@rti.org); Linda Bailey-Stone, RTI International (owls@rti.org); Lisa McCaskill, RTI International (lmccaskill@rti.org)

Obtaining a high rate of return of active, valid, signed parent permission forms from middle schoolers can be overwhelming and labor intensive for teachers and administrative staff. Studies have described the active parental permission process as an obstacle to high participation rates and have linked higher permission return rates to 1) the researchers' activities and efforts in the schools and 2) reduced burden on teachers.

This presentation examines permission form return rates from nine Southern middle schools participating in a recent national study of teen sensitive issues. Research staff coordinated and managed the permission form process to lessen the burden on teachers and increase buy-in and excitement for the study. During permission processing, some schools permitted project staff immediate access to classrooms and resources to promote the study to teachers, staff, and students; distribute permission forms; explain incentives; stress the importance of returning permission forms; collect forms; and distribute incentives. Project staff had high visibility within these schools. In other schools, administrative staff chose to handle distributing and collecting forms the first few weeks of the process, and then permitted project staff access to teachers and students to help maximize the number of returned forms.

We examined the permission return rates and pace of returns at these nine schools with the level of project staff's involvement and visibility during the schools' permission process; we observed that parent permission form return rates and pace of returns were higher in schools where project staff was more visible with immediate and frequent access to classrooms, teachers, staff, and resources, than in schools that delayed project staff's accessibility to retrieve forms.

Saturday 1:50 p.m. - 2:00 p.m.

Demonstration 3C.

Okay we have permission ... now how do we administer this survey?

Linda K Bailey-Stone, RTI International (owls@rti.org); Becky Durocher, RTI International (bdurocher@rti.org)

The purpose of this presentation is to discuss various methods for conducting in-school student surveys. We will point out the advantages and disadvantages of using different methods for a variety of school populations, depending on the type of survey being administered (sensitive issues versus academic achievement).

Surveying students in schools is a multiple-part process. Parent permission forms are distributed and returned and eligible students are identified. Then, the survey administrators must decide, in concert with the school, when to conduct the survey and how to best organize the survey administration setting while taking into consideration each school's bell schedules, cafeteria uses, and teacher attitudes about students missing class. In this presentation, we discuss various methods of administering surveys to students, the logistics involved in each, and their relative advantages and disadvantages. We'll share some techniques survey staff can incorporate into each of the methods to make the survey a smooth process (from creating hall passes and word searches to using tent cards to flow the kids in smoothly), and lessen the overall burden on school staff and resources.

In some schools, students are surveyed in individual classrooms, while in others they are surveyed in small groups. In still others, schools prefer to survey all students at once in large group settings. Each of these methods can be successful, but all present their own issues. There is no right or wrong way to conduct student surveys in schools. The decision about the methodology to use is a decision that is best made by school administrators and the survey staff. We will aim at providing field-tested ways of dealing with the challenges for each method and/or preventing potential pitfalls.

Saturday 2:05 p.m. - 2:15 p.m.

Demonstration 3D.

Ohio Electronic Health Records Survey: Increasing Response Rates Surveying Medical Practices.
Timothy Sahr, Ohio Colleges of Medicine Government Resource Center
(timothy.sahr@osumc.edu); Elizabeth Stasny, Ohio State University (stasny.1@osu.edu); Daniel
Weston, Ohio State University (weston.45@osu.edu); Lorin Ranbom, Ohio State University
(lorin.ranbom@osumc.edu)

To determine electronic health record (EHR) use rates and adoption rates for Ohio's medical practices, Ohio State University conducted a mixed-mode sample of 4,843 Ohio medical practices (primary care physicians, dentists, pediatricians, medical specialists, and nurse practitioners/nurse midwives). A sample was obtained using Medicaid's 2010 provider list and the National Provider Identifier Registry for Ohio-specific practices. The EHRS utilized a staged process of postcard notification, e-mail notification, mailing of the instrument for completion requesting either a return mail or fax response, follow-up postcard and two follow-up e-mails, with the option of an additional instrument delivery by either e-mail or postage. The survey's main purposes were to determine EHR uptake, barriers to EHR adoption, and the likelihood of EHR adoption. The fielding period was scheduled for 6 weeks. After 3-weeks, only 6.4% of potential respondents returned a completed instrument. With a target return rate of at least 20%, the research team adjusted the contact protocol and initiated a telephone calling campaign to urge practices to complete and return the survey. Options given by callers for completion included: (1) phone interviews with the caller, (2) return by fax, (3) return by e-mail, and (4) return by postage. To encourage response, a blanket approach was employed by faxing all non-responding medical practices the a cover letter and instrument, followed by phone contact. The result was a 19.3% response rate. To factor for the mixed-mode, a judgment post stratification estimator for proportion of having an EHR was employed. The mixed-mode collection was most effective for specialists, primary care physicians, and dentists, and least effective for nurse practitioners/nurse midwives. The conclusion is medical professionals are more likely to respond to multiple contact modes that include a blanketing approach and that faxing medical providers was more effective than electronic communications or postage.

Saturday, May 14, 2011

2:15 p.m. - 3:45 p.m.

Concurrent Session C

2010 Election Polls: Assessments and Mechanics.

Cell-Only Voters in the Exit Polls: An Update From the 2010 Midterms.

Michael Mokrzycki, Independent consultant (mike@mikemokr.com); Scott Keeter, Pew Research Center (SKeeter@pewresearch.org)

This paper reports on the incidence, candidate preference, demographics and attitudes of Americans who voted on Election Day 2010 and reported in response to a question on the national exit poll (in-person interviews) that they only have a cell phone vs. those who have landline telephone service at home. As in the 2008 exit poll - which asked a comparable question - cell-only voters were about one-fifth of the Election Day electorate and, overall, more likely than those with landlines to have voted for the Democratic candidate for House. Within the cell-only share of the electorate, however, there were striking differences by age. For example, while as in 2008 younger voters were more likely than their cohorts with landlines to favor Democrats - in fact those under 30 voted Democratic by an even greater margin in 2010 (40 points) than in 2008 (29 points) - cell-only voters age 40 and up joined their age peers with landlines in shifting dramatically to the GOP. Beyond the vote, there were some notable differences in the composition of the cell-only share of the electorate from 2008 to 2010 - in particular, a 9-point drop in the proportion of 18- to 24-year-old Election Day voters who were cell-only, another indication of differential turnout between 2010 and the historic election of 2008. This paper also will report on differences and similarities between cell-only and landline-reachable respondents by demographic and attitudinal variables, and it will report cell-only incidence among 2010 Election Day voters in Iowa and New Hampshire, which will be critical states in the 2012 presidential nomination fight and where phone-status questions were included in the exit polls. Throughout this analysis, broader implications for non-coverage bias in telephone survey research will be considered.

A Multi-Method Assessment of the Reliability of the 2010 Pre-Election Polls.

Mark Blumenthal, Huffington Post (mark@huffingtonpost.com); Joel D. Bloom, The University at Albany, SUNY (joeldbloom@gmail.com)

In this paper we build on our previous analysis of 2002-2008 election polls (presented at AAPOR in 2003, 2005, 2007 and 2009). We will measure reliability of the 2010 House, Senate and Gubernatorial polls against election results and analyze actual survey error as compared to election results in comparison to what probability theory would lead us to expect. As in our previous papers, we will also use multivariate methods to examine the impacts of a variety of factors, including: (1) house effects, (2) mode of interview effects, (3) likely voter models, (4) proximity to Election Day, (5) margin of victory, (6) number of polls in the race, (7) differentials in campaign spending, and others. Because of the larger than usual number of publicly-reported polls of House races and because of the high error levels in those races, we will focus particularly on those for the first time. In addition, because some pollsters had larger shifts in their results than others over time, we will look at selected pollsters over a broader time frame to see what lessons we might be able to learn about the impact of likely voter models as well as possible bias. We will use the

“pollster.com” data base of polls, making this work directly comparable both to previous work by the authors, as well as others utilizing the same data set.

How Should We Handle Undecided In Vote Preferences in Congressional District Level Political Polls? Empirical Analysis Using Micro Level Survey Datasets.

Masahiko Aida, Greenberg Quinlan Rosner Research (maida@gqrr.com)

Undecided in vote preferences in political polls offer constant headaches to pollsters and its audiences, as often, the proportion of undecided vote is so large that the outcome of the predicted winner depends on the how the future breakdowns among undecided voters.

While there are established research in the logic and methodology of imputing missing data in survey responses such as multiple imputations (Little and Rubin, 1987), in practice, many pollsters do not employ such sophisticated treatment and simply ignore undecided, report raw number or employ untested arbitrary imputation methods of their own.

Another concern is the specification of the imputation model; there are miscellaneous numbers of possibilities for the optimal imputation models that take into account of demographic trait of the respondents, and characteristics of the races (incumbency status, temporal distances from the Election Day and amount of the campaign spending for the particular district) that can potentially play a role.

Author will examine several imputation model specifications (no imputation and different imputation models) and compare estimates from the survey data with and without imputations against actual survey results using 413 congressional district political micro-level survey data sets collected with various pollsters in 2006, 2008 and 2010.

The research will try to answer, questions such as when and how imputation should be employed, and what kind of mode specification will reduce the bias the most while keeping variance as small as possible.

Mode Effects in U.S. House Race Polling: IVR versus Live Interviewer.

Patrick Murray, Monmouth University Polling Institute (pdmurray@monmouth.edu)

The 2009 gubernatorial elections in New Jersey and Virginia witnessed an interesting phenomenon. Interactive Voice Response (IVR) polls, as a group, consistently showed different results than live interviewer polls, as a group, in those two states. In the end, the final IVR polls produced more accurate estimates of the eventual outcomes.

This brought renewed attention to the IVR mode of interviewing and raised questions about the continued relevance of live interviewer methods in election polling. Has non-response bias in live interviewer polls changed so that these polls are becoming less accurate? Are there systematic differences between the types of respondent pools that participate in the two different modes? Were there issues specific to the 2009 polls that can account for the differences in mode performance, such as field period, questionnaire design, and likely voter modeling?

To answer these questions, the Monmouth University Polling Institute conducted an experiment during the 2010 midterm elections. The week before Election Day, separate polls – one IVR and one live interviewer – were conducted in three U.S. House races.

The paper discusses the results of the two modes, using both internal comparisons of weighted and unweighted sample demographics and substantive results, and an external validity check against actual election results.

The methodology for each mode was designed to be as identical as possible, e.g. sampling frame, question wording, likely voter modeling and weighting, and field dates. Some modifications were made to account for inherent differences in administering the two polls. These differences are discussed in the context of real-world applications of the two modes.

"To Vote Or Not To Vote, Are We Asking The Right Questions?: Testing Different Likely Voter Models Across 23 Different Midterm Election Races".

Clifford Alexander Young, Ipsos Public Affairs (clifford.young@ipsos.com)

Overall, the polls did a good job of picking the winners on election day November 02 2010. However, there were criticisms, especially from campaign pollsters who argued that the public polls were, in many cases, not as accurate as campaign polls. In particular, this criticism came into clear relief when many public polls suggested an Angle victory over Reid in Nevada which never materialized.

The methodological criticisms of public polls varied, including low response rates and underrepresentation of Hispanic voters, but the primary one centered on the robustness of their likely voter models. Most public polls use some derivation of the old "Gallup" model typically includes items summated into an index. Likely voters, then, are defined by a "cut point". The criticism of such models is that the likely voter profiles can be unstable overtime, due to sampling and measurement error especially in such low turnout elections, like mid-terms. Instead, the argument goes that using past data from voter lists and other sources provides more robust estimates of who actually will vote on election-day.

We are intrigued by this debate. In our paper, we want to test directly the relative efficacy of the traditional Gallup model versus models-based on voter lists in hopes of improving public pre-election polling.

To do this, we propose a meta-analysis of the Ipsos-Reuters mid-term election polls which would include 23 different midterm races (Gubernatorial, Senatorial, and Ballot initiatives), with a total n size of 13,800 interviews. We would test the accuracy of three different likely voter models: (1) a modified Gallup model used by Ipsos, (2) a voter profile model based on voter lists, suggested by campaign pollsters and (3) as a benchmark, a voter profile model based on past CPS estimates of those who voted.

Comparing Probability & Non-Probability Sample Surveys.

The Use of Web Panels to Characterize Rare Conditions.

John M. Boyle, Abt SRBI Inc. (j.boyle@srbi.com)

There are many uncommon conditions in the United States, which are largely unstudied outside of clinical settings. The prevalence rates of these conditions are too low to be represented in even the largest health surveys. Consequently, there are no reliable estimates of the prevalence, disease or treatment characteristics of these populations.

The population prevalence for one such condition – primary immune deficiency disease – was established by a national RDD telephone survey of 10,000 households in 2005. This survey identified 23 individuals with legitimate diagnoses of primary immune deficiency diseases out of approximately 28,000 persons in these households. The household sample yielded a precise estimate of the prevalence of these diagnoses in the United States, but the patient sample was too small to reliably characterize the disease and treatment.

The current study used a large national Census-balanced online panel as an alternative approach to obtaining a larger, community based sample of patients with a rare disease. A total of 850,000 unique panelists were sent generic invitations to participate in a survey. Ultimately, the online panel produced a total of 159 respondents (with 174 qualifying patients in the household or immediate family) who met the diagnostic criteria for inclusion in the study. This survey provided the first national, community based estimates of the treatment of this rare disease with a large enough sample for reliable estimates. The findings from the non-probability online survey were remarkably consistent with those from the small, but probability prevalence survey. Although no substitute for probability samples, this study suggests that online panels may play a useful role in the monitoring of treatment and health outcomes in rare diseases where adequate community samples cannot be obtained from population based sampling frames.

Measuring Intent to Participate and Participation in the 2010 Census and Their Correlates and Trends: Comparisons of RDD Telephone and Non-probability Sample Internet Survey Data.

Josh Pasek, Stanford University (josh@joshpasek.com); Jon A. Krosnick, Stanford University (krosnick@stanford.edu)

Do Internet surveys of non-probability samples yield conclusions similar to those produced by RDD telephone surveys with regard to distributions of variables, relations between variables, and trends over time? This study explored whether probability sample telephone survey data and data from non-probability sample Internet surveys yielded similar results regarding intent to complete the 2010 Census form and actual completion of the form, the correlates of these variables, and changes in these variables and their correlates over time. Using data collected between January and April, 2010, we found that the telephone samples were more demographically representative of the nation's population than were the Internet samples. Furthermore, the distributions of opinions and behaviors were often significantly and substantially different across the two data streams, as were relations between the variables and changes over time in the variables. Thus, research conclusions would often be different depending on which data stream was used. Because the telephone data collection methodology rests on well-established theory of probability sampling and produced the most demographically representative samples, the substantive results yielded by these data may also be more accurate than the substantive results generated with the non-probability sample Internet data.

More comparisons of Probability and Non-Probability Sample Internet Surveys: The Dutch NOPVO Study.

Rebecca Weiss, Stanford University (rjweiss@stanford.edu); Jon Krosnick, Stanford University (krosnick@stanford.edu); David S. Yeager, Stanford University (yeager@stanford.edu)

Measurement of public perception via survey methods is an expensive endeavor. This is partially due to the cost of probability sampling in recruiting survey respondents. As a result, non-probability sampling has grown in popularity in survey research. While there appear to be some benefits to Internet-based survey methods such as easier administration of survey questionnaires, reduced social desirability bias and reduced survey satisficing, the relative accuracy of non-probability samples has recently been shown to be less reliable than probability samples (Yeager et al., 2009). However, if non-probability samples can be weighted or otherwise analyzed in ways that yield representative results, the merging of Internet data collection and non-probability samples may still be a viable approach to yielding accurate survey measurements.

In this paper, we will present an analysis of non-probability and probability sampling survey data collected in the Netherlands. The Dutch Online Panel Comparison Study (NOPVO) involved 19 Internet survey companies collecting data from non-probability samples of the Dutch population using the same

questionnaire, which asked questions on demographic and non-demographic topics. This paper covers a comparison between the accuracy of these companies and probability samples performed at the same time, using benchmarks from the Dutch Census to determine accuracy. The results of this analysis will further illuminate the relationship between accuracy and web-based survey methods, particularly with respect to the viability of weighting methods and non-probability sampling.

Can a Non-Probability Sample Ever be Useful for Representing a Population?: Comparing Probability and Non-Probability Samples of Recent College Graduates.

Cliff Zukin, Rutgers University (zukin@rci.rutgers.edu); Jessica Godofsky, Rutgers University (jgodofsk@eden.rutgers.edu); Carl Van Horn, Rutgers University (vanhorn@rci.rutgers.edu); Wendy Mansfield, Knowledge Networks (wmansfield@knowledge networks.com); J. Michael Dennis, Knowledge Networks (mdennis@knowledge networks.com)

In spring 2011 the John J. Heldrich Center for Workforce Development at Rutgers University will survey recent college graduates to examine the extent to which their education prepared them for jobs in the US labor market. Contacting this sample--defined as those 22 to 29 years of age and having graduated with a bachelor's degree between 2006 and 2010--is like finding a very expensive needle in a haystack if a typical RDD design is used. The costs are cut considerably by using Knowledge Networks' panel, a nationwide probability sample of some 50,000 members. KN has some 700 such members of their panel who fit the desired profile of respondents.

However, even if we managed to reach and interview two-thirds, we would have fewer than 500 respondents. To augment this number, surveys will be completed with a like number of similarly profiled respondents selected through an opt-in non-probability sample. While we are not sanguine about employing a non-probability design, neither are we sanguine about probability surveys where the response rate may be 20% or so, as calculated by AAPOR standard definitions. Yet, this is where we find ourselves at the beginning of a new decade. Response rates for RDD surveys are threatening to reach single digits, and traditional RDD landline sampling frames miss close to 40 percent if Cell Phone Mainly households are added to those who are Cell Phone Only.

Clearly it is time for experimentation. This paper compares and contrasts respondents selected through a KN probability sample with those obtained through an opt-in non-probability sample. We'll test the null hypothesis that there is no difference between the samples on economic, workforce, political, social and psychological variables. We'll also examine the extent to which we can weight our way out of this problem as survey research enters the

Cost & Operational Efficiencies in Cell Phone Surveys.

A Comparison of Cell and Landline Dialing Patterns

Robert Montgomery, NORC (robert.montgomery@gmail.com); Wei Zeng, NORC at the University of Chicago (zeng-wei@norc.org); Meena Khare, National Center for Health Statistics, CDC (mxk1@CDC.GOV); Karen Wooten, NCIRD, CDC (kgw1@CDC.GOV)

One of the major challenges of conducting cell telephone surveys that are both cost-efficient and achieve high response rates is creating calling rules to optimize contact probability. Comparing dialing patterns from two surveys consisting of both cell telephone and landline samples, we find many similarities and differences as well. We compare calling patterns from data from two CDC-sponsored telephone surveys--the fourth quarter of the 2010 National Immunization Survey (a landline telephone survey with

an additional cell-phone sample), and the 2010 National 2009 H1N1 Flu Survey (a dual-frame landline telephone and cell-phone survey). Based on the findings from the analysis, we make recommendations to increase efficiency and the likelihood of contacting survey respondents in cell telephone surveys.

Minimizing Respondent Burden in Cell Telephone Interviewing: Paths to Efficient Eligibility Screening.

Alicia M Frasier, NORC at the University of Chicago (frasier-alicia@norc.org); Heather M Morrison, NORC (morrison-heather@norc.org); Rebecca Ruthenburg, NORC at the University of Chicago (ruthenburg-rebecca@norc.org)

As landline telephone coverage declines nationwide, telephone surveys are supplementing landline random-digit-dial (RDD) samples with cell telephone RDD samples to minimize potential bias from non-coverage of wireless-only households. While designed to improve survey quality, this move increases respondent burden and may decrease participation rates. Due to legal and ethical concerns surrounding cell telephone interviewing, respondents participating via this contact method are typically subject to additional screening questions ensuring they are able to talk safely (e.g., not driving). Further, under some sampling approaches it is necessary to confirm a respondent's cell telephone status (cell-only, mostly, mainly).

The additional cell screening questions are particularly troublesome for surveys that already have a significant screening burden such as the 2009-10 National Survey of Children with Special Health Care Needs (NS-CSHCN). NS-CSHCN has a two-stage screening process. Cases are screened first for age-eligibility (presence of children under 18 years of age) and second for special-needs eligibility (special health care needs in any of those children). In our initial cell telephone interviewing, determination of age and special-needs eligibility was preceded by cell screening questions regarding respondent safety and cell status. Importantly, the majority of cell respondents then went on to screen out as age-ineligible – making the burden of the cell questions unnecessary.

As a result, we modified our cell interviewing procedures to screen immediately for age-eligibility via a cognitively simple screening question – placing it prior to respondent safety questions. Only age-eligible respondents continued in the survey and were administered the full cell screening significantly minimizing burden for the majority of cell respondents. In our presentation we will examine the ethical implications of this move as well as the impact on survey key indicators such as resolution, age-eligibility, and screener completion rates. We will also consider such an approach as means of containing survey costs.

Cell-to-Landline Cost Ratios for RDD Surveys of Rare Populations

Benjamin Skalland, NORC at the University of Chicago (skalland-benjamin@norc.org)

In planning surveys with a cell telephone sample component, survey researchers often think in terms of a cost ratio – the cost per complete interview in the cell sample relative to that in the landline sample. AAPOR's 2010 Cell Phone Task Force Report includes a discussion of the factors that affect this cost ratio; however, the report does not discuss the impact of screening for eligibility on the relative costs of cell telephone surveys. That is, surveys of rare populations will encounter higher cell-to-landline cost ratios than will general population surveys. In this paper, I describe why this is true and present a cost-ratio formula that separates the components expected to vary depending on the rarity of the target population from the components that should be relatively constant across surveys with differing target populations. I discuss the usefulness of this formula in planning for and monitoring cell telephone data collection, and I provide an example calculation.

Costs Error Optimization for Cell-Landline Dual Frame Surveys.

Vasja Vehovar, University of Ljubljana (vasja.vehovar@fdv.uni-lj.si); Ana Slavec, University of Ljubljana (ana.slavec@fdv.uni-lj.si)

The share of households without a landline phone, but having at least one cell phone, is steadily growing, and so is evidence of corresponding non-coverage bias in landline surveys. Consequently, surveys increasingly use dual frame sampling, typically with a landline share of 60-80%. The aim of this paper is to find the optimal value of this share (i.e. mixture parameter), while considering survey error and costs. For this purpose, the target population is divided into five strata, according to possession and availability of landlines and cell phones. Next, by optimizing the product of costs and the mean squared error across strata, we get the analytical solution (a 4th order polynomial). Finally, this approach is illustrated in eight countries with a 2008 Flash Eurobarometer dual frame survey. The questions cover various socio-economic and political issues: the support of the EU and its currency; attitudes toward economic reforms; and evaluations of household income fluctuations. The optimal mixture parameter is estimated for different variables, countries, sources of population data, and cost options. The solution depends on the landline-cell phone cost ratio. However, it generally lies in a relatively flat optimal area, from 30-70%. Surprisingly, the results are almost invariant to variables, while linear regression (using strata weights and cost ratio) gives excellent prediction of the analytical solution. In countries where the share of the cell-only segment is not very high (less than 25%), it is optimal to have more landline units, and vice versa: where the cell-only segment is larger a predominately cell sample is optimal. However, this connection can be affected by changes in cost ratio, which is the most important factor in determining the optimal design allocation. When the cost of cell interviews is very high in relation to landline phone interviews, taking more landline units is optimal in all countries.

Using Telecommunication Trends to Estimate State Cell-only Population.

Randal ZuWallack, ICF Macro (rzuwallack@icfi.com); Frederica Conrey, ICF Macro (fconrey@icfi.com)

The lack of state and local level estimates of the cell only population continues to be a barrier for weighting cell phone and landline dual frame designs. Research has shown that nonresponse within and between cell phone and landline surveys causes challenges to appropriately combining the dual frames. State-level estimates of cell-only would allow researchers to combine the cell phone sample with the landline sample in the appropriate proportions.

As the number of people who are cell phone only increases, it follows that the demand for landline service will decrease. Decreased demand for landline service is measured by the decline in end-user switched access lines. The number of end-user lines peaked in the early 2000's, but has since been declining. During that time, the number of cell phone subscriptions has increased. Our hypothesis is that trends in cell phone subscriptions and access lines will predict the trend in cell only households. If our hypothesis proves true, then these two industry statistics can be used to estimate the percentage of cell-only households at the state level.

To that end, we are developing a model to estimate the regional cell only households based on the number of end-user lines per capita. The data are from semi-annual regional estimates from the National Health Interview Survey and Federal Communication Commission from 2003 through 2009. We are using a mixed model with an autoregressive correlation structure and fixed effects for landlines per capita and cell phone subscribers per capita. Preliminary results from the modeling process indicate a strong relationship between these two industry trends and the percentage of cell-only households. Our presentation discusses the regional modeling and the model application to state and local areas. The

ultimate goal of the model is to provide state and local cell-only estimates based on local area estimation trends.

Exploring Interviewer & Respondent Error in Cross-National, Cross-Cultural Surveys.

Interviewer Gender Effects in International Surveys.

Neha Sahgal, Pew Forum on Religion & Public Life (nsahgal@pewforum.org); Juliana Menasce Horowitz, Pew Research Center's Global Attitudes Project (jhorowitz@pewresearch.org)

International surveys that are carried out through both face to face and telephone interviews are presumably susceptible to interviewer gender effects, particularly in countries that are largely gender segregated. Survey researchers often address these sensitivities by having interviewers travel in mixed gender pairs to conduct face to face surveys and callbacks in case the respondent objects to being interviewed by a member of the opposite sex.

Implementing strict gender matching is often expensive and/or not possible for logistical reasons. For example, in highly conservative countries where interviewer gender effects could have a bigger impact, female interviewers are often not available. Therefore it is particularly important to understand the role of interviewer gender effects in these contexts. Using data from the Pew Research Center's Global Attitudes Project collected in 2010, this paper examines interviewer gender effects in seventeen countries where face to face interviews were conducted and five countries where telephone interviews were conducted. The paper examines differences in responses given by men and women in different countries when they shared their interviewers gender as well as when they did not. We examine differences in interviewer gender effects by survey mode on economic, social as well as political attitudes and explore the extent to which interviewer gender is a significant predictor of attitudes in these areas. We find that while interviewer gender effects are prevalent in both survey modes, their effects are not uniform across issue areas. Social issues, particularly questions related to gender are particularly sensitive to interviewer gender effects, while economic issues less so. Survey researchers may therefore implement gender matching techniques depending on the nature of the subject matter.

Survey Comprehension Across Multiple Racial/Ethnic Groups: Evidence from the California Health Interview Survey.

Gordon B. Willis, National Cancer Institute (willisg@mail.nih.gov); Salma Shariff-Marco, National Cancer Institute (Shariffs@mail.nih.gov); Timothy Johnson, University of Illinois (tjohnson@srl.uic.edu)

Concern with cross-cultural variability in respondent understanding of survey questions remains an important problem that may bias findings from research conducted in multi-racial and multi-ethnic settings, including many parts of the United States. To address this issue, we report data collected as part of a field test nested within the 2007 California Health Interview Survey (CHIS). Specifically, we analyze behavior codes associated with answers to a newly developed racial/ethnic discrimination module for which a random sample of 500 interviews were recorded and subsequently behavior coded. This sample was stratified to include equal numbers of interviews completed with African American, Latino, American Indian, Asian/Pacific Islander and non-Hispanic white adults. A total of 13,250 respondent answers to individual survey items were behavior coded and included in a hierarchical linear model which examines the effects of race/ethnicity on question comprehension, controlling for other respondent characteristics (i.e., age, gender, education), question characteristics (i.e., length, reading level, time frame and

response format) and interviewer behavior (i.e., reading the question correctly). We will report findings of this analysis, and conclude with a discussion of cross-cultural similarities and differences in respondent understanding of these survey questions. Finally, we outline a future research agenda for continued investigation of this important challenge to the conduct of cross-cultural survey research.

Cross-Cultural Validity of Behavior Codes.

Timothy Patrick Johnson, Survey Research Laboratory (timj@uic.edu); Allyson Holbrook, University of Illinois at Chicago Survey Research Laboratory (allyson@uic.edu); Sharon Shavitt, University of Illinois at Urbana-Champaign (shavitt@illinois.edu); Young Ik Cho, University of Illinois at Chicago Survey Research Laboratory (youngcho@uic.edu); Noel Chavez, University of Illinois at Chicago (nchavez@uic.edu); Saul Weiner, University of Illinois at Chicago (sweiner@uic.edu)

Behavior coding is a potentially useful tool for investigating cross-cultural similarities and differences in respondent comprehension of survey questions. Critics, however, challenge the equivalence of meaning of the behavior codes themselves when comparisons are made across groups. Cultural differences in expressiveness and/or willingness to acknowledge difficulties in understanding can occur. Without an understanding of the origins of such differences, they may be mistaken for comprehension differences when analyzing behavior codes. This study directly addresses this issue by investigating the degree to which behavior codes consistently identify, across multiple race/ethnic populations, comprehension problems using ten survey questions deliberately designed to be problematic that were embedded within a larger survey interview. Approximately 400 CAPI interviews with adults, stratified by race/ethnicity (100 each African American, Korean American, Mexican American and non-Hispanic white) were conducted in Chicago in 2009-2010. These interviews were digitally recorded with respondent consent and subsequently behavior coded. Using bivariate (crosstabs) and multivariate (hierarchical linear modeling) analyses, we report comparisons of the degree to which members of the four cultural groups respond to these problematic survey questions in a similar manner. Consistent patterns of behavior code responses across these groups will be interpreted as evidence supporting the cross-cultural validity of the behavior coding framework. Alternatively, cross-group variability in expression of behavior code problems with these questions will call into question the appropriateness of using standard behavior codes and necessitate additional research on this topic. We will conclude with recommendations for the future use of behavior coding in cross-cultural survey research.

Comparisons of "Good" Interviewer and Respondent Behavior in Bilingual (English and Spanish) Behavior Coding Research.

Patricia L Goerman, U.S. Census Bureau (patricia.l.goerman@census.gov)

A number of bilingual (Spanish/English) behavior coding studies have shown that interviewers read questions as worded and respondents provide codeable responses on the first exchange (what we call "good behavior") at higher rates in English than in Spanish (Goerman et al. 2008; Hunter and Landreth, 2006; Childs, et al. 2007). This has been the case in the first two bilingual behavior coding projects conducted at the U.S. Census Bureau. This may be, in part, because the Spanish and English instruments were not at the same point of development in these early projects prior to the conduct of behavior coding. For example, the English language question wording had been pretested through cognitive interviews prior to being fielded, whereas the Spanish had been translated and put into the field with minimal review. The Census Bureau is in the process of developing an automated Computer Audio Recorded Interviewing (CARI) system for behavior coding. The software is being used for the first time to conduct behavior coding of the 2010 ACS Content Test, a split sample test designed to test both revised and new question wording under consideration for the ACS. Unlike past projects, for the ACS Content

Test joint pretesting through cognitive interviews of both language versions of the instrument was conducted prior to the behavior coding. This paper examines the issue of differences across language in terms of good interviewer and respondent behavior. The prediction is that rates of good interviewer and respondent behavior will match more closely across languages than in past projects where equivalent amounts of pretesting were not done prior to behavior coding. Lessons learned and areas for future research will be discussed.

Does Behavior Coding Capture Cultural Differences in Survey Response?

Martha Stapleton, Westat (marthastapleton@westat.com); Kerry Levin, Westat (kerrylevin@westat.com); Jocelyn Newsome, Westat (jocelynnewsome@westat.com); Stephanie Beauvais, Westat (stephaniebeauvais@westat.com); Salma Shariff-Marco, National Cancer Institute (shariffs@mail.nih.gov); Nancy Breen, National Cancer Institute (breenn@mail.nih.gov); Gordon Willis, National Cancer Institute (willisg@mail.nih.gov); Anne Hartman, National Cancer Institute (hartmana@mail.nih.gov)

Understanding cultural differences in survey response has become increasingly important. Although Asian-speaking participants' response behaviors are of particular interest (Grandy, 1996; Ma et al, 2003; Pan, 2004; Kleiner & Pan, 2005; Pan et al, 2007; Park & Yelei, 2009), much of the research has focused on Asian-speaking populations as a whole. We explore differences and similarities in survey response behavior among Chinese- (Mandarin and Cantonese), Korean- and Vietnamese-speaking survey participants as assessed by behavior coding of two health surveys – Tobacco Use Supplement to the Current Population Survey (TUS-CPS), and the California Health Interview Survey (CHIS).

The TUS-CPS asks questions on various tobacco-related topics such as use patterns and quit attempts. The 2003 version was translated into Chinese (Cantonese and Mandarin), Korean, and Vietnamese. As part of an evaluation of the translations, NCI fielded a small pilot study (n=418) for which the telephone-administered interviews were tape recorded and behavior coded. We will compare selected data from the TUS-CPS behavior coding to behavior coding results from the 2009 field test of a Discrimination Module (DM) in the CHIS. The DM focuses on respondents' recent and lifetime experiences of unfair treatment. We will also compare the behavior coding results of English-language TUS-CPS interviews with English-language interviews conducted in a 2007 field test of the CHIS DM.

We will present similarities and differences among the Asian-language subgroups as well as between those groups and the English-speaking respondents. We will also discuss whether the types of survey items – potentially more sensitive items about discrimination and unfair treatment versus less sensitive items about tobacco use -- elicit culturally specific responses. Implications for future research include a better understanding of cultural differences in survey responses, more accurate interpretation of survey results, and the importance of ensuring inclusion of cultural sub-groups within population surveys.

Respondent Conditioning in Panel Surveys.

Conditioning Effects of Panel Participation.

Ting Yan, NORC (yan-ting@norc.org); A. Rupa Datta, NORC (datta-rupa@norc.org); Peter Scholfield Hepburn, NORC (hepburn-peter@norc.org)

Prior participation in interviews by panel respondents is believed to affect either respondents' actual attitudes or behavior or self-reports of their attitudes or behaviors. Thus, the concern arises that any

changes in attitudes or behaviors may be due to the conditioning effects of panel participation in addition to or instead of actual changes over time.

The 2010 Census Integrated Communications Program Evaluation (CICPE), a survey conducted by NORC at the University of Chicago, provides a unique opportunity for analysis of conditioning effects. The CICPE survey includes a panel sample that was interviewed up to three times. Comparing responses from non-panel respondents to responses from panel respondents, this paper investigates whether panel participation caused conditioning effects in awareness and knowledge of, attitudes toward, and intent to participate in the Census. In addition, CICPE oversampled certain rare subpopulations such as Native Hawaiian and Other Pacific Islander and American Indians and Alaska Natives. We will examine whether or not different subpopulations are differentially subject to the conditioning effects. Of greater interest, access to administrative records from the Census Bureau will allow for further investigation of conditioning effects using actual participation in the Census by subpopulations.

Are You Burdened? Let's Find Out.

Scott Fricker, U.S. Bureau of Labor Statistics (fricker.scott@bls.gov)

For decades, survey researchers have speculated that the burden respondents experience in a survey may negatively impact response rates and data quality. The concept of respondent burden has been operationalized in a number of different ways, but the most frequently used indicator of respondent burden is interview length. The present study examines the results of a small-scale phone panel survey patterned after the Consumer Expenditure Quarterly Interview Survey which collected respondents' detailed expenditure information. In addition to measuring the actual duration of the interview, this survey directly asked respondents how burdensome they felt the survey was, and measured other commonly believed dimensions of respondent burden – e.g., interest in survey content, perceived difficulty of responding to the survey questions, perceived interview length, use of recall aids, and frequency of survey requests. Furthermore, we have information on refusal conversion from a respondent's contact attempt history that enables us to examine the association between respondent reluctance before the interview and perceived burden from completing the interview. Although there have not been consistent findings in the literature, this study provides us with a unique opportunity to better understand the extent to which these various dimensions relate to respondent perceived burden. In addition, using recursive partitioning, a common statistical technique applied in other fields (e.g., marketing, health research), we examine and report on the relative importance of these dimensions, and discuss implication of these findings for survey practitioners.

Examination of Panel Conditioning Effects in a Web-Based 2007-2008 Election Study.

J. Michael Dennis, Knowledge Networks, Inc (mdennis@knowledgenetworks.com); Yelena Kruse, Knowledge Networks, Inc. (ykruse@knowledgenetworks.com); Trevor Thompson, Associated Press (ttompson@ap.org)

Survey panels provide a useful mechanism for longitudinal measurement of within-person change and also can be cost effective compared to conducting another cross-sectional survey.

However, because of sample losses that predictably occur, the panel sample can deteriorate in terms of representativeness or else the respondents themselves might change in their self-reports to survey questions as a direct consequence of their participation in the panel. This latter set of consequences constitutes panel conditioning.

Not all panel conditioning lead to more error. Respondents might become better “survey takers” in being able to report their true attitudes and opinions in answering survey questions (i.e., improved predictive reliability). Other consequences are less helpful, such as actual changes in attitudes, opinions, behavior, and knowledge that can be attributed to panel participation.

In this paper, we’ll accomplish two things. First, we will present a typology of panel conditioning. Second, we will show the results of multivariate regression tests that isolate the impact of prior survey taking on survey responses. We have identified certain political knowledge questions for which there is evidence of limited panel conditioning effects (i.e., more panel experience, more knowledge), while a host of opinion and attitude items did not have such effects.

The data source will be the 2008 Associated Press-Yahoo! News Poll conducted by Knowledge Networks with contributions from political scientists at Harvard University and Stanford University. The study involved an eleven-wave web panel election survey of general population U.S. adults. The baseline data collection occurred prior to the onset of the political primaries (November 2007) and the final data collection took place after the November 2008 general election.

Detecting Measurement Error with Panel Surveys: The 2006-2008 General Social Survey Panel.
Tom W. Smith, NORC (smitht@norc.uchicago.edu)

The General Social Survey is switching from a replicating, cross-sectional design to a rolling panel design. As a first step in this transition, 2006 GSS respondents were reinterviewed in 2008. This paper examines 1) bias from panel attrition, 2) unreliability in the reporting of retrospective or otherwise fixed variables, 3) differences in the measurement error of specific values or responses to questions, and 4) the characteristics of people who tend to be unreliable.

Survey Research on "Don't Ask, Don't Tell" - Methodological Challenges and Solutions.

A Team Based Approach to Analyzing a Large Volume of Qualitative Data on Repeal of "Don't Ask Don't Tell".
Susan Berkowitz (SusanBerkowitz@westat.com); Cynthia Robins

In spring 2010, Secretary of Defense Gates created a comprehensive review working group (CRWG) to understand the potential impact on the military should Congress vote to repeal “Don’t Ask, Don’t Tell” (DADT), the 1993 law prohibiting gays and lesbians from “openly” serving in the military. As part of this effort, the CRWG contracted with Westat to conduct large-scale surveys of Service members and their spouses, which yielded important quantitative data about attitudes towards repeal of DADT. The CRWG also established several other mechanisms through which Service members and spouses could communicate their perspectives on the projected impact of repeal: a) responses to open-ended questions on both surveys, b) comments submitted to a DoD-sponsored inbox, c) real-time online structured dialogues with interviewers, d) town hall exchanges, and, e) small moderator-led focus group discussions. Taken together, these mechanisms generated a large quantity of textual data. The Westat qualitative analysis team developed a systematic, comprehensive approach to analyzing this enormous volume and variety of qualitative data in a short timeframe to answer two key questions: (1) What were the main issues and concerns raised about the impact and implementation of repeal? (2) How did these differ by respondent characteristics, e.g., Service, gender, sexual orientation? This presentation will discuss several key components of our approach including: development of a comprehensive coding scheme that

could be applied across multiple datasets; building of an analysis team grounded in the fundamentals of qualitative research; use of the NVivo8 computer-assisted qualitative data analysis software for efficient data management and analysis; and, strong analytic and organizational leadership of all facets of the process by two highly experienced senior qualitative analysts. We also describe the limitations of our approach, including loss of depth in exchange for breadth when conducting such large-scale qualitative research in a compressed time frame.

Military Members' and Military Spouses' Assessment of the Effects of the Potential "Don't Ask, Don't Tell" Policy.

Kimya Lee, Westat (kimyalee@westat.com)

In spring 2010, Secretary of Defense Gates created a comprehensive review working group (CRWG) to examine the potential impact on the military of Congress's possible repeal of "Don't Ask, Don't Tell" (DADT), the 1993 law prohibiting gays and lesbians from serving openly in the military. As part of this effort, Westat conducted large-scale surveys of service members and their spouses to assess the impact of the repeal of DADT on military readiness, military effectiveness, unit cohesion, recruiting, retention, and family readiness.

This presentation will describe design and operational aspects of the surveys:

- The initial sample size was 200,000 military members and 150,000 military spouses. The military member sample size was increased to 400,000 after the CRWG decided to include a larger portion of the armed services in the study.
- Survey instruments were developed in a team approach establishing the report research questions; identifying concepts that would be measured; assembling existing survey items; developing existing and new items into the draft instrument; and cognitively testing items and conducting expert reviews of the full survey.
- A web-based survey system was established for the service member surveys. Spouses were surveyed by mail. Data were collected over a 6 week period. Response patterns and rates were examined carefully and efforts made to increase participation. A nonresponse analysis was also conducted.
- Extensive team data analysis was conducted in conjunction with the CRWG to focus in on subgroups with and without areas of specific concern related to the possible repeal.

We will discuss several key issues in the survey effort, including: working as a team with the CRWG; conducting the survey in a highly political climate; time constraints to collect data, analyze and convert analyses into findings; deidentification of confidential data; and public comment on the study design.

RAND Survey of Gay, Lesbian, and Bisexual Military Personnel.

Sandra H Berry (berry@rand.org); Ryan A Brown; Terry L Schell

Typically, when a policy change is being considered the key stakeholder groups are actively engaged in the debate and their views are solicited. However, in recent debate about repeal of the Don't Ask, Don't Tell (DADT) policy this has been difficult. Speaking out as a member of the affected stakeholder group would put someone at risk for loss of military career and benefits. RAND's 2010 update of the 1993 RAND report on Sexual Orientation and U.S. Military Personnel Policy provided an opportunity for a

survey effort aimed to allow the key stakeholders to have a say in the DADT debate through a systematic survey.

Because there is no list of service members who are gay, lesbian or bisexual and because there was considerable concern about having only qualified active duty service members participate, we adopted a peer-to-peer approach, building on some of the methods developed in respondent driven sampling. We engaged with nine advocacy organizations that credibly claimed to have vetted active duty personnel among their members and gave them PINs that allowed linking to an Internet questionnaire. Upon completion of the questionnaire, we provided respondents who qualified with 5 additional PINs and asked them to send the links to five other service members who they knew would qualify for the survey. This built the sample in waves, with each qualified respondent having an opportunity to engage others. Of course, it was essential to protect the personal safety and confidentiality of survey participants, so we adopted strict confidentiality procedures and directed the survey to individuals who were directly affected by DADT and the others who worked with them in their units. Thus, a service member did not have to be gay or lesbian to complete the survey or indicate that they were sending it to someone else who was believed to

The Role of the Interviewer in Survey Data Quality.

Interactional Environments, Requests, and Participation in the Survey Interview.

Nora Cate Schaeffer, UW Survey Center, UW Madison (schaeffe@ssc.wisc.edu); Dana Garbarski, Department of Sociology, University of Wisconsin Madison (dgarbars@ssc.wisc.edu); Douglas W. Maynard, Department of Sociology, UW Madison (maynard@ssc.wisc.edu); Jeremy Freese, Northwestern University (jfreese@northwestern.edu)

Previous research, which uses conversation analytic methods applied to detailed transcripts of calls for participation in a survey, has proposed that the actions of sample members may provide encouraging, discouraging, or ambiguous interactional environments for interviewers soliciting participation in surveys. The encouraging or discouraging character of the interactional environment may in turn condition whether the request to participate is delivered and, if it is, how the actions of the interviewer are associated with acceptance or declination of the request. To analyze features of the interactional environment quantitatively, we use audio recordings from the 2004 wave of the Wisconsin Longitudinal Study (WLS) and an innovative design that controls for sample members' propensity to participate in the survey. The propensity score is based on information collected in prior WLS waves. The data are provided by an extensive coding of interviewers' and sample members' actions the characteristics of those actions, and their sequential location in the interaction.

Our model suggests that the encouraging or discouraging actions of sample members initially may be a reaction to the politeness of the interviewer – as indicated either by features of the interviewer's actions (e.g., “Hello” versus “Hi”) or their placement (e.g., identifying herself before versus after requesting to speak to the sample member) -- in the first moments of the call. We also analyze whether a sample member's subsequent actions (e.g., a question about the length of the interview versus a statement of lack of interest) constitute an encouraging, discouraging, or ambiguous environment within which the interviewer must produce her next action. This paper is part of a larger project whose aim is to capture and model the structures of action and interaction in survey introductions.

Effects of Speech Rate, Pitch, and Pausing on Survey Participation Decisions.

José R. Benkí, Survey Research Center, University of Michigan (benki@umich.edu); Jessica Broome, Joint Program in Survey Methodology, University of Michigan (jsbroome@umich.edu);

Fred Conrad, Survey Research Center, University of Michigan (FConrad@isr.umich.edu); Robert Groves, U.S. Census Bureau (Robert.M.Groves@census.gov); Frauke Kreuter, Joint Program in Survey Methodology, University of Maryland (fkreuter@survey.umd.edu)

Telephone survey interviewers vary widely in their success in recruiting respondents, by as much as 15-20 percentage points. This variation implies that the verbal attributes of interviewers, both what they say and how they say it, influence potential respondents in deciding whether to participate. The present study focuses on the prosodic qualities of both interviewers and telephone "answerers" in survey invitations, namely speech rate, median vocal pitch, variability of vocal pitch, and empty pauses, and the relationship of these qualities to the outcome of specific telephone survey invitations: agree-to-participate, scheduled-callback, refusal and hang-up. These relationships are explored in a corpus of 1380 audio-recorded contacts, sampled from five studies conducted by the University of Michigan Survey Research Center for which each utterance is transcribed, acoustically measured and coded in detail. We hypothesize that participation will increase when interviewers speak more quickly, speak with more variable vocal pitch, and pause moderately often, given published reports that speech rate and pitch variability are positively correlated with speaker credibility and animation, and our own previous work showing greater success when interviewers are moderately disfluent (i.e., moderate use of ums or uhs). Preliminary results largely support our hypotheses on these measures. Interviewers speak at about 3.5 words/sec, considerably faster than answerers, and slightly faster when they are successful. Interviewers are significantly more variable in their pitch than answerers, consistent with their role as persuaders in the invitation. Interviewers in successful contacts have short pauses about every other speaking turn, while less successful interviewers pause half as often. We speculate that the longer infrequent interviewer pauses reflect more complex speech planning regarding what they say, while shorter frequent pauses reflect simpler planning regarding how they say it. These results suggest that interviewer prosody is quite relevant to interview participation rates.

Interviewer's Rewording of Questions in CATI Surveys.

Yfke Ongena, University of Groningen (y.p.ongena@rug.nl); Mike Huiskes, University of Groningen (m.huiskes@rug.nl)

Although correct question reading is a fundamental assumption of standardized interviewing, in any survey, interviewers will not always read all questions exactly as worded. We hypothesize that interviewers may change question wording for a variety of reasons. Firstly, interviewer's question rewordings may be considered as deliberate actions. Interviewers may try to make their own job easier by shortening questions that are too lengthy or just difficult to read. They may also be motivated to make the respondent's job easier, by replacing difficult words, adding definitions, or simply shortening question wording. Furthermore, questions that are judged to be offensive may also be paraphrased. Secondly, question rewordings may be more accidental. For example, the more often an interviewer has conducted the same interview, the better the availability of the question wording in the interviewer's memory, and the more likely that interviewers automatically reproduce the question from memory.

In order to test our hypotheses, we analyzed question rewordings in several CATI interview surveys, including the European Social Survey and the Nebraska Social Indicators Survey. Contrary to our expectation, the number of interviews an interviewer had conducted, was positively related to correct question reading. However, question rewordings occurred more often in interviews that were conducted later within a project. This might show that interviewers become sloppier in question reading the more time has elapsed since interviewer training, but it also indicates that interviewers who conduct a large number of interviews are the more reliable ones. Within an interviewing evening, the number of interviews or the amount of time of interviewing was not related to question rewordings. However, within an interview, more rewordings occur towards the end of the interview. This might show that interviewers are more specifically adapting question wording to respondents, since this especially occurred with older and lower educated respondents.

Difficult Situations in Telephone Interviews.

Wojciech Jablonski, University of Lodz, Institute of Sociology (wjablonski@uni.lodz.pl)

The aim of this presentation is to outline the results of the methodological study that was carried out among CATI interviewers from October 2009 to August 2010. 12 major Polish research organizations as well as 2 companies in Norway and Iceland participated in the research (Norwegian and Icelandic modules of the project were co-financed by technical assistance funds of the EEA Financial Mechanism and the Norwegian Financial Mechanism within the framework of the Scholarship and Training Fund). The research was based on a standardized self-completion questionnaire (in total 942 interviewers were surveyed) and in-depth interviews (which were conducted with 49 experienced CATI interviewers). The presentation focuses on the selected results of this project, it investigates the issue of difficult situations encountered by the interviewers. I will outline the results of two questions included in the questionnaire. In both of them the interviewers were given a list of different activities which may result in difficult situations (among others: conducting a long interview, refusal conversion, introducing an interview with an aggressive respondent, conducting many different projects within a single shift). The interviewers were asked to define to what extent they find such situations stressful (first question) and how easy or hard it is for them to perform their duties (second question). Additionally, the qualitative data obtained in in-depth interviews will be presented. The interviewers were encouraged to describe difficult respondents they talk to while working in a CATI studio and to point out the techniques they implement in order to reduce the difficulties. Moreover, the interviewers were asked to elaborate on the relationship between the probability of encountering difficulties while talking to the

respondents and different features, such as b2b/b2c character of the project, socio-demographic characteristics of the respondent, sampling procedures (RDD versus list-assisted samples), etc.

Strategies and Techniques for Improving the Accuracy of Interviewer Observations.

Brady T West, Institute for Social Research (bwest@umich.edu)

Ideal auxiliary variables for use in post-survey nonresponse adjustments are associated with both survey variables of interest and response propensity. Auxiliary variables having these properties will generally reduce the bias and variance in survey estimates. Unfortunately, auxiliary variables available for both respondents and nonrespondents to a survey request seldom have strong associations with key survey variables in practice. As a result, large face-to-face household surveys have started to request that field interviewers record estimates and judgments about selected characteristics of all sampled housing units. Although these auxiliary variables may be associated with survey variables of interest in theory, they will be prone to measurement error, and large amounts of measurement error in these observations may have negative implications for survey estimators in terms of the bias and variance introduced by the nonresponse adjustments. Strategies and practical techniques for reducing the error in these observations are therefore needed in the field. This study first presents results from an analysis of an intervention that was implemented prior to the 15th quarter of the recently completed Continuous National Survey of Family Growth (NSFG). The intervention was designed to provide field interviewers with observable predictors of a key auxiliary variable for which they were recording observations. Analysis of the intervention shows evidence of a significant improvement in the quality of the observations. Next, the results of a qualitative analysis of 3,992 open-ended interviewer justifications for their judgments of the key auxiliary variable are presented, examining the observational strategies used by both accurate and inaccurate interviewers. The study concludes with a discussion of directions for future work in this area.

Using Sample Surveys to Measure the Effect of Messaging.

The Effect of Anti-Smoking Messages Volume and Sources on Motivation to Quit.

Cecilia Hyunjung Mo, Stanford University (mo_cecilia@gsb.stanford.edu); Jon A. Krosnick, Stanford University (krosnick@stanford.edu); Neil Malhotra, University of Pennsylvania (neilmal@sas.upenn.edu)

Public education campaigns often feature messages from a single source, as when the U.S. Surgeon General announced to the American public in the 1960s that cigarette smoking is dangerous. What would have happened if tobacco companies had magnified this information dissemination by admitting publicly what they knew to be true about smoking's health risks? To explore this question, an experiment embedded in a national survey randomly assigned current cigarette smokers to read (1) quotes from public health authorities, (2) quotes from tobacco company executives, (3) twice the number of quotes from both sources, or (4) no quotes, before answering questions measuring motivation to quit smoking. Motivation to quit increased only after people were exposed to the greater volume of messages from both sources. This effect was confined to young smokers (who presumably had more to gain from quitting) and well-educated smokers (who presumably could recognize the credibility of messages from sources with contradictory motivations). The risk messages increased smokers' belief in the connection between smoking and lung cancer, which in turn increased motivation to quit, reinforcing prior findings that such risk perceptions motivate quitting attempts.

How Seeing Movies Changes Attitudes, Perceptions, and Consideration of Military Service: Findings from a Pre-Post Factorial Design Field Experiment.

Robert P. Daves, The Everett Group (rob@everettgroup.com)

The Department of Defense and individual service branches wanted to better understand how their Hollywood movie public affairs efforts affect Americans' image of the military. Using a 2008 Knowledge Networks panel, the Everett Group found high positive attitudes about the military, but after showing respondents movie clips, it found that several variables had different effects on attitudes, and in certain cases seeing the clips actually hurt the military's image. Researchers and military public affairs practitioners posited several causes, including differences in movies' target audiences and seeing clips rather than the full theater experience.

The Everett Group designed a field experiment to provide further insight. It recruited approximately 1,200 subjects, about half of whom were youth, and the other half, "influencers," adults who have influence in a young person's decision to serve in the military. Subjects were randomly assigned to three movie treatments: a military human drama in which military public affairs counselors supported; a military action-adventure in which military public affairs counselors supported; and a control – an animated movie with no military theme or support. Assignments to treatments were controlled to ensure adequate cell sizes for those who had prior exposure, and adequate cell sizes of youth and influencers.

The experiment took place in August, 2010, in the Minneapolis-St. Paul metropolitan area. Subjects completed a questionnaire prior to seeing the movie, and a similar questionnaire after seeing the movie. Data were analyzed using repeated-measures analysis of variance.

Preliminary results suggest that exposure to the human drama movie had strong, positive effect on image. Exposure to the animated control film had somewhat negative overall effects on image of the military. Exposure to the action/adventure film had mixed effects, depending on audience, prior exposure and other independent variables. The paper will more fully explore the findings, including attitudes about control institutions measured.

Effects of Message Framing and Political Disposition on Policy Opinions regarding Direct-to-Consumer Genetic Testing.

Jocelyn Landau, University of Pennsylvania (Jocelyn.Landau@uphs.upenn.edu); Sarah Gollust, University of Minnesota (sgollust@umn.edu); Joseph N Cappella, University of Pennsylvania (jcappella@asc.upenn.edu)

Research has shown that message framing plays an important role in shaping public opinion on emerging issues in science and medicine, but little is known about the influence of elite discourse surrounding genetics on public attitudes and policy opinions. This study examines the effects of message framing and political disposition on public opinion regarding direct-to-consumer (DTC) genetic testing using an Internet-based survey experiment administered to a representative sample of American adults (N=1,012).

Participants were randomly assigned to a message describing expert disagreement over DTC genetic testing. The messages were balanced; one "stripped" message described expert disagreement but included no justification for the arguments, while the "framed" message included pro and con arguments derived from elite discourse. We hypothesized that participants would be more likely to express opinions in the framed message condition and that dispositional characteristics would be more predictive of opinion in the framed condition, indicating more 'anchored' opinion.

As hypothesized, results show that participants were more likely to express an opinion about DTC genetic testing following exposure to the framed message, as compared to the stripped message. The results also demonstrate that political dispositions can moderate the framing effects of elite discourse on policy preferences. In particular, those with higher levels of government trust were more likely to support safety

and efficacy standards for DTC genetic testing after exposure to the framed message, but not the stripped message. Partisanship moderated the message effects on support for relying on the marketplace, not regulation, to govern DTC genetic testing; support for this policy moved in opposite directions among Republicans and Democrats in response to the framed message, but not the stripped message. This study contributes to the literature by documenting the conditional nature of message framing effects, and builds on theories of opinion formation about emergent health policy issues.

Social Disparities, Communication Inequalities, and HIV/AIDS-Related Knowledge and Attitudes in India.

Leland Keith Ackerson, University of Massachusetts Lowell (leland_ackerson@uml.edu); Shoba Ramanadhan, Dana-Farber Cancer Institute (shoba_ramanadhan@dfci.harvard.edu); Monisha Arya, Baylor College (monishaa@bcm.edu); Vish Viswanath, Dana-Farber Cancer Institute (vish_viswanath@dfci.harvard.edu)

India has the third largest number of HIV-infected persons in the world. In the context of a low income country with high rates of illiteracy, knowledge about HIV/AIDS is necessary in order to allow individuals to take measures to protect themselves from infection. The purpose of this project is to determine: (1) how the use of television, radio, newspapers, and movies is socially patterned in India; (2) the extent to which social determinants such as wealth and education are associated with HIV/AIDS-related awareness, knowledge, and stigma in India; and (3) how these associations may be attenuated by exposure to mass media. The data for this study come from the 2005-2006 National Family Health Survey, the Indian version of the Demographic and Health Surveys. The survey was completed by 124,385 women and 74,369 men for individual response rates of 94.5% and 87.1%, respectively. Analyses were conducted using gender-stratified multilevel Poisson regression models to calculate relative risk and 95% confidence intervals. Analyses revealed that there were substantial communication inequalities, with wealthier and more highly educated individuals most likely to use all forms of mass media. Education and wealth were positively associated with awareness of HIV/AIDS, positively associated with knowledge about prevention and transmission of AIDS, and negatively associated with HIV/AIDS-related stigma. These associations were significantly reduced in models that accounted for individual use of the four types of mass media, with television use showing the strongest attenuating effect. These results indicate that mass media exposure may be helpful in changing HIV/AIDS-related awareness, knowledge, and stigma in ways that reduce the risk of the spread of the virus. While substantial communication inequalities exist in India, addressing these inequalities could help to eliminate social disparities in HIV-related illness.

Positions, Perceptions, and Persuasion: Reexamining the Effects of Candidate Messages in Presidential Campaigns.

Andrew Therriault, Department of Politics, New York University (therriault@nyu.edu)

More than half a century of debate among political scientists over the existence of campaign effects has so far only delivered mixed results. The "minimal effects" viewpoint argues that the lack of strong evidence for campaign effects suggests that they play a minor role. In this paper, I contend that the lack of conclusive findings instead results from limitations of our data and analyses--political scientists have failed to find strong evidence of campaign effects because we have lacked the tools to see them. With the arrival of much-improved data sources over the past decade, I revisit the question of campaign effects using panel and rolling cross-sectional data from the 2000, 2004, and 2008 Annenberg Election Surveys. By merging this data with precise measures of television advertising from the CMAG/WiscAds datasets, I show that presidential candidates' messages have much stronger and more diverse impacts on voters than shown in previous research. Voters learn about candidates' issue positions from advertising and

update their perceptions accordingly, and can also be persuaded to change their own policy preferences-- even in the absence of partisan cues. These findings suggest a need to reevaluate our dismissal of the role of campaign messages and events in determining the outcomes of elections.

Sunday, May 15, 2011

8:30 a.m. - 10:00 a.m.

Concurrent Session A

Coverage Follow-up in the 2010 Census.

2010 Coverage Follow-up Instrument: Issues and Recommendations.

Sarah E Brady Clark, U.S. Census Bureau (sarah.e.brady.clark@census.gov); Kelly Govern, U.S. Census Bureau (kelly.a.govern@census.gov)

As part of the 2010 Census, the Coverage Follow-up (CFU) Operation was designed to improve coverage by calling households identified as having potential within household coverage error. While this operation provided unique challenges for universe selection, there was also the challenge in the instrument design beginning with the formation of the questions through available user features. The 2010 CFU Operation was a Computer Assisted Telephone Interview which incorporated several user interface features and complex navigation ability. An early version of the 2010 CFU instrument was tested as part of the 2008 Census Dress Rehearsal leading to several improvements prior to 2010. This paper discusses the instrument issues observed during the 2010 CFU Operation as well as provides recommendations for potential instrument design.

Examining Service Quality and Data Quality in Telephone Interviews.

Julia F Coombs, U.S. Census Bureau (julia.coombs@census.gov); Kelly Govern, U.S. Census Bureau (kelly.a.govern@census.gov)

While the benefits and limitations of measuring service quality in telephone surveys have been studied, not much attention has been given to measuring data quality. The Coverage Follow-up (CFU) operation, a telephone survey that resolved coverage errors in 2010 decennial census data, included two independent quality measures: one to evaluate service quality and one to evaluate data quality. Thus, CFU provides a valuable opportunity to assess the usefulness of monitoring both aspects. The Service Quality Assurance (SQA) program scored recorded interviews on interviewers' telephone courtesy and accuracy of data capture for question groups, while the Data Quality (DQ) program scored recorded interviews on interviewers' script fidelity and accuracy of data capture for certain selected survey questions. The subtle differences between SQA and DQ provide an interesting look at the close relationship between the effectiveness of an interviewer and the quality of the data collected in a telephone survey. This paper looks at the development and design of the two quality measures, compares the information collected from both, and details how real-time SQA and DQ results catalyzed operational changes.

Overview and Results of the 2010 Coverage Follow-up Operation.

Kelly Govern, U.S. Census Bureau (kelly.a.govern@census.gov); Elizabeth Poehler, U.S. Census Bureau (Elizabeth.Poehler@census.gov)

Research during the 2010 intercensal tests focused on the improvement of within household coverage in the Census through expansions to the Coverage Follow-up (CFU) operation. The CFU re-contacted households which indicated potential errors in their household count on their initial enumeration. Historically, the census has been affected by both missed persons and erroneous counts that affect certain demographic groups or living situations more than others. The Census Bureau researched different methods to identify households for CFU such as the inclusion of two questions on the initial enumeration to identify such potentially missed persons and erroneously counted persons as well as the use of administration records.

During the CFU interview, all households were asked a series of questions to identify the missed persons and those erroneously counted. After all of the data were collected, the Census residence rules that state the persons must be counted where they usually live or stay was applied to determine whether the persons should be counted as part of the household. This resulted in either persons added to or deleted from the initial enumeration roster.

This paper provides a brief overview of CFU and the methods used to identify coverage errors as well as presents the CFU results, in terms of percentage of households with either added or deleted people.

Selecting and Monitoring the Workload for the 2010 Coverage Follow-up Operation.

Elizabeth Krejsa Poehler, U.S. Census Bureau (elizabeth.poehler@census.gov); Sarah Brady Clark, U.S. Census Bureau (Sarah.E.Brady.Clark@census.gov)

The 2010 Coverage Follow-up (CFU) operation was designed to make sure everyone was counted once, only once, and in the right place in the census. The CFU was a computer-assisted telephone interview in which we re-contacted households if there was an indication that the initial household count was not right, either because someone was missing from the roster or because someone was counted erroneously. Research this decade indicated that these follow-up interviews could be conducted with approximately 18 million households. Based on budget and time constraints, the workload was prioritized and the CFU operation contacted just under 7.4 million households as part of the 2010 Census. Calls were made from 11 call centers spread across the U.S. with a centralized dialer that could route cases to any available operator at any call center. Each center had its own management structure with the overall program monitored and coordinated centrally.

This paper discusses how we determined which households had potential coverage error, the prioritization of the workload for this operation, and how a workload selected on a flow basis was managed and monitored. It also discusses how the program was implemented, the challenges associated with developing and managing a one-time survey of this size, and lessons that were learned.

Census 2010 Coverage Follow-up Experiment: Mod Q Questionnaire.

Timothy Dale Stewart, U. S. Census Bureau (timothy.d.stewart@census.gov)

One of the reasons the U.S. Census Bureau conducted the 2010 Census Coverage Follow-up (CFU) telephone interviews was to determine if changes should be made to the household rosters. We followed up with households in which we believe someone may be erroneously omitted (undercount persons) or counted in error (over count persons) according to the census residence rules.

Sometimes, no changes were made to the household rosters during the CFU interviews. One potential cause was that these CFU interviews did not provide sufficient cues for respondents to add undercount persons or delete over count persons. An experimental module, called Mod Q, was designed and added to the end of the CFU interview to probe the thought processes of some CFU respondents. We would like

to gain insight into why some CFU respondents did not mention any undercount persons and understand the complex living situations of over count persons. A sample of households were asked the Mod Q questions if no changes were made to the household rosters during the CFU interviews.

This paper focuses on the development of the Mod Q questions and its implementation in the CFU instrument. It also contains a brief discussion on what happened during the 2010 Census CFU operation. It includes some high-level preliminary Mod Q results as well as some recommendations for the future tests in preparation for the 2020 Census.

Cross-National Illustrations of ABS & Other Sample Designs.

Address Based Sampling (ABS) in Puerto Rico.

Mansour Fahimi, Marketing Systems Group (mfahimi@m-s-g.com); David Malarek, Marketing Systems Group (dmalarek@m-s-g.com); Edward Cohen, Arbitron, Inc. (Ed.Cohen@arbitron.com)

Increasingly, the Computerized Delivery Sequence File (CDSF) of the USPS is becoming a standard source for sample survey applications. Through proper augmentations with geodemographic and other ancillary data items, this basic delivery database can evolve into a bona fide sampling frame suitable for complex surveys. While improving coverage, such enhancements enable researchers to develop more efficient sample designs as well as broaden their analytical possibilities through an expanded set of covariates for hypothesis testing and related inferential tasks. Unfortunately, however, at the present time the CDSF does not cover any of the US Territories. In order to bridge this gap, Marketing Systems Group in collaborations with Arbitron, Inc. has developed an innovative technique for generation of addresses in Puerto Rico. Akin to the process of frame construction for RDD sampling methodology, this technique can generate all possible delivery points in Puerto Rico by relying on address parameters and assignment protocols established by the USPS. While providing a complete coverage for the entire Island, the resulting ABS frame will also include addresses that currently do not correspond to existing dwellings. The authors will provide an overview of this promising methodology and discuss the results of a pilot test based on a sample of 3,120 generated addresses in Puerto Rico.

Overcoming Challenges to Sample Design in Iraq.

Matthew Warshaw, D3 Systems, Inc. (matthew.warshaw@d3systems.com); John Willingham, D3 Systems, Inc. (john.willingham@d3systems.com); David Peng, D3 Systems (david.peng@d3systems.com); Nelson Gunter, D3 Systems, Inc. (nelson.gunter@d3systems.com); Kevin Tillman, D3 Systems, Inc. (kevin.tillman@d3systems.com)

Iraq is an extremely difficult environment for conducting survey research. Security concerns coupled with the lack of recent census population parameters present difficult challenges to sample design. There is a need for accurate sampling in Iraq as policy-makers and analysts attempt to shape Iraq's future in light of seven years of societal conflict. D3 Systems has continuously explored innovative approaches in sample design and implementation to overcome these challenges in Iraq since 2003. The main goal is to continue to provide accurate representation of the thoughts and opinions of the Iraqi population.

In the absence of a new census, 2005 estimates from the Iraqi Central Statistical Office are widely accepted as the best source of data for the distribution of survey respondents. However, the lack of up-to-date, consistent, verifiable sample frames raises issues about the representativeness of any sample.

Therefore, D3 began an investigation, with the help of experienced outside reviewers, to develop options for improved sampling in Iraq that do not solely rely on current population frameworks available for Iraq.

The new sampling approach involves eight stages, extending from provincial stratification to the selection of households for interviewing while incorporating the use of satellite imagery technology, combined with ground-level knowledge by our field team. It relies on the use of an automated grid-cell system in Google Earth Pro to create potential primary sampling units and starting points via two additional layers of simple random selection in urban areas. Rural sampling points are drawn using simple random selection of all known rural settlements inside the sub-district.

D3 plans to introduce its new, tested methodology at AAPOR.

Within-Household Selection: Searching for a Better Method Conditional on Household Size Information.

Kien Le, Social and Economic Survey Research Institute, Qatar University (kienle@qu.edu.qa); Abdoulaye Diop, Social and Economic Survey Research Institute, Qatar University (adiop@qu.edu.qa); Darwish Alemadi, Social and Economic Survey Research Institute, Qatar University (dalemadi@qu.edu.qa); Michael Brick, WESTAT (mikebrick@westat.com)

A random selection of a survey respondent at the household level (so called within-household selection) is critical for any valid statistical inference with the survey data. In this paper, we will review existing selection methods. Some methods ensure the randomness but require a lengthy and intrusive process. Some methods provide a quick and simple selection at the expense of the randomness. Although household size information is collected in most of these methods, this information is not fully used. The question that these methods try to answer is how to randomly select a person. In this paper, we argue that the method should be developed not to answer this question but to answer the question about how to randomly select a person conditional on (or given) the household size. Compared to current methods, which are mostly "one size fits all", it is always better to use a method that allows for different ways of selection for different household sizes. The randomness for the whole sample is ensured as long as the selection used for each household size is random. To identify the most suitable selection for each household size, we propose a framework to evaluate various existing methods (Kish, modified Kish, oldest/youngest males/females, last/next birthday, minimally intrusive method) and then either pick one from existing methods or create a completely new one to best fit a particular household size. The final product is a method with different ways of selection for different household sizes. We then used it in a national survey in Qatar, a country with large household size population and most people do not know other household members' birthday, typical characteristics in the Middle East and Developing countries. We will show difficulties when existing methods are applied to these countries and how our method can overcome these difficulties

Response Quantity, Response Quality, and Costs of Building an Online Panel via Social Contacts.
Vera Toepoel, Tilburg University (v.toepoel@uvt.nl)

More than 50% of all survey data in the Netherlands are collected via Internet. However, these data may not adequately represent the views of the Dutch people. The majority of the Dutch people are not willing to join a web panel, and from the people that are in a panel the minority (20%) fills out the majority (80%) of the questionnaires (NOPVO, 2006). Therefore, the answers obtained from web panels can differ significantly from the general population. It is well known that panels contain too many (heavy) Internet users and too few ethnic minorities. So how can we get people into a panel that would normally not join and (hopefully) make the results more reliable?

An unconventional approach is used for building this panel: via social networks. Traditionally one could make the distinction between probability and volunteer opt-in panels. Although most survey researchers agree that probability panels are needed for representativeness, the majority of web surveys is based on volunteer opt-in panels because of budget restraints. Volunteer opt-in panels are prone to selection bias, however. This new way of recruitment may increase representativeness compared to volunteer opt-in panels (recruitment is on invitation only; respondent driven sampling can be used for difficult to reach groups) while keeping the costs at a minimum. By asking respondents via friends and relatives to join the panel, respondents that are normally not willing to join a panel might be persuaded to join. The starting point of building the panel are administrative records of Breda University of Applied Sciences in the Netherlands (about 7000 students with a national spread). I will investigate response quantity, response quality, and costs and give suggestions about when to use this type of recruitment. Note that the Internet penetration rate in the Netherlands is about 90% in 2010.

Election Polling Methods and Considerations.

A Comparative Analysis of Voter Misreport in Two Modes of Interviewing: Telephone and Online.
Kenneth Winneg, Annenberg Public Policy Center (kwinneg@asc.upenn.edu)

One of the great challenges in any research on voting behavior is to base conclusions on self-reported voting measures found in surveys. There are a number of explanations for this, most of which relate to the accuracy of recall in telephone surveys. Generally, surveys ask respondents to report behavior that is either misremembered, or misreported because of social desirability. Further, survey respondents generally are more likely to participate in the political process than those who do not, thus researchers are left with an over-representation of likely participants and a dearth of non-participants. In some instances, fielding a survey of likeliest participants is a beneficial circumstance. However, when one is trying to study the entire population including those who do and do not participate, this over-representation may prove to be problematic. In this paper, I examine the difference in the accuracy of self-reported vote found in two modes of the 2008 National Annenberg Election Survey (NAES): the NAES Online Panel and the NAES Telephone Post-election panel. In doing so, I pose one hypothesis and one research question. I hypothesize that ongoing survey contact increases the likelihood of both self-reported voting and actual voting. While addressing the repeated contact hypothesis, I also answer the research question, which mode of interviewing produced more accurate levels of self-report of voting: Telephone or online. The method I follow to address the research is to match the self-report vote data from the two surveys against a matched list from Catalyst, a voter database company.

How Polling Locations Influence Voting Behavior.

Clint W. Stevenson, Edison Research (cstevenson@edisonresearch.com)

The 2010 National Election Pool Exit Poll conducted by Edison Research provides an excellent and unique opportunity to measure voting behavior and how environmental factors can play a role. The data for this discussion was collected from a large-scale nationally representative Election Day sample of polling location across the United States on November 2nd, 2010. The detailed examination presented here is conducted on how environmental factors can influence an individual's vote for a particular candidate and on particular ballot issues. Several polling location types will be addressed and compared including churches, schools, libraries, and other private/public buildings. To fully compare and understand the voting behavior at these locations this discussion will control for other factors ranging from voter demographics (age, sex, marital status, etc.) to political views to last-minute voting decisions to the

general geographic location of the polling place. The conclusion of this discussion will present how the polling place location at which a person votes can play a role in a voter's decision-making process.

Political Prediction Markets vs. Aggregated Polling Models: Capturing the Dynamic Nature of Political Campaigns.

Gregory Holyk, Washington and Lee University (holykg@wlu.edu); Alan Marco, Washington and Lee University (marcoa@wlu.edu)

Political scientists and polling experts are beginning to recognize the accuracy and usefulness of political prediction markets to predict election outcomes. Recent efforts (e.g., Rothschild, 2009) have directly compared the efficacy of political prediction markets and aggregate polling data.

This relatively young area of research is growing and our project seeks to add to the expanding literature on the topic. One area that requires further analysis is the dynamic nature of the campaign process and its effects on candidate support. Specifically, we examine the sensitivity of political prediction markets and aggregate polling data to significant focusing events and changes during the campaign process.

We examine this question within the context of Senate races in the recent 2010-midterm elections using FiveThirtyEight.com as the aggregate polling models and Intrade as the political prediction market models. We capture the dynamic nature of the campaign process by using Bayesian inference modeling, which has the advantage over most other statistical modeling of elections because it takes into account the previous probability of the outcome (election victory). Bayesian modeling allows us to capture the dynamic updating process of election campaigns and determine whether political prediction markets are better than aggregate polling at reflecting the "updating" that occurs during campaigns.

Conducting Online Pre-election Polls Using Registration Based Sampling.

Michael Barber, Princeton University (mbarber83@gmail.com); Chris Mann, University of Miami (cmann@mail.as.miami.edu); J. Quin Monson, Brigham Young University (Quin.Monson@byu.edu); Kelly D. Patterson, Brigham Young University (kelly_patterson@byu.edu)

This paper presents a method for pre-election surveys that combines the best features of other pre-election survey methods. Surveys are administered online to a probability sample drawn from a voter registration list. A Probability Proportionate to Size (PPS) sample is drawn using a regression model employing variables available in the voter file to produce a probability of voting for each individual in the voter file. The paper examines data from seven elections that cover a variety of electoral situations including primaries and general elections in Utah, Colorado, and Florida in both 2008 and 2010.

The methodology accurately estimates election results despite very low response rates. A counterintuitive possibility of this research is that unlike most surveys where lower response rates increase the chance of non-response error, the low response rates in these surveys may actually decrease the chance of non-response error by improving the coverage of the population of likely voters. This may occur because the time and effort to go online and complete the questionnaire mimics the act of voting and may require a similar level of political interest. Other analysis presents parallel exit polling for some of the surveys that allows for a closer examination of the composition and accuracy of the pre-election poll samples with that of available exit poll data. Finally, in early versions of the methodology the PPS sampling strategy is compared to a Simple Random Sample. The PPS sample is especially accurate in the low turnout elections where producing an accurate likely electorate has historically proven to be most difficult.

Insights from Cognitive Interviewing.

Cognitive Research on Experimental Race and Hispanic Origin Questions Translated into Spanish.

Rodney L Terry, U.S. Census Bureau (rodney.terry@census.gov); Marissa Fond, U.S. Census Bureau (marissa.fond@census.gov); George Higbie, U.S. Census Bureau (george.e.higbie@census.gov)

Surveying and enumerating hard-to-reach populations in the U.S. continues to be a challenge for organizations that provide critical data about the nation's people and economy. Among these populations are Spanish-speaking individuals who may be underserved when questionnaire materials in Spanish are not readily available. As part of a continuing Census Bureau effort to enumerate this population, Spanish translations of 11 Alternative Questionnaire Experiment (AQE) panels of race and Hispanic origin questions were cognitively tested among monolingual Spanish-speaking respondents. Part of the 2010 Census Program of Experiments and Evaluations, these panels are an effort to design questions that reflect the evolving complexity of race and Hispanic origin identification. The results of this test were used to develop moderator guides for AQE focus groups with monolingual Spanish-speaking respondents.

Thirty-three cognitive interviews were conducted with monolingual Spanish-speaking respondents from North, Central, and South American countries. The interviews showed that: (1) most respondents had difficulty selecting a race category, although all respondents were able to select a Hispanic origin; (2) the majority of respondent feedback consisted of difficulty with figuring out how to properly respond to the questions as they were presented, with few opinions about the subtle grammatical details of the questions; (3) respondents with children born in America had difficulty reporting a Hispanic origin for their children; and (4) the most-preferred panels were a combined race and Hispanic origin question that allowed write-in origin lines for most or all of the main race and Hispanic origin categories.

Suspicious and Non-Suspicious Response Patterns which Are and Are Not Problematic: Cognitive Interviewing as Tool for Exploring Quantitative Findings.

Pamela Campanelli, The Survey Coach (dr.pamela.campanelli@thesurveycoach.com); Michelle Gray, National Centre for Social Research (Michelle.Gray@natcen.ac.uk); Margaret Blake, The National Centre for Social Research (Margaret.Blake@natcen.ac.uk); Steven Hope, National Centre for Social Research (Steven.Hope@natcen.ac.uk)

As part of a three-year grant project in Great Britain to inform questionnaire decisions about when and how to mix modes, an experiment was run which not only randomly assigned respondents to different modes of data collection but also to different question formats. These formats included: short versus long scales, rating versus ranking, agree/disagree statements versus balanced forced choice questions, 'yes/no for each' versus 'mark all that apply', branched versus non-branched scales, fully-labeled versus end-labeled scales and showcard versus no showcard on long lists in CAPI. A unique aspect of the project was that it used cognitive interviewing as a follow-up study (on a subset on the same respondents) rather than as a pretesting method. This paper focuses on the question format part of the experiment rather than mixed modes, and contrasts quantitative findings with what was found in the cognitive interviews. Examples will be presented which demonstrate how cognitive interviewing can cast new light on quantitative results. These include instances where quantitative findings would indicate satisficing (e.g., agreeing to opposite statements or rating all questions with the same rating category), but where cognitive interviewing would not or where similar quantitative distributions could imply similar respondent thought processes, but cognitive interviewing shows these are different (e.g., respondents'

processing of attitudinal and behavioral end-labeled scales). The paper also explores some surprising instances where very unlikely quantitative findings (e.g., significant differences in two simple factual questions in 3 and 7 category formats) could be considered a Type I Error, but cognitive interviewing exposes unforeseen respondent difficulties and where choosing a middle category in an aural end-labeled format is not as easy as it may appear to be. The paper concludes with a discussion of the value of using cognitive interviewing to illuminate quantitative findings.

Usability Issues from Testing a Census Web Survey: Results from Testing of the Census Quality Survey (CQS).

Kathleen T. Ashenfelter, U.S. Census Bureau (kathleen.t.ashenfelter@census.gov); Temika Holland, U.S. Census Bureau (temika.holland@census.gov); Victor Quach, U.S. Census Bureau (victor.quach@census.gov); Elizabeth Nichols, U.S. Census Bureau (elizabeth.may.nichols@census.gov)

Although the Census Bureau would like to reduce paper use as well as costs by moving from a paper form to an online version of the Decennial Census, there is concern that many United States residents do not have internet access and it is uncertain whether the situation will change before 2020. Additionally, there is evidence that offering both a mail and an internet option for taking a survey actually lowers the response rate (Smyth, Dillman, Christian and O'Neill, 2010). The Census Quality Survey (CQS) was conducted in order to estimate measurement error, such as simple response variance, from a census Internet questionnaire compared to that from a census paper questionnaire (Hill, J., Reiser, C., and Bentley, M., 2010).

When considering an online version of the Census, its overall usability must be taken very seriously, especially for an instrument that needs to be completed by every resident of the United States. For a data-collection Web survey to be successful, its user interface must support the user in completing the survey in an efficient, effective, and satisfying way. The Census Bureau's Usability Lab conducted two rounds of usability testing of the online Census CQS instrument in April and June of 2010. The goal was to identify elements of the user-interface design that were problematic and led to ineffective and unsatisfying experiences for potential respondents of the survey. Usability issues identified during testing (such as participants having trouble with the auto-tabbing functionality of the "Date of Birth" question) will be discussed along with potential suggestions for the improvement of future online surveys.

Developing a Sexual Identity Measure for the National Health Interview Survey: Results from Cognitive Testing.

Heather Ridolfo, National Center for Health Statistics (hridolfo@cdc.gov); Kristen Miller, National Center for Health Statistics (ksmiller@cdc.gov)

Patterns of poorer health have been noted for sexual minorities, specifically for lesbian, gay, bisexual, and transgender people, leading to the establishment of a reduction in health disparities by sexual identity groups as a federal priority in the United States. The Questionnaire Design Research Laboratory (QDRL) at the National Center for Health Statistics has been tasked with developing a measure of sexual identity for the National Health Interview Survey (NHIS). The goal is to develop a measure of sexual identity that can serve as a gold standard for which others surveys can compare their estimates. However, sexual identity is a complex concept and has been proven difficult to measure on survey instruments due to the multiplicity of identities and how personal identity is tied to particular social locations and experience. The QDRL has conducted multiple evaluations using cognitive interviewing methods on various sets of questions designed to capture information on sexual identity. Analysis consistently shows that the questions produce measurement error due to problems with respondents' comprehension of terminology,

as well as the use of labels that are not consistent with the way respondents identify themselves. Consequently we believe that it is necessary to develop a new measure of sexual identity based on findings from previous cognitive interviews. Furthermore we contend that it is necessary to ask this question using an Audio-CASI instrument, in lieu of the CAPI instrument traditionally used for the NHIS. Using this type of instrument will allow us to explain complex concepts in a consistent way to all respondents. In this presentation we will provide an overview of the research the QDRL has conducted over the years to develop a valid measure of sexual identity, and discuss the progress of our development of a sexual identity measure for the NHIS.

New Relationship and Marital Status Questions: A Reflection of Changes to the Social and Legal Recognition of Same-Sex Couples in the U.S.

Theresa DeMaio, U.S. Census Bureau (theresa.j.demaio@census.gov); Nancy Bates, U.S. Census Bureau (nancy.a.bates@census.gov)

In August 2009, the Secretary of Commerce directed the Office of Management and Budget to establish a task force to research issues related to the collection and tabulation of marriage and relationship data. One focus of the research was family relationships, particularly with respect to same-sex couples who report being married. This paper reports on the second phase of qualitative research conducted by the Census Bureau under the auspices of the OMB working group.

The first phase involved focus groups conducted primarily with persons in cohabiting same-sex relationships. The groups explored the meaning and interpretation of the current decennial Census and American Community Survey relationship and marital status items (see Bates, DeMaio, Robins and Hicks, 2010). Major conclusions were: (1) both items were perceived to be asking about a legal status; (2) few same-sex couples that were not legally married chose "husband/wife" or "now married;" (3) most legally married same-sex couples reported "husband/wife" and "now married" (regardless of place of residence); (4) the marital status categories were viewed as inadequate by unmarried same-sex couples; and (5) relationship categories should be re-ordered according to functional equivalence.

As a result of the focus groups and expert panel review, two alternatives were developed for both relationship and marital status. This paper will report testing of these alternatives via one-on-one cognitive interviews conducted with persons from both same- and different-sex households. To gauge level of sensitivity and potential backlash to the alternatives, some of the interviews will be conducted among heterosexual married persons residing in traditionally conservative states. We will conclude the paper with recommended wording to be further tested in larger scale quantitative content tests.

Methodological Briefs: Using Incentives to Increase Survey Participation.

Impact of Differential Incentive Amounts on Early and Final Survey Response Rates.

Yan Wang, American Institutes for Research (ywang@air.org); Jason Hill, American Institutes for Research (jhill@air.org)

In order to determine an optimal incentive amount for the Beginning Teacher Longitudinal Study (BTLS), sponsored by National Center for Education Statistics, an experimental study was carried out during the 2009-10 administration of BTLS to test the relationship between differential incentive amounts and response rates. The results showed that a higher incentive (\$20 vs. \$10) was associated with both higher early (one month since the beginning of the collection and before the telephone follow-up) and final response rates.

Roughly 2,000 teachers in BTLS cohort were randomly assigned to one of two experimental groups – a \$10 group, or a \$20 group. Teachers received the cash incentives in mail around the same time they received the email to the online BTLS instrument. Comparisons were made between the two incentive groups on the number of interviews before the telephone follow-up date, the number of final interviews, and the number of completed surveys using chi-square tests for association between incentive amounts and different outcome variables.

The result shows that 49 percent of the teachers in the \$10 group and 56 percent in the \$20 group completed the survey or the required items of the survey by the telephone follow-up date. The chi-square test result shows a significant relationship between the number of early study interviews and the incentive amount (chi-square = 10.3463, 1 d.f., $p = .0013$). By the end of the data collection, 86 percent of the participants in the \$10 group and 90 percent in the \$20 group were counted as the study interviews. The chi-square test result shows a significant relationship between the number of final study interviews and the incentive amount (chi-square = 7.6216, 1 d.f., $p = .0058$). However, the chi-square test result shows no significant association between the completeness of the BTLS survey and the incentive amount.

Full-Factorial Experiment on the Effect of Tailoring Incentives, Mail Survey Delivery Options and Questionnaire Length to Increase Response Rates and Reduce Nonresponse Bias.

Sonja Ziniel, Harvard Medical School; Children's Hospital Boston

(sonja.ziniel@childrens.harvard.edu); Jonathan Amatruda, Children's Hospital Boston

(jonathan.amatruda@childrens.harvard.edu); Robert Green, Children's Hospital Boston

(robert.green@childrens.harvard.edu); Noelle Huntington, Children's Hospital Boston

(noelle.huntington@childrens.harvard.edu); Ingrid Holm, Children's Hospital Boston

(ingrid.holm@childrens.harvard.edu)

Much research has been conducted over the past decades aimed at optimizing response rates in mail surveys (Dillman 2007). With the increasing awareness of nonresponse bias, we used a pretest with an embedded experiment to determine the contribution of three factors to response dynamics: the length of the questionnaire (short vs. long), the incentive (\$2 bill vs. lottery), and the postage of survey packet (first class vs. nonprofit). The first goal of this study was to assess which combination of these variables produces the greatest overall response rate. The second goal was to determine whether a certain combination was more successful in gaining participation from subgroups of the population whose nonparticipation could likely introduce nonresponse bias in our statistics of interest (e.g. underrepresented racial and ethnic minorities).

The survey was sent to a random sample of 400 parents of children seen at Children's Hospital Boston over the summer of 2010. The survey contained questions about parents' opinions and attitudes toward participating in a DNA research repository that allows researchers to use participants' DNA to examine genetic causes or moderators for diseases. In addition, the questionnaire assessed opinions about if and how participants would like to receive research results. Each respondent was randomly assigned to one of eight experimental conditions.

Because the sample of the participants was drawn from the medical records, we have access to demographic and other data that will allow an analysis of whether a certain combination of the factors described above minimizes nonresponse bias in several statistics of interest.

The Impact of Gift Cards versus Checks: Comparisons of Daily Response Rate in Panels.

Adam Gluck, Arbitron (adam.gluck@arbitron.com); Yelena Pens, Arbitron

(Yelena.Pens@arbitron.com)

Arbitron Inc. has developed the Portable People Meter (Meter), a new technology for media and marketing research. A panel-based methodology is used to collect Meter related data, and produce media ratings.

Only panelists who wear their Meter for a minimum amount of time each day contribute to the ratings, and these panelists are considered "in-tab" for the day. The percent of eligible panelists that are in-tab each day is considered the daily in-tab or response rate. Having a higher in-tab rate is desirable, as this reduces the potential for non-response bias.

In late November 2009, Arbitron offered panelists an additional monetary bonus to carry their meter during Thanksgiving week. Approximately 8000 panelists in multiple markets were sent a mailing informing them of the bonus opportunity. Half of the panelists were sent a mailer informing them they would be paid any bonus money via check (this is how we typically pay panelists). The remaining panelists were sent an unloaded Visa gift card in addition to their informational mailer. This latter mailer stated that any bonuses they earned would be loaded onto the gift card.

During the Thanksgiving week of 2009, panelists' daily in-tab rates were tracked. In this paper, the differences between the daily in-tab rate of the "check" group and the "gift card" group are analyzed to determine the impact of gift cards on panelists' daily in-tab rates.

Differential Mailing Methodologies on Response Rates: Testing Advance Notices, Blast Telephone Messages, and Package Design.

Yelena Pens, Arbitron (yelena.pens@arbitron.com); Robin Gentry, Arbitron (robin.gentry@arbitron.com)

Arbitron Inc., a provider of radio rating data, conducted a test using an address frame sample with a mail-based approach to recruit respondents for a web-based survey. Since web-based surveys historically have had response issues, there were several initiatives in place not only to increase the response rate but also to find the optimal mailing strategy for mail recruitment.

The first mailing experiment included advance notice mailings. Since the most important mailing asking households to participate in an online survey included a small cash incentive in addition to a promised incentive, the purpose of the advance notice postcard was to highlight that the main package was on its way. The second experiment incorporated telephone blast voice messages along with postcards as an advance notice mailing approach to emphasize the main invitation. Another experiment was a new approach with the idea of a different package design and communication. The purpose of this experiment was to establish whether a more official looking design with a focus on completing a survey as an act of civic duty would appeal to households rather than the current approach of direct mail marketing materials such as flashy bulk packages, brochures, and non-monetary gifts.

In this presentation, we will determine whether advance notice mailers alone increase response rates, whether telephone blast messages enhance the effect of advance notice mailers, and whether the more official looking package increases response rates. Finally, we will present the optimal mailing strategy for mail-based recruitment for an online survey.

Increasing (or Decreasing) Response Rate by Changing the Subject of Email Invitations.

Vani Henderson, Google (vani@google.com); Mario Callegaro, Google (callegaro@google.com)

We ran experiments to study the role of email subject lines on improving response rates to Web surveys in nine countries. In the first study (US and Canada) subjects received an email invitation and two email reminders; we randomly assigned potential respondents to receive different email subject lines for each. For the initial invitation, the subject line was either a question ("Would you like to provide your feedback?") or a polite request ("Please provide your feedback"). The polite request received 21% more responses. For the first reminder, the subject line either appealed to personal interests by reminding them of the incentive ("Provide feedback & receive a mug") or played on the bandwagon effect ("Join other advertisers in providing feedback"). The incentive reminder was 94% more effective. The final reminder either provided a deadline for completing the survey or did not; including a deadline in the subject line was 42% more effective.

We replicated the experiment in Europe, with modifications: no incentive was offered, and subjects were randomized to condition for each email. The body of the email invitation was kept the same across conditions--only the subject line differed. As in North America, the polite request outperformed the question, improving the response rate by 21%. The second experiment tested the 'bandwagon' wording against a strong directive: Provide your feedback on...; the directive improved RR by 18%, a statistically significant difference for 5 of the 7 countries. Finally, in Europe we did not see a difference between a reminder that included a deadline and one that did not; we hypothesize the lack of incentive lead to a lack of urgency to complete the survey on time. We will discuss these findings in light of the literature on email subject lines and provide suggestions for using subject lines to improve RR to Web surveys.

Combining Prepaid and Promised Incentives: Impact on Survey Quality in a Mail Survey of Young Adults.

Luciano Viera Jr., Fors Marsh Group, LLC (lviera@forsmarshgroup.com); Scott Turner, Fors Marsh Group, LLC (sturner@forsmarshgroup.com); Sean Marsh, Fors Marsh Group, LLC (smarsh@forsmarshgroup.com)

The general population has been increasingly less likely to participate in surveys, regardless of mode or sponsorship. As a result, a large body of research has focused on the use of incentives to encourage survey participation. In general, the findings show that monetary incentives are more effective than nonmonetary incentives and that prepaid (unconditional or non-contingent) incentives are more effective than promised incentives that are conditional upon survey return (Church, 1993; Singer, Gebler & Raghunathan, 1999). However, very little research has examined the benefits of combining prepaid and promised monetary incentives on survey quality.

This paper presents the results of an experiment designed to test the impact of offering a promised monetary incentive in addition to an upfront, \$2 prepaid incentive on the survey quality of a national mail survey tracking the future career plans of young adults. Specific measures of survey quality included survey response rates vs. costs, respondent profiles, and key metrics. In addition, because this experiment also tested the feasibility of fielding a longer questionnaire, analyses also examined whether adding promised incentives could counteract any potential decreases in response rates. In order to conduct these tests, a total sample of 30,000 young adults ages 16-24 were randomly assigned to one of six experimental conditions (5,000 each) manipulating survey length (short vs. long questionnaire) and amount of promised incentive (\$0 vs. \$5. vs. \$10). With the exception of content varying according to these six experimental conditions, all materials sent to youth were identical. Everything else was kept consistent to minimize confounds. All six conditions received a \$2 initial cash incentive included in the first survey invitation package.

Implications for existing survey practice and directions for future research will be discussed.

The Effect of Customized Materials on Response Rates.

Michelle A Cantave, Arbitron, Inc (michelle.cantave@arbitron.com)

The Arbitron Radio Ratings recruits all household members (aged 12+) for a one-week diary survey of radio listening. Traditionally we mail the same informational materials to our diary households regardless of the age composition of the household members. In an effort to increase the return rate of diaries from households containing a young adult between the ages of 18-34 we re-designed our mailed materials (the box mailer and the informational brochure) in a manner that we believed would be more appealing to the young adult demographic. In the Fall of 2010 we conducted a test to see the effect of sending customized materials to these households. In this test we sent the modified materials to half of the households with the presence of a young adult and the other half of the households served as the control group. The purpose of this test was to assess the difference in return rate when using the customized materials, whether or not it improved returns from young adults and if there was any negative return impacts on the older people who live with young adults. In this presentation we will discuss the effect sending these materials had on overall response rates and response rates of young adults.

Sampling Racial/Ethnic Minorities in the U.S.

Improving Recruitment Efficiency in Minority Populations: Lessons Learned from The Genetic Study of Nicotine Dependence in African Americans (AAND).

Stacey Weger, RTI International (sweger@rti.org); Christy Hottinger, RTI International (chotti@rti.org)

Since the early 1990s much attention has been focused towards correcting the historical under-representation of minorities in genetic and health research studies. However, traditional population-based recruitment strategies often fall short of adequate sample yield and there is limited literature to guide researchers in the most cost effective minority recruitment methods. Case/Control research studies targeting specific minority populations are presented with the further challenge of recruiting participants who are representative of the community while also controlling for the introduction of potential biases that may result from nontraditional recruitment methods such as treatment seeking, referrals, other community-based recruitment and advertisement.

The Genetic Study of Nicotine Dependence in African Americans (AAND) is an ongoing project with such a design. Data collection efforts for this study are restricted to include only African Americans between the ages of 25-44 in the greater Chicago, IL area, who have smoked a minimum of 40 cigarettes during their lifetime, and who meet eligibility criteria to be enrolled as either a case or a control based on Fagerström Test for Nicotine Dependence scores. AAND was designed to utilize a targeted sampling frame of 100,000 listed telephone numbers from Chicago census blocks known to be 80% or greater African American with an oversampling of households aged 25-44 for enrollment of 1000 cases and 1000 controls.

Research suggests the primary driver of difficulty recruiting minorities is establishing contact with prospective participants. Throughout data collection AAND has faced obstacles in efficiently reaching African American respondents for recruitment via listed telephone sampling methods. In an effort to increase the recruitment efficiency rate for AAND, several supplemental strategies have been implemented with varying degrees of success. This paper will discuss the challenges encountered in

reaching eligible minority household members, outline the various supplemental recruitment methods introduced, and highlight lessons learned from this complex study.

Representative Sampling by Korean Surnames.

Jibum Kim, NORC (kim-jibum@norc.org); Hee-Choon Shin, NORC (shin-hee-choon@norc.org); Diane Lauderdale, University of Chicago (lauderdale@uchicago.edu); Yongmo Lee, Konkuk University (ymlee@konkuk.ac.kr)

Korean surnames are unusual because just a few names account for the majority of the population. Shin and Yu (1984) proposed sampling Korean Americans by simply selecting persons with the surname "Kim," both the most popular Korean surname and one that is distinctively Korean. To support the idea that Kims would be representative of all Koreans they demonstrated that – in Korea – persons named Kim were distributed proportionately across 11 provinces and that the percentage of Kims was constant across several groups (e.g., an alumni association) in Korea. However, they noted that better evidence that Kims were representative would require census data showing the socioeconomic status of persons with different surnames, data they lacked.

Here we use the Korean 1985 Census microdata (N=1,044,742) to examine the geographic, demographic and socio-economic status of persons with the five most common Korean surnames (Kim, Lee (Yi), Park, Choi and Chung). For persons with each of the five surnames, we compare the proportions in categories defined by religion, age, gender, education, number of children, marital status, home ownership, urbanicity and province. We find very little variation across categories for the different surnames. For example, 31.0% of the entire population is single. That percentage only varies from 30.6% of Jeongs to 31.2% of Lees. Similarly, while 15.8% of the entire population is protestant, the percentage only varies from 14.8% of Jeongs to 15.9% of Kims and Chois. There is slightly greater evidence of geographic variation, but the percent living in urban locations just ranges from 64.9% of Jeongs to 65.7% of Lees.

Our results suggest that surname sampling for any of the common Korean surnames, not just Kim, would produce highly representative samples in Korea. However, further work is needed to demonstrate that surname sampling is representative for Korean Americans.

Sampling Methods in Latino Populations: Comparison of Geographic Over Sampling, Listed Frames and Respondent Driven Sampling.

Todd Rockwood, University of Minnesota (rockw001@umn.edu); Melissa Constantine, University of Minnesota (cons0026@umn.edu); Sheldon Swaney, Assessment Unit, Human Services and Public Health Department (Sheldon.Swaney@co.hennepin.mn.us); Karen Virnig, University of Minnesota (Retired) (karen.virnig@gmail.com)

This paper will compare the basic socio-demographic characteristics of respondents three studies conducted in the Twin Cities area in which different sampling methods were utilized. The studies were conducted between 2007 and 2009. One study utilized the delivery sequence file (DSF) with over sampling of geographic areas in which there were high concentrations of Latino's (based on Census estimates). This study was a general public health screening survey conducted using the mail mode with a telephone follow-up (for addresses in which a number was appended). Another study utilized a surname screened listed frame telephone and face-to-face modes were used. The final study utilized respondent driven sampling (RDS) and was conducted using the telephone mode. The last two studies were primarily focused on acculturation and tobacco issues and utilized the same base instrument. Initial analysis demonstrates that there are significant differences between the three different studies relative to language the survey was conducted in (percent Spanish: DSF 61%, listed 67%, RDS 94% X² p <.01),

gender of respondent (percent Female: DSF 59%, listed 49%, RDS 79%, $X^2 p < .01$), birth place (percent US Born: DSF 25%, listed 25%, RDS 9%, $X^2 < .01$), age ($F p < .01$), years in the US for immigrants ($F p < .01$). While not an experimental design, comparison of the profile of respondents across these studies indicate that there are fundamental differences in the respondents and thus potentially the populations being represented by each study. Further analysis will focus on health related quality of life and health behaviors.

Do Different Recruitment Methods Reach Different People?

Hyunjoo Park, RTI International (mpark@rti.org); Lu Liu, RTI International (lliu@rti.org); Mandy Sha, RTI International (mscha@rti.org); Michelle Yuan, RTI International (myuan@rti.org)

This paper takes a closer look at the effectiveness of cognitive interview recruitment strategies for reaching monolingual Chinese and Korean populations in the United States. Results from a preliminary study (Liu et al, 2010) suggest that ethnic newspaper advertisement is advantageous in terms of time efficiency and reach-out capacity, but that publicizing the study via word-of-mouth is effective in reaching people with certain demographic characteristics such as a specific age group or education level.

Our study uses the recruitment data of approximately 800 respondents collected through a cognitive interviewing study of the American Community Survey (ACS) Language Assistance Guide, conducted by the U.S. Census Bureau. Based on prior research, we utilized four different recruitment methods: advertisement in ethnic in-language newspapers, posting flyers at businesses frequented by potential respondents, online advertisement, and word-of-mouth. All forms of advertising share the following message:

- 1) All are directed toward Chinese and Korean speakers;
- 2) All contain information about the purpose of the study;
- 3) All state the length of the interview and the incentive; and
- 4) All mention the survey sponsor.

Once interested candidates responded to the call for participation, the recruiters administered a short, structured questionnaire to screen the respondents. The screening information was then used to select qualified monolingual speakers that represented diverse demographics similar to the actual ACS respondents who requested non-English language materials.

Using the recruitment data, we examine whether there are considerable differences in demographic characteristics among the recruited individuals by the four different methods. For example, older monolingual Chinese and Korean speakers may be more likely to respond to newspaper advertisements than they would to online advertisements. From this analysis, we aim to identify effective recruiting methods for reaching specific subgroups. The results will be of interest to practitioners of cross-cultural studies focusing on the U.S. Asian populations.

Bias in Surveys of Hispanics.

David Dutwin, Social Science Research Solutions (ddutwin@icrsurvey.com); Mark Lopez, Pew Hispanic Center (MLopez@PewHispanic.org); Eran Ben-Porath, SSRS (ebenporath@ssrs.com); Jocelyn Kiley, Pew Hispanic Center (jkiley@PewHispanic.org)

As the largest and fastest growing minority population in the United States, it naturally follows that survey research of Hispanics has become more widespread. As with any branch of survey research, there are a myriad of options regarding how the population is sampled and how the sample plan is operationalized. Typical survey designs include simple but expensive RDD; stratified RDD, “top markets” designs, and

surname designs. To complicate matters, all but the last of these designs can include both landline and cell phone sampling. Every one of these designs covers a different percent on the U.S. Hispanic population. As well, many of these designs sample certain Hispanics at different proportions to other Hispanics. Additionally, how such designs are executed, particularly with regard to the allocation of bilingual interviewers, has tremendous impact on the data acquired. And finally, there is the question of quality, for example, whether response rates, number of call attempts, and the amount of effort given to refusals and callbacks all have an impact on bias in survey estimates of Hispanics. The following paper explores these three foci; sampling, interviewing language, and interviewing effort, by utilizing a meta-analytic approach to the five most recent Pew Hispanic Center National Surveys of Latinos, 2004 to present. We report on what we find to be best practices, and the consequences of failing to enact these practices, as measured by bias in survey estimates of Hispanics.

Sampling/Weighting in Multi-Frame Designs Involving Cell Phone Populations.

The Impact of Cell Phones in the 2010 Elections: More than Just the Cell Onlys.

Leah Melani Christian, Pew Research Center for the People & the Press
(Ichristian@pewresearch.org); **Scott Keeter, Pew Research Center** (skeeter@pewresearch.org);
Michael Dimock, Pew Research Center (mdimock@pewresearch.org)

The number of Americans who rely solely or mostly on a cell phone has been growing for several years, posing an increasing likelihood that pre-election polls conducted only by landline telephone will be biased. Although many national surveys are now conducted using dual frame samples, many state and local surveys still do not interview people on cell phones. An analysis of Pew Research Center pre-election surveys conducted in 2010 finds that weighted estimates based only on landline interviews show greater support for Republican candidates and less support for Democratic candidates than weighted estimates based on the combined landline and cell interviews. The Republican lead among likely voters in surveys conducted in the fall was on average 5.1 percentage points larger in the landline sample than the dual frame sample. That difference is twice as large as in 2008, when Barack Obama's lead over John McCain was on average 2.4 percentage points smaller in the landline sample than in the combined sample.

Preliminary analysis finds that this difference in candidate preferences is a result not only of the inclusion of the cell only, but also because of the addition of dual users from the cell phone frame. Respondents with both a landline and cell phone who were interviewed on their cell phone differed demographically and politically from those reached on their landline phone. This paper will investigate the differences among various types of duals to help further understand why estimates between the combined sample and the sample only of landline interviews differ. In addition, we will explore how the size of the difference between the landline sample and the dual frame sample varies by age and other demographics.

From Dual-Frame to Triple Frame: An Assessment of Coverage Bias in a Telephone Survey Design Combining RDD, Directory-Listed And Cellphone Samples.

Thomas M. Guterbock, Center for Survey Research, University of Virginia (tmg1p@virginia.edu);
Abdoulaye Diop, Social and Economic Survey Research Institute (ADiop@qu.edu.qa); **James M Ellis, University of Virginia Center for Survey Research** (jme2ce@virginia.edu); **John Lee P Holmes, University of Virginia Center for Survey Research** (jlh2r@virginia.edu); **Trung Kien Le, Social and Economic Survey Research Institute** (kienle@qu.edu.qa)

Much recent research explores dual-frame telephone samples drawn from the landline RDD and cell phone RDD frames (AAPOR Task Force, 2010).

Our 2009 AAPOR poster and forthcoming journal article (2011) consider the feasibility of combining Electronic White Pages (EWP, or directory-listed) samples with cell-phone RDD, thus eliminating the usual landline RDD component of the design. We have shown that this proposed “EWP+cell” sampling design would leave only one segment of the telephone population uncovered: unlisted landline households that have no cell phone. Our experience in several local surveys indicates that households with unlisted landlines and no cell phones are quite rare. Differences between the landline RDD and EWP samples are vastly overshadowed by the contrast of each of these with the cell phone sample. Therefore, estimates derived from EWP+cell samples differ little from estimates derived from RDD+cell samples.

This paper describes a large trial of a triple-frame design that combines landline RDD, EWP and cell-phone RDD samples. Our 2009 telephone survey (n=2,600) of the National Capital Region focused on how people would respond to a hypothetical terrorist attack. By oversampling EWP numbers but including all three sampling frames, the triple-frame design blends the complete coverage attainable from the RDD+cell frame with the efficiency and cost-savings attainable from the EWP+cell frame. Because landline RDD households are asked about listedness, we are able to estimate the percentage of landline phones that are unlisted, and weight accordingly. After poststratification weighting on key demographics, we find little substantive difference in results between our new design and the standard dual-mode design (RDD+Cell).

A triple-frame design can deliver closely comparable results at substantially reduced cost and thus offers a practical approach to adapting sampling methods to fit the emerging future of telephony in the United States.

The Impact on Accuracy of Estimates of Increasing Cell Telephone Sample to Correct Coverage Error in the Survey of Consumer Attitudes.

Mahmoud Elkasabi, ISR - University of Michigan (mkasabi@umich.edu); Zeynep Suzer-Gurtekin, ISR - University of Michigan (tsuzer@umich.edu)

The Survey of Consumer Attitudes (SCA) is a monthly national RDD landline telephone survey conducted by the Survey Research Center at the University of Michigan. SCA has since 2009 added a small sample of cell telephone RDD samples to supplement the present landline RDD sample. Combining landline and cell samples requires a composite weight based on an overlapping dual frame estimation design. Probability of selection for a household is computed based on telephone domain (landline only, cell telephone only, and dual users) and frame. Probability of selection weights are poststratified to the national telephone domain proportions obtained from the National Health Interview Survey. Since the cell telephone samples are small, relatively large weight values are generated for cell telephone frame sample cases, leading to inefficient estimates. In addition, the number of cell- only households, which are not covered in the landline frame sample, is small. Data from 18 consecutive months collection were used to simulate what would happen if substantially larger sample allocations were made to the cell frame sample. Simulated allocations were compared in terms of precision and changes in survey estimates and data collection cost. Potential increases in variance of key survey estimates is substantially reduced by the higher allocations to the cell telephone frame, while relatively minor changes in five out of eight key survey estimates occurred for higher cell telephone allocations; change in the other three key estimates was moderate.

Accuracy of Geographic Stratification in a Cell-phone Survey

Robert Montgomery, Xian Tao Kirk Wolter NORC at the University of Chicago and Stacie Greby Erin Kennedy, Centers for Disease Control and Prevention. Kirk M. Wolter, NORC and University of Chicago (wolter-kirk@norc.uchicago.edu)

Detailed geographic area information is needed from respondents to monitor implementation of public health programs; however, stratification by detailed geographic area is difficult or impossible in surveys sampling cell telephones. We analyze data from the November 2010 Rapid Flu Survey, sponsored by the Centers for Disease Control and Prevention, in which the cell telephone sampling design employed stratification by county and where county was derived from the wire center code. Our main analysis compares the county employed in stratification to the county obtained from the respondents' reports of actual physical location. We measure the degree of inaccuracy for each area where the data allow.

Assessment of Bias from Incomplete Frame Coverage and Other Sources in a Random Digit Dial Survey: Applications of a Supplement to the National Health Interview Survey.

Kennon R. Copeland, NORC at the University of Chicago (copeland-kennon@norc.org); Meena Khare, National Center for Health Statistics (mkhare@cdc.gov); Philip J. Smith, National Center for Immunization and Respiratory Diseases (pzs6@cdc.gov); Reiping Huang, NORC at the University of Chicago (huang-reiping@norc.org); Lin Liu, NORC at the University of Chicago (liu-lin@norc.org); Kirk M. Wolter, NORC at the University of Chicago (wolter-kirk@norc.org)

This paper examines differential bias (due to all sources of sampling and nonsampling error) between estimated vaccination coverage rates obtained from the National Immunization Survey (NIS) and the National Immunization Survey-Teen (NIS-Teen), both landline RDD surveys, and from a supplement to the National Health Interview Survey (NHIS), an address-based survey administered in person. While bias may arise in all surveys due to a variety of sampling and nonsampling errors, a main concern for the NIS surveys is coverage bias due to the increased size of the cell-phone-only population.

The NIS and NIS-Teen provide annual estimates of vaccination rates for children age 19-35 months and 13-17 years, respectively. Both surveys collected information on doctor's visits, general health information, insurance status, and demographics. With parental consent, a mail survey (called the Immunization History Questionnaire, or IHQ) was sent to the children's immunization providers to collect vaccination histories. Beginning in 2009, parental consent was requested for NHIS age-eligible children and an IHQ was sent to the children's immunization providers to collect vaccination histories – the resultant data are referred to as the NHIS-PRC (Provider Record Check). As a result, vaccination rate estimates were derived for the NHIS sample which covers the entire population regardless of telephone status. Comparisons of the NHIS estimates by telephone status allows estimation of coverage bias of a hypothetical landline-only survey and further comparisons to the NIS and NIS-Teen estimates allow for indirect assessment of coverage and other sources of bias in the NIS by assuming estimates from NHIS are less biased overall.

We compare vaccination rate estimates from 2009 NIS and NIS-Teen to those from the 2009 NHIS-PRC, and analyze socio-demographic and telephone status distributions to explain differences in estimates.

Understanding Panel Attrition.

Panel Attrition and the Survey Experience.

Sunshine Hillygus, Duke University (hillygus@duke.edu); Trevor Thompson, Associated Press (TThompson@ap.org)

Panel surveys are an important source of attitudinal and behavioral data, but analyses from panel data can be threatened by attrition, which occurs when individuals who participate in the first wave of a panel drop out in later waves. Panel attrition reduces the effective sample size and can introduce bias in survey estimates if the tendency to drop out is systematically related to the substantive outcome of interest. In this paper, I will examine the predictors on panel attrition in the General Social Survey, the American National Election Study (both the GSS and ANES recently added a panel component) as well as the 2008 Associated Press/Yahoo News Election Study. This comparison will be particularly useful because these studies come from different academic disciplines, have different panel designs, and use different survey modes. The analysis will examine the effect of individual-level characteristics that might predict attrition, such as mobility, citizenship status, household status, education, etc, as well as paradata such as interviewer experience, duration in online panel, initial contact efforts, length of interview, extent of item nonresponse, and individualized incentives. Preliminary analysis finds that many of these latter factors are among the strongest predictors of attrition. Finally, the analysis will compare those who drop out of the survey entirely with those who drop out from a wave but then return to a later wave.

Identifying Potential Attrition Bias using Paradata, Sampling Frame Information and Survey Data.
Darryl Creel, RTI International (dcreel@rti.org)

Longitudinal surveys typically experience some amount of attrition from one round of data collection to the next. The major concern that arises because of attrition is the potential for estimates computed from the respondent data to be biased because respondents alone may not accurately represent the survey population. With a focus on attrition, this paper investigates the potential for attrition bias in the second round of a longitudinal survey. We evaluate the relationship between potential attrition bias and a wide array of data, including paradata from the data collection process, sampling frame data, and respondent data from the first round of the survey. In order to screen such a large number of possible variables that may be related to attrition bias, we use a tree-based methodology to identify a smaller group of variables to use in multiple regression and multiple logistic regression modeling to compare the proportional distribution of characteristics of respondents and nonrespondents to determine if nonresponse bias exists. This analysis is conducted for all respondents and important subgroups. To minimize potential bias, we discuss how we adjusted the weights using information from this evaluation. This research adds to the growing literature on the utility of paradata in evaluating nonsampling bias and provides researchers with a guide to identify and account for attrition bias.

Respondent Engagement, Panel Fatigue and Other Determinants of Attrition in the Gallup Panel.
Robert J. Torongo, The Gallup Organization (Bob_Torongo@gallup.com); Tim Gravelle, The Gallup Organization (tim_gravelle@gallup.com)

Longitudinal survey panels have some unique and useful qualities of interest to researchers. As the longitudinal panel develops over time, panel researchers gain an in depth knowledge of the characteristics of their panel members. However, these characteristics can only be utilized if panel membership stable and attrition (a respondent choosing to leave the panel) is low. Also, since recruiting new panel members is both costly and time consuming, panel researchers need to be aware of the factors driving attrition and take steps to mitigate those factors when administering a series of surveys over time. Using the Gallup Panel, a probability sample-based multi-mode consumer panel, we will examine the effects of respondent burden (understood not just as survey length, but also the number of survey invitations over time), survey mode, respondent engagement and respondent demographics on panel attrition. This paper is a useful addition to the field for panel researchers and will present data that will ultimately aid in further understating panel attrition.

Comparing Two Telephone Panel Surveys Conducted 25 Years Apart: Similarities and Differences in Factors affecting Panel Attrition.

David R Johnson, Pennsylvania State University (drj10@psu.edu); Veronica Roth, The Pennsylvania State University (vlr124@psu.edu)

It is well known that response rates have declined in the last decade for RDD telephone surveys. Less well known is the impact of the more challenging telephone survey climate on response rates for longitudinal panel studies, administered over the telephone. Attrition rates for follow-up panel studies may be higher in the recent decade than in earlier years, but the reasons for nonresponse and potential nonresponse bias may also differ. In this study, we compare two similarly designed longitudinal panel studies—one (Marital Instability over the Life Course) conducted in 1980 with a second wave interview in 1983, and another study (National Survey of Fertility Barriers) conducted in 2005 with a second wave interview in 2008. Both surveys were based on nationally representative RDD samples. Because the earlier survey was restricted to married respondents age 18-55, and the second survey to women age 25-45, we restricted the samples used here to married women age 25-45. For the 1980 study, 856 respondents met the criteria, while 2,002 met the criteria in the 2005 study. Panel attrition in both studies was modeled with logistic regression. Next, the similarities and differences in demographic and paradata factors that accounted for the attrition were scrutinized. Based on these models, sample weights were developed to adjust for the nonresponse, and the design effects were contrasted for a number of variables that were measured the same way in both datasets. The paper concludes with a discussion of the challenges of conducting panel studies based on RDD telephone samples in today's survey research climate.

Characteristics Influencing Biomeasure And DNA Cooperation Rates In A Longitudinal Study Of Older Adults.

Katie Lundeen, NORC at the University of Chicago (lundeen-katie@norc.org); Jane McPhillips, NORC at the University of Chicago (mcphillips-jane@norc.org); Krishna Chokshi, NORC at the University of Chicago (chokshi-krishna@norc.org); Angela Jaszczak, NORC at the University of Chicago (jaszczak-angie@norc.org)

Over the past several years, population-based surveys that integrate an in-home interview with biomeasure collection have become increasingly more common. Respondents that participate in these surveys have the right to refuse any biomeasure, which results in missing biomeasure data. Missing data may impact the representativeness of the biomeasures being collected, which would impact the study's overall data quality. In previously published literature, the first wave of the National Social Life, Health and Aging Project (NSHAP) and the 2006 Health and Retirement Study (HRS) began to investigate the willingness of respondents to cooperate with biomeasure collection within surveys. This paper will expand on this biomeasure cooperation literature using data from the second wave of NSHAP, a longitudinal survey of approximately 3,000 older adults that examines the interaction between aging, social relationships, and health outcomes. In both waves of the study, non-medically trained interviewers administer up to thirteen biomeasures to respondents including blood spots, DNA, urine, actigraphy, and self-administered vaginal swabs for female respondents. In this paper, we will present the biomeasure cooperation rates for Wave 2 of NSHAP and examine characteristics that may promote or preclude cooperation. Specifically, we will examine variation in cooperation by respondent characteristics, interviewer characteristics, and spouse's participation in the study. In addition, we will explore whether biomeasure cooperation remains consistent across both waves of data collection for the project and individual respondents. Finally, we will compare DNA cooperation rates across two collection modes offered to respondents in the Wave 2 pretest. Respondents could collect the sample during the in-home

interview with the interviewer present or collect the sample after the interview and mail it back. We will explore whether collection mode impacts respondent cooperation, as well as data quality.

Sunday, May 15, 2011

10:15 a.m. - 11:45 a.m.

Concurrent Session B

Assessing Public Attitudes Towards Minorities.

Conflict Over Discrimination Against Ethnic Minorities, Smokers, and the Overweight: Australian Evidence. S. M.C. Kelley, Yale University (sarah3@international-survey.org)

Public conflict over job discrimination against immigrants and ethnic minorities has been widespread for decades in many western nations. In recent years prejudice and threats of discrimination have also emerged against other groups, notably smokers and overweight people. Using Australian data, I offer quantitative estimates of the level of prejudice, social distance, and willingness to discriminate against immigrants from Vietnam, Greece, Britain, and the USA, and investigate the sources of those attitudes in social structure, economic circumstances, and religion. I then extend the analysis to smokers and the overweight. Data are from the discrimination module of the 2004 International Social Science Survey/Australia, a representative national sample of Australian adults. N=1533. Factor analysis confirms three clear and reliable multi-item measures of prejudice against immigrants of different nationalities; social distance from them; and willingness to discriminate against them in hiring. I assume a causal order in which prejudice potentially creates social distance and both in turn are potentially causes of discrimination. Regression analysis shows that both prejudice and social distance have strong effects on willingness to discriminate economically against immigrants, as has long been known. Attitudes toward different immigrant groups are broadly similar, with the public making little distinction between them. There is great prejudice against smokers – more than against migrants – which leads to substantial social distance and willingness to discriminate against them. Sympathy with migrants is not correlated with sympathy for smokers. Levels of prejudice against overweight people are similar to those against migrants, with similar consequences. Willingness to discriminate against migrants, smokers, and fat people are all highly correlated. This suggests that there are general views about whether or not it is just to discriminate at all, as well as varied and differentiated feelings of sympathy and social distance in regard to migrants, smokers, and the overweight.

Do Individual Characteristics Matter? An Analysis of Americans' Opinions Towards the Korean, Vietnam, Persian Gulf, Iraq Wars, and the War on Terror.

Emanuel Boussios, Nassau Community College (emanuel.boussios@ncc.edu)

This paper presents the results of an analysis of public attitudes towards war for the Korean, Vietnam, Persian Gulf, Iraq wars and the War on Terror. In total, 61 surveys conducted by the Gallup Organization starting in August 1950 (near the beginning of the Korean War) and ending February 2006 (middle of the Iraq War) were analyzed. This paper examines the social characteristics of those people who are more or less likely to oppose war and is a continuation of the author's earlier work analyzing the Iraq War. The research reported here was arrived at answering the following question: are there some groups (e.g. Democrats) who are more likely to oppose war? Using binary logistic regression methods, this research analyzes why some groups are more likely than others to oppose or support war. This cross-sectional analysis studies groups by age, education, gender, race, religion, party identification, and political ideology, in their general likelihood to oppose war. In addition, this analysis studies respondents' attitudes

towards war which includes whether the respondent was a hawk or a dove, an internationalist or isolationist, and optimistic or pessimistic about the war's outcome. Several traditional expectations on war attitudes, including a "race gap" and a "liberal-conservative gap", were unfounded. In some cases where correlations were found, the relationship could be explained predominately by the correlation of demographic characteristics with partisanship. This study has demonstrated, over time, just how much more powerful party identification has become in the United States. This trend analysis is the closest look at U.S. public opinion on wars after WWII since the ground breaking and prize winning work of John Mueller (1973, 1994).

Public Attitudes Toward the Homeless.

Robert P. Agans, Survey Research Unit, Dept of Biostatistics, UNC-CH (agans@unc.edu)

Homelessness in the United States has increased substantially since the late 1970s and has been a central political topic for several decades. A lack of affordable housing, the deinstitutionalization movement, substance abuse, and unemployment all contribute to the increase in homelessness. Consequently, the U.S. has one of the highest rates of homelessness among developed nations as a significant number of individuals are either chronically homeless or experience short bouts of homelessness. Studying homeless populations, however, presents major methodological challenges because the living situations of these individuals can change daily, making them difficult to locate and contact. Research data plays an important role in supplying valuable information to government agencies and nongovernmental organizations who provide services to these vulnerable populations. Two surveys were conducted in Los Angeles City and County as part of the 2009 Greater Los Angeles Homeless Count: i) a face-to-face survey that randomly sampled 3,073 adult homeless persons living on the streets or in shelters; and ii) a telephone survey of 739 adults living in households with landline telephone access. Results found that there is a large discrepancy between public opinion on the causes of homelessness and the actual causes of homelessness as reported by the public at large and among the homeless themselves, respectively. In addition, we identify predictors among the general public that are likely to be linked with increased sympathy for the homeless as well as a willingness to help the homeless.

Political, Social and Demographic Factors Underlying Public Reaction to the "Ground Zero" Muslim Community Center and Mosque.

Mark A. Schulman, Abt/SRBI (m.schulman@srbi.com); Benjamin Phillips, Abt SRBI (b.phillips@srbi.com); Seth Brohinsky, Abt SRBI (s.brohinsky@srbi.com)

A controversial plan to build a Muslim community center and mosque near the World Trade Center site in New York City earlier this year stirred passions and unleashed a wave of anti-Muslim sentiment nationally. In the midst of this controversy, Abt SRBI conducted a national cross-section sample of 1,000 adults in August 2010 to assess not just views on the mosque, but some of the underlying reasons for those views.

We analyze attitudes toward Muslims across a number of dimensions: general feelings toward Muslims; views of Islam encouraging violence against non-believers; views of Muslims' patriotism; perceptions of the community center; views of Muslims as compared to other religious groups building religious center near home.

Specifically, political, social and economic factors that may underlie reaction to the mosque siting will be modeled to examine and illuminate underlying attitudinal structures. We will also examine the extent to which the reaction was a manifestation of current political polarization. A number of features of the survey

in addition to standard sociodemographic measures like education, income, party identification, and political ideology support particularly nuanced contextual analysis. We measure general religious prejudice, defined by views of other groups besides Muslims (Protestants, Catholics, Jews, and Mormons) and views of a religious center being established near home (Catholics, Jews, and Mormons). Respondents were asked whether they were personally familiar with a Muslim, allowing us to test whether personal familiarity with Muslims lessens perceptions of threats. We also have a measure of fringe views in the form of the respondent's belief that the president is a Muslim. Finally, the fact that the survey includes multiple measures of attitudes toward Muslims enables us to analyze potential disjunctions in survey responses (e.g., individuals with negative views of Muslims who support the right to build the center).

Investigating the Terrain of Hate Crimes in the Oklahoma Public Education System: Are we Facing a Decline in Tolerance in an Era of Purported Ethnic, Racial, and Social Diversity?

Amy Sue Goodin, University of Oklahoma Public Opinion Learning Laboratory (amysgoodin@ou.edu); Natalie Jackson, University of Oklahoma Public Opinion Learning Laboratory (nataliemjackson@ou.edu); Sheryll Brown, Oklahoma State Department of Health (sherylls@health.ok.gov); Andrea Hamor Edmondson, Oklahoma State Department of Health (AndreaH@health.ok.gov)

Since that fateful day on September 11, 2001 the United States has been confronted by growing tensions and societal conflict that not only transcend national borders, but which are also affecting people at the sub-national level. One particular group that has been enveloped in societal conflict is children. Children have long been considered a "special population" due to their vulnerability in society, but some of the biggest problems children face today are other children and it represents a new frontier for exploring societal conflict. The challenges children are facing run the gamut from public ostracism to sexual or physical assault to sexting and cyber bullying, and sometimes these actions are severe enough to be considered hate crimes that have devastating consequences. Many of these types of interactions begin at school, but with the ease and convenience of technology, the problems now easily follow children home and are having devastating consequence on this vulnerable population. This paper examines these issues from the perspective of public educators and parents of school-aged children throughout Oklahoma. The data presented herein focuses on information obtained from statewide surveys of parents and educators specifically designed to shed light on what issues these groups see children dealing with and how they try to help students cope with negative interactions between themselves and their peers. By exploring these problems from the perspectives of those who are responsible for this special population, maybe we can understand the problem better and learn how to prevent the growing levels of intolerance that are damaging children in the formative years of their lives. This is particularly important in Oklahoma, which has one of the highest sexual assault rates in the nation and is also known for one of the youngest victims of bullying that has brought international notoriety to the State.

How Recalcitrance Affects Data Quality.

The Relationship Between Response Propensity and Measurement Error in Survey Reports of Program Participation: An Exploration Using Linked National Health Interview Survey Data.

Renee Gindi, National Center for Health Statistics (RGindi@cdc.gov); James M Dahlhamer, National Center for Health Statistics (JDahlhamer@cdc.gov); Aaron Maitland, National Center for Health Statistics (AMaitland@cdc.gov)

The National Health Interview Survey (NHIS) is a complex, multi-stage household survey and a principal source of information on the health of the civilian noninstitutionalized population of the United States. Linkage of the NHIS with records collected by the Social Security Administration (SSA) is primarily intended to maximize the scientific value of survey data for NHIS users. Linked survey and administrative data may also be used to evaluate whether measurement error is present in key survey estimates. In this paper, we explore the extent of measurement error in NHIS estimates of income and insurance program participation, and whether measurement error varies with survey response propensity. Do efforts to minimize nonresponse bias through the recruitment of low response propensity households lead to greater measurement error?

Data from the 2005 NHIS were linked with SSA files containing records on Medicare; Social Security Old Age, Survivors, and Disability Insurance (OASDI); and Supplemental Security Income (SSI) program participation. We examined agreement between reported program participation and administrative record program participation in linked records. Survey response propensity was modeled using logistic regression; with paradata measures (e.g. concerns expressed by householders) included as key explanatory variables. Survey participants were then classified into three response propensity categories (high, medium, and low), and statistics on agreement between survey and administrative data were calculated for each category. In a preliminary analysis of more than 45,000 linked NHIS records, agreement was high for all three types of programs. However, statistics on agreement between survey and administrative data did not differ by response propensity category. Regression models explored the associations between response propensity and reporting discrepancies; findings varied by program type. These results suggest that while NHIS estimates are generally in concordance with SSA records, additional research may further improve survey measurement of program participation and benefits.

Do Reluctant Responders Provide Poor Data? Evidence From The Face-To-Face Recruited Internet Survey Platform (FFRISP).

Emily Gilbert, Institute for Social and Economic Research, University of Essex (eegilb@essex.ac.uk); Nick Allum, Department of Sociology, University of Essex (nallum@essex.ac.uk); Ana Villar, Institute for Research in the Social Sciences, Stanford University (villar@stanford.edu); Jon Kroshnick, Stanford University (kroshnick@stanford.edu)

Although many theories have been offered to explain the occurrence of measurement error in a multitude of situations and under different conditions, less attention has been paid to measurement error, caused particularly by satisficing, among difficult-to-recruit respondents. This is an important issue because much time and effort is expended on attempting to elicit responses from reluctant respondents in a bid to reduce non-response bias (Groves, 2006). In their push to achieve a high response rate, organizations spend an increasing amount of time and money, aiming to convert those sample members who initially refuse to participate. Yet this overlooks the possibility that reluctant respondents may not be providing data comparable to that imparted by more willing participants (Green, Kroshnick & Holbrook, 2001; Olson, Feng & Witt, 2008). As such, reducing non-response error may unintentionally increase total survey error.

This paper uses data from the The Face-to-Face Recruitment Internet Survey Platform (FFRISP) survey, a face to-face recruited internet survey, to investigate how satisficing is related to response propensity. A series of randomized split-sample experiments were embedded within the survey, which permitted estimates of the consequences of satisficing on measurement error to be made for several items in terms of response order effects, non-differentiation, acquiescence and 'no opinion'. Paradata were used to generate measures of recruitment difficulty, focusing particularly on contact attempts and refusals. Self-reported measures of satisficing were also gathered, where respondents were directly asked about perceived task difficulty and level of effort expended. Results indicate that reluctant responders provide

data of a quality that differs from easy-to-recruit participants but the effects are not as pronounced nor as straightforward as expected. In particular, difficult-to-recruit respondents appeared to nondifferentiate and acquiesce more. However, the propensity to answer in these ways differed depending on number of contact attempts and number of refusals in nonlinear ways.

Non-consent Error, Nonresponse Error, and Measurement Error: Assessing the Overall Quality of Linked Survey and Administrative Data.

Joseph W. Sakshaug, University of Michigan (joesaks@umich.edu)

Survey records are increasingly being linked to administrative databases to enhance the survey data and increase research opportunities for data users. A necessary prerequisite to performing exact record linkage is obtaining informed consent from respondents. Obtaining consent from all respondents is a difficult challenge and one that faces significant resistance due to societal concerns regarding privacy and data confidentiality. As a result, data linkage consent rates vary substantially from study-to-study and sometimes fall below response rates (Woolf et al., 2000; McCarthy et al., 1999). Several studies have found significant differences between consenters and non-consenters on socio-demographic characteristics, leading to concerns that inferences obtained from linked data sources may be biased (Kho et al., 2009). However, these studies are limited because they rely solely on survey data to assess consent biases. No study has attempted to determine whether systematic differences exist between consenters and non-consenters on key administrative outcomes. Estimating consent biases for administrative outcomes is complicated by the fact that administrative records are generally not available for the non-consenting portion of the sample. I estimate consent biases in the 2006/2007 German Labor Market and Social Security Survey (PASS). The PASS survey obtained an 80% consent rate to link survey records with federal employment and benefit reciprocity records. With permission from the German Institute for Employment Research (IAB), I obtained access to administrative records for both consenters and non-consenters and estimated consent biases for several administrative variables. This is the first study to estimate consent biases for linked administrative outcomes. In addition, I use call record data and variables common to both the survey and administrative data to estimate nonresponse and measurement error biases. This paper compares the relative contribution of each error source (consent, nonresponse, and measurement) to the overall error observed in the linked data set.

Identifying Causes of Verification Refusals on a Large Nation-Wide Field Study Using a Multilevel Model.

Dustin Williams, RTI (dwilliams@rti.org); Christina Touarti, RTI (ctouarti@rti.org); Christine Clark, RTI (chclark@rti.org); Jason Butler, RTI (jbutler@rti.org)

Verification data that are accurate and collected according to project protocols are an essential part of field studies. In particular, verification offers assurance to clients and the public that data are valid and reliable. On the National Survey on Drug Use and Health (NSDUH), respondents are asked to provide contact information so that project staff may call to check on the quality of completed work. Refusal of such information by respondents impedes the ability to verify said work or, at best, introduces delays and added expense to the process. Identifying the causal factors for the absence of verification contact data allows for remedial actions, thus reducing costs and increasing quality.

This paper will present the results of analyzed NSDUH data from all four quarters of 2009 and the first quarter of 2010. First conducted in 1971, NSDUH is sponsored by the Substance Abuse and Mental Health Services Administration (SAMHSA) and provides national, state and sub state data on substance

use and mental health in the civilian, noninstitutionalized population age 12 and older. Approximately 140,000 household screenings and 67,500 interviews are completed annually.

Questions persist as to whether high verification refusal rates are caused by interviewer performance and noncompliance with protocols, privacy concerns among respondents, protocols and forms needing improvement, or community characteristics. This paper will use multilevel modeling to further examine the effects of field interviewer performance measures and area demographic characteristics on collection of verification contact information. Field interviewer performance will be measured by production, cost, and data quality indicators. The demographic characteristics of sampling areas, or segments, will be based on Census data. We will discuss the effect each of these variables has on verification data in order to identify explanations for patterns in verification refusal rates.

Have You Really Seen This Ad? An Investigation of Measurement Error in Self-Reported Exposure Through Confirmed Awareness and Mode-Specific Cueing.

Peter Scholfield Hepburn, NORC (hepburn-peter@norc.org); Ting Yan, NORC (yan-ting@norc.org); A. Rupa Datta, NORC (datta-rupa@norc.org); W. Douglas Evans, George Washington University (sphwde@gwumc.edu)

The 2010 Census Integrated Communications Program Evaluation (CICPE) is a survey conducted by NORC at the University of Chicago to evaluate the effectiveness of the communications campaign launched by the Census Bureau to promote participation in the 2010 Census. Critical to the evaluation of the campaign is public awareness of and exposure to the campaign. The survey itself, therefore, hinges on respondents' self-reports of exposure to various elements of the campaign.

For the paid media component of the campaign, respondents were asked whether they recalled having seen or heard a particular widely-circulated Census advertisement. Such a measure of self-reported awareness is subject to measurement error: respondents may report awareness because of social desirability or selective attention bias or because they were confusing the target ad with other ads. To investigate the measurement error in self-reported exposure, we take advantage of two validation techniques. The first is the confirmed awareness technique, through which respondents' self-reported awareness could be confirmed as either true or false awareness. The second method uses mode-specific cueing to compare recognition of the ad to the self-reported awareness (which is based on recall). Both methods will help reveal the direction, size, and correlates of measurement error in the self-reported awareness measure. We will also examine which of the two methods is more effective in reducing measurement error in the self-reported awareness measure.

Media Effects: Framing & Priming.

Powerful Pictures: Examining the Effects of Visual Priming on Public Perceptions of Controversial Scientific Information.

Kajsa Dalrymple, University of Wisconsin - Madison (kdalrymple@wisc.edu)

Research has consistently shown that mass media play an important role in simplifying information in ways that make science issues easier to digest by the public. Specifically, media theory has pointed to the importance of journalistic priming and framing in presenting controversial and complex information to the public. For instance, media primes and frames can help reduce the complexity of certain issues and encourage audiences to connect scientific information to relevant public discourse. These cues can indirectly influence how individuals interpret media messages and form opinions about scientific issues.

Little research, however, has explored the possible relationships between the journalistic use of visual primes, individual's retention of information from news stories, and subsequent public opinion and discourse.

Using the topic of nanotechnology as a case study, we examine these relationships using experimental data collected at a large mid-western university (N=180). Specifically, subjects received one of five conditions and were assigned to read a news article about nanotechnology with one of four visual primes (relating to either: 1. consumer products, 2. scientific research, 3. technology, or 4. politics/policy) or were assigned to read an article with no visual. Our results indicate that subjects who were exposed to visual primes related to scientific or technological research were significantly less likely to retain information from the provided scientific news story, in comparison to the control group. These findings suggest that including scientific or technological visual primes with a controversial science news story may negatively influence an individual's understanding of that topic and subsequent public attitudes. This result supports previous research suggesting that the use of images in scientific news stories may, in fact, distract readers rather than aid them in their processing of controversial information. This demonstrates the importance of visual priming when conveying controversial information, which may subsequently have influence on public

Religious Media, Political Knowledge, and Attitude Polarization: Cheap Framing in Focus on the Family.

Colene J. Lind, University of Texas at Austin (lindc@mail.utexas.edu)

More people consume religious media than regularly attend church, and religious programming contains political information (Abelman, 1994). Consequently, religious media may be affecting political attitudes. Zaller (1992), as well as studies on partisan selective exposure (Stroud, 2006), found only partisans with relatively high levels of political knowledge were polarized by news-media exposure. On the other hand, Baum (2003) found political information presented via "cheap frames" sometimes reached low-knowledge audiences. As a possible new dimension of Zaller's theory of elite influence on public opinion, I argue religious media present political information via cheap frames, making their persuasive cues accessible to those who are relatively low in political knowledge. Therefore, this study tests whether religious-media audiences hold polarized attitudes irrespective of their political knowledge.

Data from a 1996 Pew Research Center for the People and the Press survey are analyzed (n = 1751). Religious-media use is measured via a survey question on frequency of listening to "religious radio shows such as Focus on the Family." Respondent attitude toward homosexuality serves as the measure of attitude polarization. Through regression analysis the study finds religious-media use predicts attitudes toward homosexuality. In contrast, the interaction of religious-media use and political knowledge is not significantly related to attitudes toward homosexuality. In comparison, the interaction of political knowledge and listening to Rush Limbaugh—a conservative source with few religious cues—is a slightly stronger and significant predictor of attitudes toward homosexuality.

Results offer tentative evidence that religious-media audiences are relatively polarized in their attitudes, even when low political knowledge might prevent them from understanding persuasive cues from other sources of political information. The present study cannot determine causation, but future investigation of political messages in religious media and how such messages are interpreted is warranted.

See No Evil, Hear No Evil: National Identity, Drone Warfare, and Culturally Resonant Frames.

Charles M Rowling, University of Washington (rowlingc@u.washington.edu); Penelope Sheets,

University of Washington (sheetsph@u.washington.edu); Timothy M Jones, Bellevue College (taojones@u.washington.edu)

This study involves an experiment in which we explore differential effects of frames related to the use of predator drones by the U.S. military in Afghanistan. Specifically, we exposed respondents to a news story about an incident in which a drone strike mistakenly killed 23 innocent civilians. Driven by a real-world event, we wanted to explore how official messages rationalizing the incident would be received by the public. As a first step, we constructed and tested three types of national identity-protective frames—minimization, disassociation, and reaffirmation—derived from previous scholarship on the framing of military incidents by White House and military officials (Rowling, Jones, & Sheets, in press). We then tested whether contestation of these frames by Congressional opponents affected how the frames were received by respondents. As Entman's (2003) Cascading Activation Model suggests, culturally resonant frames—such as those that protect the national identity—tend to cascade past contestation and into the public. However, minimal scholarship has examined this possibility experimentally. We, therefore, conducted a four-by-two experimental design. Respondents were randomly assigned to one of the three frames or all of the three frames combined. These frames were then either echoed or contested within the news story. Respondents' attitudes about the scope and responsibility for the incident, whether justice was served, and the overall impact of this incident on America's image were assessed. In particular, we examined whether the national identity frames shaped how the respondents came to understand the incident and to what extent contestation offered by rival officials undermined the power of these official frames. Our results have implications for the scholarly debate over framing effects as well as the dynamic interaction among political officials, the U.S. press, and the public. We hope this study will illuminate these processes in the context of current, controversial military policy.

War of words: Framing of the United States in Selected Belarusian Newspapers in 2009.
Natalie Manayeva, University of Tennessee (natalie.manaeva@gmail.com); Dzmitry Yuran, University of Tennessee (dyuran@utk.edu)

In the last decade America's global image has declined world-wide. In order to avoid or manage negative outcomes of the anti-American phenomenon it is important to understand its origins and mechanisms. The purpose of this study was to examine the peculiarities of framing of the United States in selected Belarusian newspapers during first six months of the Obama's Presidency. This study was focused on the two main questions: first, what is the difference between the U.S. image in Belarusian state-run and independent newspapers, and second, what is the mechanism of its creating. In order to provide comprehensive answers to both questions the multi-method approach (involving methods of content and framing analysis) was chosen.

This study demonstrated that the state-run and independent newspapers present a very different image of the U.S.: state-run newspapers present the U.S. within a scope of strong negative frames, while the independent newspapers mostly used neutral-positive frames. The results of the content analysis showed that negative images of the U.S. are not promoted through direct evaluating judgmental statements, but are rather initiated by means of selecting certain negative facts for publication. By concentrating attention on crime, natural catastrophes, manipulating statistical data, omitting sources of information and selecting foreign experts who are critically inclined against the U.S media create a negative image of the United States. These methods used by Belarusian media illustrate the evolution of the ways a negative image of the U.S is formed since Soviet propaganda times. As a result of this tactic a pseudo-objective image is being created, but it still remains strongly negative.

Secondary analysis of public opinion polls from Belarus reveals a strong correlation between the way mass media frame and present the U.S. and the way Belarusian public perceives the United States.

An Analysis of the Presentation of Research Data in News Media.

Maria Ivancin, American University (ivancin@american.edu)

Research is a powerful tool in helping decision makers to determine the policies our leaders support, what products will be sold, and even what medical treatments may be pursued. All these decisions affect citizens, but research presented in the news can actually have a direct impact on the individual. A political poll presented in the news may encourage someone to vote in a particular way, a medical study may encourage someone to change their lifestyle, and a product study may cause someone to buy a particular product. If the research information presented in the news media is not accurately portrayed – and in a manner that can be understood by the public – it can adversely affect individuals and our society as a whole.

Many of the public opinion polls and other research studies that are reported in the media on a regular basis are poorly designed, have an outright bias, or are not accurately presented by the media.

To understand the extent of the media's presentation of research, a content analysis was conducted on articles appearing in the top ten major newspapers over the course of a 12-month period. All the articles either reported on or cited research. Specifically, this study determines (1) the nature, scope, and quality of research presented in the media (types of studies presented; nature of the articles; source of the research, e.g. university, company, nonprofit); (2) the accuracy of the presentation (is methodology explained, sponsor disclosed, limitations discussed, whether conclusions are true to the findings, etc.).

Understanding the role research plays in our society, especially as it is presented in the news media, will allow us to use research more effectively and ensure that it is not used to confuse or misrepresent.

Questionnaire Design: Response Options and Order Effects.

Designing Effective Rating Scales in Customer Satisfaction Surveys.

Carey Stapleton, Service Management Group (cstapleton@smg.com)

There has been considerable debate within the market research and general survey research fields about survey scaling and which approach is best suited to capture respondent attitudes and opinions. Many of these debates have focused on the appropriate number of scale points, whether each scale point should be labeled and what type of verbal label proves most effective for each point.

In order to understand what scale creates the highest possible data quality in customer satisfaction research, a test was developed with four test cells:

1. 5-Point Fully Labeled Satisfaction Scale
2. 5-Point Fully Labeled Agree Scale
3. 5-Point Anchor Only Labeled Satisfaction Scale
4. 10-Point Anchor Only Labeled Satisfaction Scale

Working with a panel provider, a sample of 750 completed surveys was collected for each of the four test cells. Questions included customer satisfaction questions where the respondent was asked to rate a recent shopping experience, covering a variety of different aspects, with a major retail brand.

Results will be presented comparing acquiescence bias, straightlining, response distributions, predictive ability, factor loadings, and inter-item correlations across the four different scales. Based on these findings, recommendations will be made for satisfaction surveys about the optimal number of scale points, labeling scale options, and what the verbal labels should be.

Measuring perceptions of probabilities: Verbal or numerical response options?

Yphtach Lelkes, Stanford University (ylelkes@stanford.edu); Jon Krosnick, Stanford University (Krosnick@Stanford.edu)

Researchers across disciplines often attempt to measure respondents' perceptions of probabilities. For example, respondents may be asked to estimate the likelihood that they will vote in an upcoming election, or to estimate the percent chance that they will lose their job in the next year. However, the measurement of probability questions has been inconsistent: Participants are sometimes asked to respond using verbal response options, such as, "very likely" and "very unlikely," and other times asked to respond using numerical response options, such as "0%" or "100%." Since data quality is contingent on the validity of the measures, guidelines are needed so that researchers write the best possible question to gauge perceptions of probabilities. As a step in that direction, the concurrent validity of a question that utilized verbal response options was compared to that of the same question framed in numerical form. Respondents (N=1203) were asked, as part of the 1998 ANES pilot study, to estimate their perceived probability of voting in upcoming elections and were randomly assigned to receive the question and response either in verbal or numerical form. The association between question form and a number of correlates of vote turnout were compared between groups. In line with our predictions, the association between criterion variables and the numerical form of the question was consistently stronger than that of the verbal form of the question. Surprisingly, this effect was not stronger among the more educated, whom we expected to be more apt at using percentages, than the lower educated.

Re-Examining the Effects of Grouping the Options in Factual Questions.

Nuttirudee Charoenruk, Program of Survey Methodology, University of Michigan (nuttirud@umich.edu); Roger Tourangeau, the Joint Program in Survey Methodology, University of Maryland (rtourangeau@survey.umd.edu); Cleo Redline, the Joint Program in Survey Methodology, University of Maryland (credline@survey.umd.edu)

This paper builds upon previous work that examined the effects of grouping response options in factual questions with many response categories (e.g., items on educational attainment). In the earlier research, the response options were presented in a single column or in three columns. The response options were also grouped by content with a heading given for each group (e.g. "Grade school or less;" "High school;" and "College and above"). We found that grouping long lists of response options and providing headings confused respondents more than it helped them and significantly increased response times. This paper presents the results of a follow-up experiment that examined how response times are affected by the formatting of the headings (comparing bold, underlined, gray, and italicized headings); these formats were tested against a control question that contained no grouping of the options. The results of this experiment are in line with our previous findings. Again, we found that headings significantly increased the response times. The formatting of the headings did not seem to affect the level of confusion. In addition to slower times, respondents often changed their answers (clicked more than once) in the presence of headings. With the headings, some respondents inferred they were supposed to select one answer from each group of response options. The control condition was found to confuse respondents the least and to produce the fastest response times. This suggests that minimal formatting may significantly reduce cognitive burden.

Question and Respondent Predictors of Response Latencies to Survey Questions Measuring Opinions, Personal Characteristics, and Factual Knowledge.

Allyson L Holbrook, Survey Research Laboratory, University of Illinois at Chicago (allyson@uic.edu); Timothy P Johnson, Survey Research Laboratory, University of Illinois at Chicago (timj@srl.uic.edu); Young Ik Cho, Survey Research Laboratory, University of Illinois at

Chicago (youngc@srl.uic.edu); Sharon Shavitt, University of Illinois Urbana-Champaign (shavitt@illinois.edu); Noel Chavez, University of Illinois at Chicago (nchavez@uic.edu); Saul Weiner, University of Illinois at Chicago (sweiner@uic.edu)

Response latencies (e.g. a measure of how long it takes a respondent to answer a question or make a judgment) have been widely used in psychological research and have more recently been used in a small number of surveys as an indicator of the underlying process involved in answering survey questions. One of the challenges of using response latencies in evaluating survey questions is that their theoretical relationship to data quality is not straightforward. For example, one could imagine that some processes that introduce error into survey responses (e.g., satisficing) would likely lead to faster response latencies whereas other processes that introduce error (e.g., difficulty comprehending a question or choosing to give a socially desirable response) would likely lead to slower response latencies. To gain greater insight into the predictors of response latencies, we use data from a survey in which approximately 400 respondents participated in an hour-long CAPI interview. Response latencies for approximately 100 questions measuring opinions, beliefs, personal characteristics and behaviors, and factual knowledge were measured. Respondent-level (e.g., education, race/ethnicity, individual differences) and question-level characteristics (e.g., type of judgment, response format) are used to predict response latencies to identify factors that influence response latencies in surveys.

Not Getting it Right: Incorrect Answers vs. “I Don’t Know” in Knowledge Surveys.

Gregory A. Smith, Pew Research Center (gsmith@pewforum.org); Rob Suls, Pew Research Center (rsuls@pewresearch.org); Mike Mokrzycki, Independent Consultant (mike@mikemokr.com); Jocelyn Kiley, Pew Research Center (jkiley@pewresearch.org); Scott Keeter, Pew Research Center (skeeter@pewresearch.org)

Having a sense of how much (or how little) the public knows about a variety of topics is crucial for understanding modern society, and especially for public opinion researchers. As such, many surveys in recent years have sought to gauge the public’s knowledge in a variety of domains, including politics, current events, science, and religion. Analyzing the patterns and statistical correlates of correct answers to knowledge questions is a relatively straightforward process. But what are scholars to make of the large number of people who do not answer knowledge questions correctly?

This paper seeks to better understand patterns in failures to provide correct answers in knowledge surveys. We are interested in respondents who provide each of two different types of answers – providing an answer that is factually incorrect, and answering “I don’t know” in response to knowledge questions. Drawing on a variety of Pew Research Center knowledge surveys conducted in recent years (including the Center’s bi-annual political knowledge tests, a 2010 religious knowledge survey and a 2009 poll of scientific knowledge) we investigate a variety of specific questions: What kinds of respondents tend to offer incorrect responses, and what kinds of respondents simply say they don’t know the answer to a question? Is there variance across domains of knowledge in the characteristics of respondents who offer incorrect answers and those who say “don’t know?” How well can interviewers estimate the frequency with which respondents are guessing in response to knowledge questions? And which kinds of respondents – those who offer incorrect answers or those willing to admit they don’t know answers to questions – are more likely to break-off of a knowledge survey? Providing answers to these questions will help researchers better understand the results and the real-world implications of knowledge surveys.

Religious Identification and Opinion Formation.

Evolving a Better Measure: Developing Alternative Measures of Beliefs in Human Origins.

William Bryan Higgins, ICF International (bhiggins@icfi.com); Randall K. Thomas, ICF International (randall.k.thomas@gmail.com); George Bishop, University of Cincinnati (bishopgf@ucmail.uc.edu)

Beliefs about the origins of human life are complex and nuanced. Gallup has assessed public belief concerning human origins using 3 possible beliefs – God-guided evolution, evolution, and creationism. In a series of web-based studies we compare the Gallup measure to a 4 belief measure that distinguished between the beliefs of God-guided evolution (God both created and guided evolution), God-originated evolution (God created but evolution occurred naturally), evolution, and creationism. Along with comparing how responses to the Gallup measure mapped out against the 4 category measure, we also examined the association of their responses with religious experience and beliefs, along with their attitudes toward science and religion. In addition, we examined how locus of control related to the various beliefs. Study 2 also was a web-based survey and compared Gallup's 3 belief measure with an even more detailed 5 belief measure (rating the likelihood of each being true), also comparing responses with attitudes toward science and religion. We replicated the relationship between human origin beliefs and locus of control and need for cognition.

Religious Inter marriage and Its Relationship to Religious Commitment and Politics.

Scott Clement, Pew Research Center - Forum on Religion and Public Life (sclement@pewforum.org); Christine Brickman Bhutta, Pew Research Center - Forum on Religion and Public Life (chrisbhutta@yahoo.com); Greg Smith, Pew Research Center - Forum on Religion and Public Life (gsmith@pewforum.org); Allison Pond, Pew Research Center - Forum on Religion and Public Life (apond@pewforum.org)

An abundance of research documents the strong connection between public opinion, political behavior and religious commitment. Other research shows that religious commitment is itself influenced by marriage and family structure, with getting married and starting a family linked with increases in religious commitment. However, there has been less focus on how religious inter marriage, which rose steadily throughout the twentieth century, relates to these attitudes and behaviors. In this paper, we first analyze the relationship between inter marriage and religious commitment. Using data from the 2007 Pew Religious Landscape Survey (RLS), we show that religious inter marriage is generally linked with lower levels of religious commitment, though the degree to which this is true depends on the respective religious backgrounds of spouses. We supplement these findings by analyzing panel data and find a causal relationship, with inter marriage leading to lower levels of commitment. We then use the RLS to examine the extent to which type of marital union influences political attitudes and behaviors. Relative to being in a same-faith union, inter marriage to a religiously unaffiliated person increases the odds that a Protestant or Catholic is politically liberal. Conversely, inter marriage to a Protestant or Catholic increases the odds that an unaffiliated respondent is politically conservative. We also find that any type of inter marriage decreases the odds that a religiously-affiliated individual voted in the 2004 election. However, inter marriage has no effect on the voting behavior of the unaffiliated. We hypothesize possible explanations for these results and offer suggestions for future work.

Catholics and Confidence in Religious Institutions.

Ann Arthur, Gallup Research Center, University of Nebraska-Lincoln (ann.arthur@huskers.unl.edu); Tarek Albaghal, Gallup Research Center, University of Nebraska-Lincoln (talbaghal@gmail.com); Lei Huang, Gallup Research Center, University of Nebraska-Lincoln (leihuang03@hotmail.com); Ana Lucia Cordova Cazar, Gallup Research Center, University of Nebraska-Lincoln (al.cordovacazar@huskers.unl.edu); Lauren Walton, Gallup Research Center,

University of Nebraska-Lincoln (lauren.walton8008@gmail.com); Meryem Ay, Gallup Research Center, University of Nebraska-Lincoln (meryemay@huskers.unl.edu); Allan L. McCutcheon, Gallup Research Center, University of Nebraska-Lincoln (amccutch@unlserve.unl.edu)

How many Catholics express confidence in the Catholic Church? Using data from the Gallup World Poll, this research examines the confidence of Catholics in religious institutions from three perspectives. First, are Catholics in countries where Catholicism is the religious minority more confident in the Catholic Church than Catholics in countries where Catholicism is the religious majority? Previous research has indicated that Catholics in Poland, a country where Catholicism is a religious majority, are much more likely to break with the teachings of the Catholic Church than in the United States, a country where Catholicism is a religious minority (McCutcheon and Nawojczyk, 1995). Second, has this confidence changed since 2005 and 2009, in the wake of a Papal election and numerous clergy sex abuse scandals? Finally, Catholics' confidence in religious institutions will be compared to that of other religious groups.

Comparative analyses will examine the nature of confidence in religious institutions on self-identified Catholics. Specifically, we will focus on Catholics in the United States and Germany, where Catholics comprise the religious minority, as well as in Mexico, Poland, and Italy, where Catholics comprise the religious majority.

Preliminary analyses indicate differences across countries, with American Catholics being the most confident (83.6%), and European Catholics being significantly less confident (averaging 65.7%). Across time, there are some apparent changes, but is unclear whether this is due to true change or it is an artifact of the data collection. Additionally, contrary to expectation, Catholics in all of the examined countries appear to be more confident in religious institutions than other religious groups in these same nations.

“He Who Lies With Another Man as With a Woman”: Attitudes Towards Homosexual Sex in Australia.

C. G. E. Kelley, Yale University (claire1053@gmail.com); S. M. C. Kelley, Yale University (sarahkelley1759@gmail.com)

Homosexuality continues to be a highly contentious social issue, with debates informed particularly by religion and education. Widely publicized conflicts surrounding gay marriage, homosexuals in the military, and adoption by gay couples have further polarized the public. We seek to understand some of the sources of these attitudes and their origin in religion and social structure through a quantitative analysis of data from the sexuality module of the 2004 International Social Science Survey/Australia, a representative national sample of Australian adults. N=1533. The main question is part of a battery of questions measuring various areas of sexual permissiveness, including attitudes towards premarital heterosexual sex, casual sex, and homosexual sex. The question asks “What do you think about sexual intercourse between two men?” And respondents are asked to answer on a scale from 1 (absolutely wrong) to 7 (absolutely right). The same question is asked about sex between two women. We combined these two very highly correlated questions into a single scale. Independent variables include two reliable multi-variable scales measuring general heterosexual permissiveness and strength of religious belief, as well as standard demographic and background variables. We found that heterosexual permissiveness was the single biggest predictor of acceptance of homosexual sex. In addition, religion and church attendance also have strong effects but operate entirely through influencing respondents' general sexual permissiveness. To gain more insight into how religion affects these attitudes, I categorized denominations based on official church position towards homosexuality as well; it has a modest indirect effect. It is especially striking that most demographic background variables, as well as religious variables, influence attitudes towards homosexuality almost entirely indirectly through their influence on permissiveness towards heterosexual sex. Views on homosexuality are part of a broad social conflict which we may understand better by analyzing attitudes towards homosexual sex.

Response & Nonresponse Issues in Multi-Mode Surveys.

More Positive or More Extreme? A Meta-Analysis of Mode Differences in Response Choice.

Cong Ye, University of Maryland (eyecong@msn.com); Jenna Fulton, University of Maryland (jannefulton@gmail.com); Roger Tourangeau, University of Maryland (rtourangeau@survey.umd.edu)

Some researchers have argued that respondents give more extreme answers to questions involving response scales over the telephone than in other modes of data collection, but others have argued that telephone respondents give more positive answers. We conducted a meta-analysis based on 18 experimental comparisons between telephone interviews and another mode of data collection. Our analysis showed that telephone respondents are significantly more likely than respondents in other modes to give extremely positive answers (for example, the highest satisfaction ratings in a customer satisfaction survey) but are not more likely to give extremely negative responses. This tendency to give highly positive ratings appears to be related to the presence of an interviewer and it may reflect respondents' reluctance to express bad news, a tendency some social psychologists have dubbed the MUM effect.

Examining Response Rates and Patterns in a Multimode Experiment: A Study of Department Chairs/Heads in STEM Programs at Research Intensive Universities.

Kelly N. Foster, University of Georgia (knfoster@uga.edu); Monica Gaughan, University of Georgia (monica.gaughan@gmail.com)

The expansion of the Internet, email, and virtual worlds has lead to new avenues for data collection. However, virtual data collection methods have left many researchers wondering if these avenues are an adequate substitute for the historically reliable methods of collecting survey data such as telephone interviewing or paper surveys. Virtual data collection can often be done at a fraction of the cost of conducting similar data collection via traditional methods (Dean et al, 2009) thereby becoming an attractive option for quick, cost-efficient data collection. However, the extent to which these virtual methods impact question response patterns and response rates compared to traditional modes is uncertain and often varies by population.

Giving survey respondents the choice of mode often results in higher response rates than by a single mode alone. This is not surprising, particularly when examining response rates for online data collection, given the technological accessibility and capability of the general population. However, would this be the case with a population that not only has access to high speed Internet but who is also a frequent consumer of virtual technologies such as the World Wide Web and email?

The authors seek to explore this area by examining the response rates and impact of differing mode of survey among 776 University Department Heads/Chairs - a highly educated, professional population with virtually 100% Internet accessibility. Respondents were randomly placed in one of three strata: paper only survey, web only survey, or mixed mode. Despite the generally high level of accessibility and use of the Internet, there are significantly different response rates and response patterns across the 3 groups. It is the authors' intent to provide an overview of the study experiment, discuss the potential impact of the differing response rates and patterns, and to discuss future research with this population.

Mode Differences in a Mixed-Mode ABS Design.

Rene Bautista, NORC (bautista-rene@norc.org); Ting Yan, NORC (yan-ting@norc.org); A. Rupa Datta, NORC (datta-rupa@norc.org); Peter Scholfield Hepburn, NORC (hepburn-peter@norc.org)

Surveys employing an Address-Based Sampling (ABS) design have the luxury of contacting and recruiting respondents via multiple modes of administration. Given the increasing popularity of ABS surveys, it is time to start research on potential mode differences in a mixed-mode ABS design. Empirical questions regarding what types of respondents are recruited by which modes are not as yet answered. This paper makes use of data from the 2010 Census Integrated Communications Program Evaluation (CICPE) to examine the differences in demographic characteristics and cooperativeness of respondents recruited by different modes.

The CICPE was designed as a mixed-mode study using an ABS sample. Selected sample addresses with matched phone numbers were initially fielded as telephone cases; nonparticipants in the phone survey and sample without matched phone numbers were approached to complete an in-person interview. In later waves, panel respondents were offered self-administered paper questionnaire (SAQ) and web completion options in addition to the telephone and in-person options. This mixed-mode design provides a great opportunity to study mode differences. We will compare distributions of demographic characteristics and measures of cooperativeness by modes of interview (telephone, in-person, SAQ/web) to see if there were indeed differences by mode. In particular, CIPCE oversampled certain rare subpopulations (such as Native Hawaiian and Other Pacific Islander and American Indians and Alaska

Natives). We will examine the extent to which these rare subpopulations were recruited by different modes.

Pay to Play: An Incentive Experiment.

Nicholas Redel, Mathematica Policy Research (nredel@mathematica-mpr.com)

Authors: Nicholas Redel, Laura Kalb, Nancy Cole, Mark Denbaly and John Kirlin.

In an effort to better understand the food choices of U.S. households, the Economic Research Service (ERS), U.S. Department of Agriculture (USDA), plans to collect detailed food acquisition data from a nationally representative sample of 5,000 households. Respondents will be asked to undergo a 60-minute training, record seven days of household food acquisition data, and make three 15-minute telephone calls to report household food acquisitions. Respondents will also complete three 30-minute interviews (one CATI and two CAPI) about household characteristics. These data collections impose a significant burden on households (over three hours of data collection and training), which will be offset with a monetary incentive. In order to assess the optimal incentive for this study, a Field Test with 400 completed interviews will be conducted between January and March 2011. During this test, households will be randomly assigned to one of two incentive levels (low or high). Both incentive levels include a base amount for the primary respondent (\$50 and \$100 respectively), a supplement for each additional household member who provides information about their food acquisitions (\$10 or \$20 depending on the age of the household member), and a “telephone bonus” (up to \$30) to be paid if the primary respondent initiates the telephone reporting of food acquisitions. The telephone bonus is designed to reduce overall data collection costs since calls initiated by respondents are completed at significantly lower cost than outgoing calls. This paper discusses implementation of the dual incentive scheme and preliminary results from the incentive experiment.

The Impact of a Mixed-Mode Data Collection Design on Response and Non-Response Bias on a RDD Landline Telephone Survey.

Sean Hu, Centers for Disease Control and Prevention (shu@cdc.gov); Carol Pierannunzi, Centers for Disease Control and Prevention (cPierannunzi@cdc.gov); Lina Balluz, Centers for Disease Control and Prevention (lballuz@cdc.gov)

Like other RDD telephone surveys, response rates in the Behavioral Risk Factor Surveillance System (BRFSS) have declined steadily since the mid-1990s. Nonresponse bias has become a considerable concern in all telephone surveys conducted in the U.S. Although response rates do not directly measure the degree of nonresponse bias in the risk factor and health condition estimates, the demographic and socioeconomic distributions of the design-weighted BRFSS state samples can differ substantially from external sources such as Claritas, the Current Population Survey, and the American Community Survey. Because some risk factors and health conditions are highly correlated with these demographic and socioeconomic characteristics, there is a potential for nonresponse bias in the estimates. In order to improve state-level BRFSS response rates and to assess nonresponse bias, BRFSS is using a mixed-mode data collection design to achieve these goals. A mail follow-up survey to non-respondents of the RDD landline telephone survey is implemented at 6 states in 2010. The telephone numbers of non-respondents are reverse-matched against a commercial database of residential addresses. The telephone numbers with an address match are mailed the questionnaire. We demonstrate how combining data from the mail follow-up survey with those from landline samples in the ongoing BRFSS can help to reduce the potential for bias in key survey estimates of health conditions and risk behaviors. Furthermore, we used the expression relative bias to estimate the non-response bias in landline telephone surveys.

Sampling Frame Methods – Geographic, Network and within Households.

Eliminating Invisible Boundary Problems in Area Samples, with an Application to the National Children's Study (NCS).

Colm O'Muircheartaigh, Harris School- University of Chicago (colm@uchicago.edu); Ned English, NORC at the University of Chicago (english-ned@norc.org)

We present a method for eliminating the problem of invisible boundaries in segment definition while maintaining the appropriate probabilities of selection. We apply the method to the sample design of the National Children's Study (or NCS). The ultimate area units in probability samples are typically small areas, known as 'segments'. The boundaries of these units must be defined for data collection purposes. Segments are often groups of Census blocks or block groups; census blocks themselves are defined by a variety of features, including both visible (e.g., roads, rivers) and invisible (e.g., municipal) boundaries. At issue is the impact of invisible boundaries, which can challenge field staff in numerous practical ways, including simply determining the extent of their workloads. Such vague boundaries can potentially result in the erroneous inclusion or exclusion of households from selected segments, creating under- or over coverage. The challenge is how to re-define the boundaries post-hoc without adversely affecting the selection probabilities. We propose and apply a rule-based method that maintains the integrity of the sample at minimal cost.

Failure of the Half-Open Interval Missed Housing Unit Procedure.

Stephanie Eckman, Institute for Employment Research (steph.eckman@gmail.com); Colm O'Muircheartaigh, University of Chicago (colm@uchicago.edu)

The half-open interval procedure seems to offer an inexpensive method of reducing undercoverage in housing unit frames during data collection: units missing from the frame are linked to covered units and given a chance of selection. This paper provides details on how the procedure should work, and then discusses problems both in principle and in practice. Reporting the results of an experiment in three U.S. cities, we show that the procedure often fails to reduce undercoverage and can introduce overcoverage. On the basis of this evidence, we conclude that the half-open interval procedure should not be considered a cure for undercoverage and should be used only with great care and training.

A Field Validation Study of Frame Construction for the National Children's Study: Enhancing the Delivery Sequence File.

Ned English, NORC at the University of Chicago (english-ned@norc.org); Colm O'Muircheartaigh, Harris School- University of Chicago (colm@uchicago.edu); Katie Dekker, NORC at the University of Chicago (dekker-katie@norc.org)

This paper presents a field evaluation of alternative frame construction methods using the National Children's Study (NCS). The NCS is based on an equal-probability sample of addresses within selected segments. Decisions were made prior to data collection as to the method of address frame construction that would be most appropriate for each segment. Segments were either traditionally-listed, employed the USPS delivery-sequence file (or "DSF") as it was, or used a version of the DSF that was edited during field checking. It is this last approach, known as "enhanced listing", that was most commonly employed in the NORC segments. During enhanced listing field staff add any missing addresses, delete those not present, and otherwise edit the "raw" DSF in order to maximize coverage, while still realize efficiency benefits from an a-priori list. At question is what coverage and efficiency benefits were realized through enhanced listing in comparison with the traditional or "raw" DSF approaches otherwise employed. We

would expect improved coverage if households were added through field checking that would not have been present on the original DSF. Similarly, we may anticipate efficiency gains by removing addresses from the list that were falsely included or not housing units. Our analysis compares the qualities of households that were present on the original DSF vs. those that were added by enhanced listing in order to understand any coverage advantages of list enhancement. We do so by first describing the types of housing units contained in either category. Secondly, we examine eligibility rates of each type of household for the National Children's Study. Thirdly, we compare member-level characteristics of both categories of households.

Comparing the Selection of One Person per Household to the Selection of All Household Members: Can Less be More?

Vincent G. Iannacchione, RTI International (vince@rti.org); Bonnie E. Shook-Sa, RTI International (bshooksa@rti.org)

The decision to select one or all eligible members of a sampled household (HH) is influenced by a number of factors including burden on the HH, cost, and the sampling variance of survey estimates. Design effects quantify the influence of a complex sampling design on the variance of survey estimates. Restricting a sample to one eligible person per sampled HH can have counteracting impacts on design effects. A one-person per HH sample requires subsampling in multi-person HHs which increases the design effects attributable to unequal weighting. Conversely, selecting one person from each sampled HH could reduce the design effects attributable to clustering because the potential intra-household correlation among respondents in the same HH is avoided. If this reduction is larger than the increase caused by subsampling, a one-person per HH sample can achieve the same sampling variance as a multi-person sample with less cost and burden.

We present the results of a simulation study that evaluated the design effects associated with the selection of one person per HH on personal victimization rates based on the 2008 National Crime Victimization Survey (NCVS) which currently selects all persons 12 and older for the survey. We selected replicate samples of one respondent from each HH from the 2008 NCVS public-use database and then compared the design effects associated with victimization rates for a one-person per HH sample to those of the existing multi-person sample. We found that the increase in design effects caused by unequal weighting associated with a one-person sample was significantly greater than the decrease caused by the elimination of intra-household correlation. We discuss the pros and cons of a one-person per HH design, and estimate the number of HHs that would be needed to equalize the sampling variances of the current multi-person sample design and a one-person design.

Cognitive Testing Results of Personal Network Size Questions: Implications for Respondent Driven Sampling Weights.

HarmoniJoie Noel, National Center for Health Statistics (hnoel@cdc.gov); Heather Ridolfo, National Center for Health Statistics (hridolfo@cdc.gov)

Respondent driven sampling (RDS) is a chain-referral sampling strategy similar to snowball sampling that is often used for hard-to-reach populations. To produce population estimates RDS data is weighted by participants' personal network size to reduce some of the biases associated with chain-referral sampling. The quality of the weights is dependent on how accurately personal network size is measured. Yet, standardized measures of personal network size do not exist. This study will provide the cognitive testing results of several questions used to measure personal network size from the National HIV Behavioral Surveillance (NHBS) System. NHBS uses respondent driven sampling to conduct surveys among injecting drug users and heterosexuals at increased risk of HIV. The goal of this project is to understand the process respondents go through when answering personal network questions. In particular, do

respondents interpret these questions in the same way? Do respondents count people in their networks in the same way? Who do respondents include in their network? Do the network questions accurately reflect how respondents think of their relationships? To answer these questions cognitive interviews will be conducted with injecting drug users and heterosexuals at increased risk of HIV. Respondent narratives will be analyzed to better understand how accurately personal network size is measured in the NHBS.