60TH AAPOR ANNUAL CONFERENCE

FINAL PROGRAM

MAY 12-15, 2005

FONTAINEBLEAU HILTON RESORT

MIAMI BEACH, FLORIDA

AAPOR 2005

American Association for Public Opinion Research
60th Annual Conference

Improving Survey Quality

Miami Beach, Florida
May 12-15, 2005

WELCOME to the ...

American Association for Public Opinion Research
60th Annual Conference, May 12-15, 2005

Fontainebleau Hilton Resort
4441 Collins Ave.
Miami, Florida 33140
Phone: 305-538-2000 Fax: 305-673-5351

TABLE OF CONTENTS

Registration/Information Desk

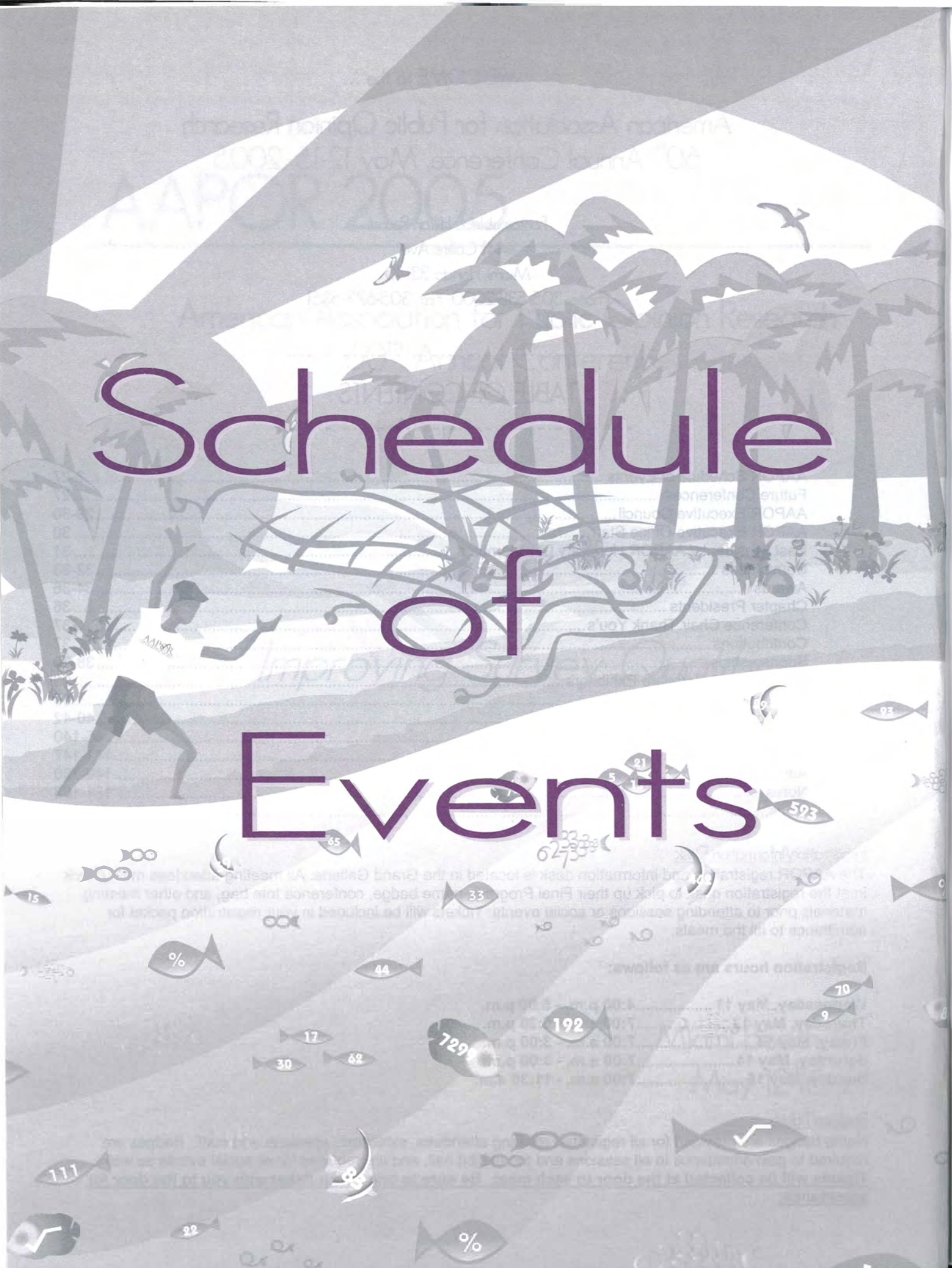
The AAPOR registration and information desk is located in the Grand Gallerie. All meeting attendees must check in at the registration desk to pick up their Final Program, name badge, conference tote bag, and other meeting materials prior to attending sessions or social events. Tickets will be included in your registration packet for admittance to all the meals.

Registration hours are as follows:

Wednesday, May 11 .....4:00 p.m. - 8:00 p.m.
Thursday, May 12 .....7:00 a.m. - 6:30 p.m.
Friday, May 13 .....7:00 a.m. - 3:00 p.m.
Saturday, May 14.....7:00 a.m. - 3:00 p.m.
Sunday, May 15 .....7:00 a.m. - 11:30 a.m.

Badges/Tickets

Name badges are provided for all registered meeting attendees, exhibitors, speakers and staff. Badges are required to gain admittance to all sessions and the exhibit hall, and are required for all social events as well. **Tickets will be collected at the door to each meal. Be sure to bring your ticket with you to the door for admittance.**



Schedule of Events

AAPOR 60TH ANNUAL CONFERENCE

SCHEDULE OF EVENTS

WEDNESDAY, MAY 11

4:00 p.m. - 8:00 p.m. AAPOR REGISTRATION DESK OPEN - Grand Gallerie

THURSDAY, MAY 12

7:00 a.m. – 6:30 p.m. AAPOR REGISTRATION DESK OPEN - Grand Gallerie

7:00 a.m. AAPOR Golf at Miami Shores – Sponsored by Gallup

AAPOR SHORT COURSES
Thursday - 8:30 a.m. – Noon

Doing Contingent Valuation in Surveys - Imperial I
Robert Cameron Mitchell

Thursday - 8:30 a.m. – 12:30 p.m.

Survey Translation: Issues and Procedures - Imperial II
Janet Harkness & Alisú Schoua-Glusberg

Break Sponsored by: Survey Sampling International

Noon - 5:00 p.m. AAPOR EXECUTIVE COUNCIL MEETING/LUNCH - Lafayette

AAPOR SHORT COURSES
Thursday - 1:30 p.m. – 5:00 p.m.

Latent Class Analysis - Imperial I
Allan L. McCutcheon

Designing Scales for Survey Questionnaires - Imperial II
Jon Krosnick

AAPOR CONCURRENT SESSION
Thursday - 3:30 p.m. – 5:00 p.m.

Weighting - Burgundy
Chair: Doug Usher, Mellman Group
Discussant: Trent Buskirk, East Virginia Medical School

Propensity Models Versus Weighting Cell Approaches to Nonresponse Adjustment: A Methodological Comparison
Peter Siegel, RTI International
James Chromy, RTI International
Elizabeth Copello, RTI International

Dealing with Wide Weight Variation in Polls
Richard Griffin, U.S. Census Bureau

Party ID Weighting - Experiments to Improve Survey Quality
Olena Kaminska, Univ. of Connecticut
Christopher Barnes, Univ. of Connecticut

Real-Time Propensity Models for Responsive Survey Design and Post-Survey Adjustment Through Propensity Models: Comparisons of Model Fit and Model Specification
Robert M. Groves, University of Michigan
Jim Lepkowski, University of Michigan
John VanHoewyk, University of Michigan
Paul Schulz, University of Michigan

Scaling and Response Scales – Imperial III
Chair: Larry Lushkin, ORC Macro
Discussant: Bosah Ebo, Rider University

Survey Question Length and the Reliability of Measurement
Duane Alwin, Pennsylvania State University
Ryan McCammon, University of Michigan

Rating versus Comparative Trade-off Measures: Trending Changes in Political Issues across Time and Predictive Validity
Randall Thomas, Harris Interactive
Alyssa Johnson, Harris Interactive
Susan Behnke, Harris Interactive
Mitchell Sanders, Harris Interactive

Measurement and Scaling Methods Independent of Response Style
Jon Pinnell, MarketVision Research
Lisa Fridley, MarketVision Research

Interviewer Race and Training - Brittany
Chair: Janice Ballou, Mathematica Policy Research, Inc.
Discussant: Todd Rockwood, Division of Health Services Research

Race-Matching: Interviewers' Reactions to the Race-Matching Process
Cindy Boland-Perez, Ualr- Iog Survey Research Center
P. Denise Cobb, Siue
Patricia Lebaron, Ualr/Iog Survey Research Center

Respondent Views on Perceived Racial Context and Race of Interviewer Effects
David Wilson, Gallup/Michigan State University
Jack Ludwig, Gallup

Out of the Lab and Into the Field: Developing a Protocol for Responding to Diversified Recruitment Needs
Karen Whitaker, National Center for Health Statistics
Miller Kristen, National Center for Health Statistics
Willson Stephanie, National Center for Health Statistics

An Inquiry into the Efficacy of a Short Training for Interviewers
Claire Durand, Université de Montréal
Marie-Eve Gagnon, Université de Montréal
Christine Doucet, Université de Montréal
Eric Lacourse, Université de Montréal

Morality, Religion and Politics - Champagne
Chair: Don Camburn, Research Triangle Institute
Discussant: Sid Groeneman, Groeneman Research & Consulting

The Basis of Voting Preferences Among Evangelical Christians in the 2004 Election
Kenneth Blake, Middle Tennesee State University
Robert Wyatt, Middle Tennessee State University
Holly Warf, Middle Tennessee State University

Understanding the Catholic Vote in 2004
Kate Stewart, Belden Russonello & Stewart
John Russonello, Belden Russonello & Stewart
Rachel Sternfeld, Belden Russonello & Stewart

The Impact of Attitudes toward Stem Cell Research on the Vote for President
Jon Miller, Northwestern University
Rafael Pardo, Foundation BBVA
Linda Kimmel, Northwestern University

Surveying Racial Attitudes in America: New Data Sources and Selected Results - Lemans/Bordeaux
Chair: Devon Johnson, Harvard University
Discussant: Lawrence Bobo, Stanford University

Understanding Perceptions of Racial Discrimination
Stanley Feldman, SUNY
Leonie Huddy, SUNY

The National Politics Study: Ethnic Pluralism and Politics in the 21st Century
Vincent Hutchings, University of Michigan
Cara Wong, University of Michigan
Ron E. Brown, Wayne State University
James S. Jackson, University of Michigan

A Multi-Ethnic Measure of Racial Residential Preferences: New Results from Chicago and Detroit
Maria Krysan, University of Illinois at Chicago
Reynolds Farley, University of Michigan
Mick Couper, University of Michigan

On Race and Orientation of Juvenile Court Workers: Perspective on the Difference that Difference Might Make
Geoff Ward, Northeastern University

5:15 p.m. – 6:15 p.m. NEW COMERS'/ALL CHAPTER RECEPTION – Club Atlantic

First-time AAPOR conference attendees are invited to get to know long-time AAPOR members and all AAPOR members are encouraged to learn more about AAPOR's regional chapters.

Hors d'oeuvres sponsored by: SPSS

New Member Drink Ticket sponsored by: Greenfield Online

6:15 p.m. – 7:45 p.m. PLATED DINNER (Core Meal) Fontaine/Fleur Delis

Wine sponsored by: Knowledge Networks

7:45 p.m. – 9:15 p.m. PLENARY SESSION - Fontaine/Fleur De Lis

"Total Survey Error"
Robert Groves, Paul Biemer and Ken Rasinski

9:15 p.m. – 10:30 p.m. DESSERT RECEPTION – Fleur De Lis

Join us for a nightcap and dessert. Reception will be held following the Plenary Session.

Desserts sponsored by: NORC

FRIDAY, MAY 13

7:00 a.m. – 3:00 p.m. APOR REGISTRATION DESK OPEN - Grand Gallerie

7:00 a.m. – 9:00 a.m. Breakfast - Cash & Carry in the Grand Gallerie

7:00 a.m. – 9:00 a.m. NNSP Meeting – Imperial I
WAPOR Meeting – Imperial II
Endowment Committee Meeting – Imperial III

AAPOR SHORT COURSE
Friday - 7:00 a.m. – 8:15 a.m.

Building a Career in Survey Research - Club Atlantic
Peter Miller, Elizabeth Martin, Mark Schulman & Geraldine Mooney – *Continental Breakfast Sponsored by: Westat*

8:15 a.m. – 5:30 p.m. EXHIBITS OPEN: Software, Technology & Books
Fontainebleau C-D

AAPOR CONCURRENT SESSIONS
Friday 8:15 am-9:45 am

Informed Consent - Fontainebleau A
Chair: David Kashihara, Agency for Healthcare
Discussant: Karen Goldenberg, U.S. Bureau of Labor Statistics

Development and Testing of Informed Consent Questions to Link Survey Data with Administrative Records
Nancy Bates, U.S. Census Bureau

The Effect of Introductory Consent Information on Response Rates and Respondent Understanding in a Telephone Interview
Mary Losch, University of Northern Iowa
Shelly Campo, University of Iowa
Gene Lutz, University of Northern Iowa

Do Characteristics of Consent Forms Affect the Response Rate?
Dragana Bolcic-Jankovic, UMass Boston
Brian Clarridge, UMass Boston
Floyd J. Fowler, Jr., UMass Boston

When Silence is Not an Option: Finding a Better Balance Between Confidentiality and Respondent Protection in Ethical Codes
Iain Noble, Department for Education and Skills
Nick Moon, NOP World

Making Contact - Fontainebleau B
Chair: Randall Keesling, RTI International
Discussant: Carl Ramirez, General Accounting Office

Gaining Efficiencies in Scheduling Callbacks in Large RDD National Surveys
Jeffery Stec, InteCap, Inc.
Paul J. Lavrakas, Nielsen Media Research
Charles D. Shuttles, Nielsen Media Research

Calling Strategies, Response Rates, Efficiency and Representativeness
G. Donald Ferree Jr., University of Wisconsin

An Experimental Test of Answering Machine Message Content to Improve Response Rates
Charles D. Shuttles, Nielsen Media Research
Paul J. Lavrakas, Nielsen Media Research

Administrative Interventions and Their Impact on Response Rates in Military Attitude Surveys
James Caplan, Defense Manpower Data Center
Barbara Quigley, Defense Manpower Data Center

The Web and Other Modes I - Brittany
Chair: Stephanie Eckman, NORC

Mode Effects for Hybrid Telephone/Internet Surveys and Reaching Cellphone-Only Households
Dale Kulp, Marketing Systems Group
Melissa Herrmann, ICR/International Communications Research
David Dutwin, ICR/International Communications Research
Steve Lavine, Common Knowledge

The Transition From Telephone to Online Data Collection in Time Series Measurement: The Estimation of Mode Effects
Jon Miller, Northwestern University
Linda Kimmel, Northwestern University
William McCready, Knowledge Networks
Mike Dennis, Knowledge Networks

In Search of Equivalency across Modes: Experimental Results Comparing Alternative Question Formats for Eliciting Dates in Telephone and Web Modes
Leah Christian, Washington State University
Don Dillman, Washington State University
Jolene Smyth, Washington State University

Survey Mode Effects: Comparison between Telephone and Web
Howard Speizer, Market Strategies
Reg Baker, Market Strategies
Karin Schneider, Market Strategies

How Does Social Desirability Affect Responses?: Differences in Telephone and Online Surveys
Humphrey Taylor, Harris Interactive
David Krane, Harris Interactive
Randall Thomas, Harris Interactive

Sampling I - Champagne
Chair: Craig Hill, RTI International

A Test of a Combined RDD/Registration-Based Sampling Model in Oregon's NEP Survey
Joel Bloom, University of Oregon
Warren Mitofsky, Mitofsky International
Joe Lenski, Edison Media Research

Assessment of Address Frame Replacements for RDD Sampling Frames
Michael Link, Centers for Disease Control and Prevention
Michael Battaglia, Abt Associates
Pamela Giambo, Abt Associates
Martin Frankel, Abt Associates and Baruch College, CUNY
Ali Mokdad, Centers for Disease Control and Prevention

Measuring the Efficacy of a Clustered Design in Producing Small Area Estimates in a Survey of Health Insurance Status
Tom Duffy, ORC Macro
Ronaldo Iachan, ORC Macro

The Impact of Cluster (Segment) Size on Effective Sample Size
Steven Pedlow, NORC - University of Chicago
Yongyi Wang, NORC - University of Chicago
Colm OMuircheartaigh, NORC - University of Chicago

Matched Sampling
Douglas Rivers, Stanford University and Polimetrix

Young and Restless Voters - Lemans/Bordeaux
Chair: Jim Bason, Survey Research Center
Discussant: Herb Weisberg, Ohio State Univ.

First Time Voters in the 2004 Election: Who Are They and What Was On Their Minds?
Jennifer De Pinto, CBS News
Megan Thee, New York Times

Young Cell Phone Users and Voting Behavior in Georgia in 2004
James Bason, University of Georgia

Campaign Involvement and the Internet: A Panel Study of Howard Dean Supporters
Courtney Kennedy, Pew Research Center
Scott Keeter, Pew Research Center
Cary Funk, Pew Research Center

Context and Framing - Burgundy
Chair: George Bishop, Unviversity of Cincinnati
Discussant: Jon Krosnick, Stanford University

The Impact of Question Order on the Measurement of Party Identification
Trevor Tompson, Associated Press
Mike Mokrzycki, Associated Press

Order Effects and Vote Preference in the 2004 Presidential Election
Janice Bell, IPSOS-Public Affairs

Did Ballot Order Matter at "The Epicenter"? : An Evaluation of Candidate Ballot Order Effects in the 2004 Ohio Elections
Eric Rademacher, University of Cincinnati
Jason Minser, University of Cincinnati
Kim Downing, University of Cincinnati

Improving Data Quality: Do Changes in Item Formatting and Placement Help Older Respondents Provide Better Data?
Judith Lynch, RTI International
Anne Kenyon, RTI International
Scott Scheffler, RTI International
Stephanie Rizk, RTI International
Katherine Jackson, RTI International
Jiantong Wang, RTI International
Mildred Duke, RTI International

What is Sexual Harassment? It Depends on Who Asks! Framing Effects of Sponsorship on Survey Responses
Roger Tourangeau, Joint Program in Survey Methodology
Frederick Conrad, University of Michigan
Mick Couper, University of Michigan
Mirta Galesic, Joint Program in Survey Methodology

Issues, Candidates and the 2004 Elections - Pasteur/Lafeyette
Chair: Leonie Huddy, SUNY
Discussant: Floyd Ciruli, Ciruli Associates

The Separate Realities of Bush and Kerry Supporters
Steven Kull, Program on International Policy Attitudes

Economic Voting and Political Sophistication in the U.S.: A Reassessment
Jean-Francois Godbout, Northwestern University
Eric Belanger, McGill University

The Role of Issues in the 2004 Election
Merrill Shanks, University of California, Berkeley
Douglas Strand, University of California

What Motivated Americans' Views of the Candidates and Vote Preferences across the 2004 Presidential Campaign?
Neil Malhotra, Stanford University
Jon Krosnick, Stanford University
Gary Langer, ABC News
Daniel Merkle, ABC News

Racial Attitudes - Voltaire
Chair: Bronwyn Nichols, NORC
Discussant: Manuel de la Puente, U.S. Bureau of the Census

The Interdependence of Determinants for the Strength and Direction of Social Desirability Bias in Racial Attitude Surveys
Volker Stocké, University of Mannheim

Blacks' Perceptions of their Racial Context and Prejudice Towards Whites
David Wilson, Gallup/Michigan State University
(PAPOR Student Paper Award Winner)

When Crime Spending Views Reflect Punitiveness: The Context of Racial Prejudice
Steven Barkan, University of Maine
Steven F. Cohn, University of Maine

9:45 a.m. – 10:00 a.m. REFRESHMENT BREAK
In the Exhibit Hall
Fontainebleau C-D

Sponsored by: NORC

AAPOR CONCURRENT SESSIONS
Friday 10:00 am-11:30 am

Data Quality I - Fontainebleau A
Chair: Hernando Rojas, UW Madison
Discussant: Pamela Jull, Applied Research Northwest, LLC

The Effects of Digital Recording of Telephone Interviews on Survey Data Quality
Danna Basson, University of Washington

Use of Dependent Interviewing Procedures to Improve Data Quality in the Measurement of Change
Jeffrey Moore, U.S. Census Bureau
Nancy Bates, U.S. Census Bureau
Joanne Pascale, U.S. Census Bureau
Julia Griffiths, U.S. Census Bureau

Reporting the Frequency of Household Tasks by Elderly Respondents: The Effect of Different Interview Strategies on Data Quality
Johannes van der Zouwen, Vrije Universiteit
Johannes Smit, Vrije Univ. Medical Centre
Marleen Van der Horst, Vrije Universiteit Medical Centre

Managing Data Quality on the Survey of Consumer Finances 2004
Leslie Athey, NORC
Arthur Kennickell, Federal Reserve Board

Theories and Models of Nonresponse - Fontainebleau B
Chair: John Sokolowski, NORC

How Could They Ever, Ever Persuade You? Are Some Refusals Easier To Convert Than Others?
Nick Moon, NOP World
Nickie Rose, NOP World
Nikki Steel, NOP World

A Theory-driven Approach to Reducing Nonparticipation of Twelfth Graders in National Surveys
Young Chun, American Institutes for Research

Participation Decisions Throughout and Across Surveys
Andy Peytchev, University of Michigan
Scott D. Crawford, Survey Sciences Group
Karen Kurotsuchi Inkelas, University of Maryland
Sean McCabe, University of Michigan

Exchange and Compliance: Integrating Theories of Survey Nonresponse
John Goyder, University of Waterloo
Luc Boyer, University of Waterloo
Guil Martinelli, University of Waterloo

Nonresponse Bias – Lemans/Bordeaux
Chair: Paul Lavrakas, Nielsen Media Research
Discussant: David Roe, RTI International

Potential Non-Response Bias from Hard to Reach Respondents
Peyton Craighill, Pew Research Center
Cary Funk, Pew Research Center
Michael Dimock, Pew Research Center

Are We Under-Counting Blue Counties?: Correcting for Disproportional Response Rates at the County Level in National RDD Samples
Michael Dimock, Pew Research Center
Jonathan Best, Princeton Survey Research Associates International

Unit Nonresponse and Error in a National Public Opinion Survey: A Census Matching Approach
Allyson Holbrook, University of Illinois at Chicago
Timothy Johnson, Survey Research Laboratory
Young Cho, University of Illinois at Chicago
Jaime Brugueras, University of Illinois at Chicago

Non-response Bias in the Measurement of Radio Listening: Is it Necessary to Make 13 Contact Attempts?
Robin Gentry, Arbitron Inc.

Nonresponse Bias in Reliability and Validity Estimates for Attitude Items
Zachary Arens, Gallup
(SAPOR Student Paper Award Winner)

Response Scales - Champagne
Chair: John Tarnai, Washington State Univ.
Discussant: Moh Yin Chang, Univeristy of Nebraska

Rating versus Comparative Trade-off Measures: Effects of Task, Topic, Element Differentiation, and Number of Elements on Validity
Randall Thomas, Harris Interactive
Kerri Miller, Harris Interactive
Alyssa Johnson, Harris Interactive

How Much are You Willing to Give? The Response Scales Can Influence the Size of Donations
Mirta Galesic, Joint Program in Survey Methodology
Frederick Conrad, University of Michigan
Mick Couper, University of Michigan
Roger Tourangeau, Joint Program in Survey Methodology

Response Option Ordering: Reconciling Meanings Conveyed by Rating Scale Label and Position
Philip Garland, Stanford University
Jon Krosnick, Stanford University

Comparing Check-All and Forced-Choice Question Formats in Web Surveys: The Role of Satisficing, Depth of Processing, and Acquiescence in Explaining Differences
Jolene Smyth, Washington State University
Don Dillman, Washington State University
Leah Christian, Washington State University
Michael Stern, Washington State University

Health Surveys I - Burgundy
Chair: Kenneth Steve, Nielsen Media Research
Discussant: Christopher D. Karadjov, State University of New York at Oswego

NHANES Converted Refusals: Are They Different from Willing Respondents in Socio-Demographic Composition?
Yinong Chong, CDC
Margaret Carroll, CDC
Pat Montalvan, Westat

Measuring Knowledge of General Cancer Risks - Open Listing versus Targeted Questioning in the 2002 HINTS
Michael Massagli, Dana-Farber Cancer Institute
Vish Viswanath, Harvard School of Public Health/Dana Farber Cancer Institute
Bradford W. Hesse, National Center Institute
Richard P. Moser, National Cancer Institute

Impact of Reports of Hepatitis B Vaccination on Hepatitis A Vaccination Reports
Lisa Carley-Baxter, RTI International
Doug Passaro, University of Illinois at Chicago
Paul Levy, RTI International
Susan Twiddy, RTI International
Ron Hershow, University of Illinois at Chicago

A Balance Between Quality and Cost-Conducting an Adult Tobacco Survey with Multiple Stakeholder Interests
Barbara Fernandez, ORC Macro
Randal ZuWallack, ORC Macro
Kristie Hannah, ORC Macro
Anne Gorrigan, ORC Macro
Kisha Bailly, ORC Macro
Peter Mariolis, CDC

Sampling II - Pasteur/Lafeyette
Chair: Fred Conrad, ISR/University of Michigan

Improving Web Based Intercept Surveys: A Framework For The Active User Sampling
Lars Kaczmirek, ZUMA
Wolfgang Neubarth, ZUMA

Simultaneous Sampling
Karol Krotki, RTI
Cynthia Bland, RTI

Comparison of Traditional Listings and USPS Address Database as a Frame for National Area Probability Samples
Colm O'Muircheartaigh, NORC, University of Chicago
Stephanie Eckman, NORC, University of Chicago
Ned English, NORC, University of Chicago
James Lepkowski, ISR, University of Michigan
Steven Heeringa, ISR, University of Michigan

A New and Better Method For Sample Designs With Disproportionate Sampling
Iain Noble, Department for Education and Skills
Peter Lynn, ISER, Essex University
Patten Smith, BMRB International

Assessing the Effectiveness of an Optimal-Allocation Sample Design to Obtain Reliable Estimates of Health Insurance among Minority Populations
Tom Duffy, ORC Macro
Ronaldo Iachan, ORC Macro

Recent Methodological Research on the Knowledge Networks Approach - Voltaire
Chairs: William McCready, Knowledge Networks
Discussant: Linchiat Chang, Consultant

Comparing the Knowledge Networks Web-Enabled Panel and the In-Person 2002 General Social Survey: Experiments with Mode, Format, and Question Wordings
Tom W. Smith, NORC, University of Chicago
Rick Li, Knowledge Networks
Mike Dennis, Knowledge Networks

Data Collection Mode Effects Controlling for Sample Origins in a Panel Survey: Telephone versus Internet
Cindy Chatt, Gallup
Mike Dennis, Knowledge Networks
Rick Li, Knowledge Networks
Paul Pulliam, RTI

Statistical Tests of Data Quality in a Contingent Valuation Survey Using Knowledge Networks Data
Joel Huber, Duke University
Jason Bell, Duke University
W. Kip Viscusi, Harvard University

Correcting for Self-Selection Bias Using the Heckman Selection Correction in a Valuation Survey Using Knowledge Networks
Trudy Cameron, University of Oregon
George M. Deshazo, UCLA
Mike Dennis, Knowledge Networks

MEET THE AUTHOR SESSIONS
Friday - 11:30 a.m. – 12:15 p.m.

Meet the Author Session - Fontainebleau C-D

Authors: Stanley Presser, Jennifer Rothgeb, Mick Couper, Judith Lessler, Elizabeth Martin, Jean Martin, and Eleanor Singer
Text: *Methods for Testing and Evaluating Survey Questionnaires (Wiley, 2004)*

Author: George Bishop
Text: *The Illusion of Public Opinion (Rowman & Littlefield, 2004)*

AAPOR POSTER SESSION
Friday 11:30 am-12:15 pm

Poster Session - Fontainebleau C-D

1. Analysis of Reinterview Data from NLSY97 Rounds 2 Through 5
Parvati Krishnamurty, NORC

2. A Transdisciplinary Approach to Protocol Development: A Case Study
Melissa Clark, Brown University
Suzanne Colby, Brown University
Julie Boergers, Rhode Island Hospital
Christopher Kahler, Brown University
Susan Ramsey, Rhode Island Hospital
Charles Neighbors, The Miriam Hospital
Johanna Lewis-Esquerre, Brown University

3. Too Much of a Good Thing? Working Through Establishment Gatekeepers
Andrew Zukerberg, U.S. Census Bureau
Andrew Soderborg, U.S. Census Bureau
Randy Parmer, U.S. Census Bureau
Steve Tourkin, U.S. Census Bureau

4. Dual-Incentive Experiment with New Business Owners
Janice Ballou, Mathematica Policy Research
Dave DesRoches, Mathematica Policy Research
Frank Potter, Mathematica Policy Research

5. The RAS Model: A Simple Test
Agnieszka Dobrzynska, Université de Montréal
André Blais, Université de Montréal

6. Knowledge, Question Format and Propensity to Guess
Nick Allum, University of Surrey
Patten Smith, BMRB International

7. The Use of the CSAP's Core Measures Initiative: Cross-Sectional Development but Repeated Measures Application
Mariah Storey, University of Wyoming
Eric L. Canen, University of Wyoming

8. Using Survey Data to Provide Small-Area Resource Allocation Models
Bill Blyth, TNS
Graham Moon, University of Portsmouth
Gemma Holt, University of Portsmouth

9. Interpreting Response Latencies: A Mixed Method Approach
Christine Carabain, Vrije Universiteit
Wil Dijkstra, Vrije Universiteit

10. Survey Experiences Influence Survey Attitudes and Participation Propensity
Robert Brage, Stockholm Univ.
Lars R. Bergman, Stockholm Univ.

11. Response Propensity Patterns among Respondents Selected in Multiple Years of a Repeated Cross-sectional Survey
Sherman Edwards, Westat
John Rauch, Westat

12. First-Level Interactions and Final Outcomes: Adding Value to Behavior Coding
Ashley Landreth, U.S. Census Bureau
Jennifer Hunter, U.S. Census Bureau

13. Evaluating the Impact on Data Quality of Conversion from PAPI to the iPAQ Handheld Computer for In-person Data Collection
Randall Keesling, RTI International
Scott Bell, Nielsen Media Research

14. A Profile of Self-Selecting Web Respondents
Paul Guerino, American Institutes for Research

15. The Arbitron/NMR PPM Response Rate Test
Beth Webb, Arbitron
Arianne Buckley, Arbitron

16. Measurement Issues in Behavioral Self-report Measures: Effects of Response Format, Respondent Load, Number of Elements, and Time Frame Reference
Randall Thomas, Harris Interactive
Jonathan Klein, University of Rochester
Lisa Wilding, Harris Interactive
Susan Behnke, Harris Interactive
Robin Repass, Harris Interactive

17. Effects of Suggestive and Non-Suggestive Probing on Different Types of Inadequate Answers
Yfke Ongena, Vrije Universiteit Amsterdam
Wil Dijkstra, Vrije Universiteit Amsterdam

18. Survey Non-Response in a National Area Probability Sample as a Dimension of Survey Quality: An Analysis of Community Characteristics
Jibum Kim, NORC
John Sokolowski, NORC

19. What's Up Doc? Mixing Web and Mail Methods in a Survey of Physicians
Timothy Beebe, Mayo Clinic
Giles R. Locke, Mayo Clinic
Sunni Barnes, Mayo Clinic

20. Collecting Geographic Location Information in an RDD Survey
David Grant, UCLA
Anthony Ramirez, UCLA
Sherman Edwards, Westat
John Rauch, Westat

21. The Utility of Extended Longitudinal Profiles in Predicting Future Health Care Expenditures
Steven Cohen, AHRQ
Trena Ezzati-Rice, AHRQ
William Yu, AHRQ

22. Nonresponse in Telephone Surveys: The Reporting of Response Rates and Outcome Measures
Richard Seltzer, Howard University

23. Explaining Survey Quality to Data Users: Can We Get Them to Care?
Susan P. Love, U.S. Census Bureau
Deborah H. Griffin, U.S. Census Bureau

24. The National Survey of America's Families, 1997-2002: A Project Summary of Response Rate Initiatives, Implementation, and Results
Timothy Triplett, Urban Institute
Adam Safir, Urban Institute
Natalie Abi-Habib, Urban Institute

25. Changing Response Pattern over Time: Results from Statewide Telephone Surveys Conducted in 1999 and 2004
Cyndi Garvan, University of Florida
Colleen Porter, University of Florida
R. Paul Duncan, University of Florida
Qin Li, University of Florida
Vijaya Komaragiri, University of Florida

26. Reaching Low-Income Populations for Telephone Interviews
Ellen Marks, RTI International
Murrey Olmsted, RTI International

27. Advance Notice and Confidentiality Statements for Listed-Assisted Telephone Surveys
Brian J. Meekins, Bureau of Labor Statistics
Roberta L. Sangster, Bureau of Labor Statistics

28. Does Imputation Bias Lead to Finding Significantly More Uninsured in the Current Population Survey's Estimates of Health Insurance Coverage?
Michael Davern, University of Minnesota
Lynn Blewett, University of Minnesota
Kathleen Thiede Call, University of Minnesota
Holly Rodin, University of Minnesota

29. Analysis of Behavioral Risk Factor Surveillance System Demographic Bias and Effect of Weighting on Risk Factor Estimates
Herbert Stackhouse, CDC
Pranesh Chowdhury, BCA

30. Predicting Exit Poll Interviewer Response Rates
Clint Stevenson, Brigham Young University
Daniel Magleby, Brigham Young University
Quin Monson, Brigham Young University
Kelly Patterson, Brigham Young University
Howard Christensen, Brigham Young University

31. Public Perceptions of Poll Results - Do the Details Matter?
Monika McDermott, University of Connecticut
Samuel Best, University of Connecticut

32. Who Deliberates in the Public Square and Does It Matter?
Robert Goidel, Louisiana State University
Charles Zewe, Louisiana State University
Steven Procopio, LSU Public Policy Research Lab

33. A Pre-Validation Survey of Voter Turnout
Michael McDonald, George Mason University

34. Predicting Sampled Respondents' Likelihood to Cooperate in a Mail Survey: Part III
A.T. Burks, Nielsen Media Research
Paul J. Lavrakas, Nielsen Media Research
Mildred Bennett, Nielsen Media Research

35. Partners in Preservation of Opinion Data: Developments of a Library of Congress Initiative
Lois Timms-Ferrara, Roper Center
Marc Maynard, Roper Center

36. Exemplification, War and Iraq: Examining the Effects of Exposure to Violence in Iraq on Attitudes on the Iraq War
Russ Tisinger, University of Pennsylvania
Kelli Lammie, University of Pennsylvania

37. Sound Bytes: Capturing Audio in Survey Interviews
Sue Ellen Hansen, University of Michigan
Maria Krysan, University of Illinois, Chicago
Mick Couper, University of Michigan

38. The Role of Economics, Religion and National Security in the 2004 Presidential Vote
Leonie Huddy, SUNY
Stanley Feldman, SUNY
Sarah Dutton, CBS News

12:15 p.m. – 2:05 p.m. PLATED LUNCH AND AAPOR PRESIDENTIAL ADDRESS
Fontaine/Fleur De Lis

(Core Meal)

AAPOR CONCURRENT SESSIONS
Friday 2:15 pm-3:45 pm

Cross-National Research I - Brittany
Chair: Kathy Frankovic, CBS News
Discussant: Sherman Edwards, Westat

How Do Citizens Respond To Photographs of Politicians?: A Cross-National Analysis
Robert Eisinger, Lewis & Clark College
Thomas Petersen, Institut für Demoskopie Allensbach

Daily Contact as A Proxy Measure of Personal Networks: Data from 45 Surveys in Nine Societies
Yang-chih Fu, Academia Sinica

Cross-National Contact Strategies and Effects
Oliver Lipps, Mannheim Research
Grant Benson, University of Michigan Institute for Social Research

Using Anchoring Vignettes to Assess Cross Cultural Comparability of Consumer Ratings
Patricia Gallagher, Center for Survey Research

Social Change - Fontainebleau B
Chair: Jibum Kim, NORC
Discussant: David Fan, University of Minnesota

Changes in Attitudes Toward American Institutions and Occupations
Richard Seltzer, Howard University
Rhea Roper, Howard University

Are Older Adults Losing Social Capital? 2004 and 1996 Compared
Thomas Guterbock, University of Virginia
Edna Renee Macbeth, University of Virginia

Deliberative Attitude Change: Changing Your Mind Within an Interview
Patrick Fournier, Université de Montréal
André Blais, Université de Montréal
Joanna Everitt, University of New Brunswick
Elisabeth Gidengil, McGill University
Neil Nevitte, University of Toronto

Deliberation and Its Discontents: A Comparison of Survey and Deliberative Discussion Results, 1987-2003
Frank Rusciano, Rider University

Increasing Response Rates - Fontainebleau A
Chair: Charles Shuttles, Nielsen Media Research
Discussant: James Caplan, Defense Manpower Data Center

(Inter) Net Gain? Experiments to Increase Web-Based Response
Steve Tourkin, U.S. Census Bureau
Shawna Cox, U.S. Census Bureau
Randy Parmer, U.S. Census Bureau
Andrew Zukerberg, U.S. Census Bureau

Does Type of Pre-Notification Affect Web Survey Response Rates?
Michele Harmon, Westat
Elizabeth Westin, Westat
Kerry Levin, Westat

Did You Get The Message? Using E-Mail and SMS for Prenotification in Web Surveys
Wolfgang Neubarth, ZUMA
Michael Bosnjak, University of Mannheim
Wolfgang Bandilla, ZUMA
Mick Couper, University of Michigan
Lars Kaczmirek, ZUMA

Analysis of Break-off Patterns in Web Surveys
Syed Ahsan, NORC
Ronald Broach, NORC

Interviewer Effects - Champagne
Chair: Randall Thomas, Harris Interactive

The Effects of Interviewer Gender and Race on Sensitive Health Questions
Gregg VanRyzin, Baruch College / CUNY
Martin Frankel, Abt Associates and Baruch College, CUNY
Farzad Mostashari, New York City Department of Health
Stephen Immerwahr, Baruch College / CUNY

Interviewer Effects in a RDD Telephone Pre-Election Poll in Minneapolis 2001: An Analysis of the Effects of Interviewer Race and Gender
Mario Callegaro, University of Nebraska, Lincoln
Femke De Keulenaer, University of Antwerp
Jon Krosnick, Stanford University
Robert Daves, Star Tribune

Gender-of-Voice Effects in an ACASI Study of Same-Sex Behavior
Kristine Fahrney, RTI International
Jennifer Uhrig, RTI International
Richard Wolitski, CDC

Estimating Measurement Error Using Interviewer Feedback
Jim Wolf, Indiana University
Andy Hutcherson, Indiana University
Brianne Peyton, Indiana University
Charles Hulen, Indiana University
Sharon Sidenbender, Indiana University

Will They Talk to Us? Survey Performance Rates of a Cell Phone Sample
Anna Fleeman, Arbitron

The Do Not Call Registry - Lemans/Bordeaux
Chair: Colm O'Muircheartaigh, University of Chicago
Discussant: Diane Burkom, Battelle/CPHRE

Has the National Do Not Call Registry Helped or Hurt Survey Participation Rates?
Michael Link, Centers for Disease Control and Prevention
Ali Mokdad, Centers for Disease Control and Prevention
Dale Kulp, Marketing Systems Group

The Impact of the National Do Not Call Registry on Social Survey Response
Z. Joan Wang, REDA International
Elham-Eid Alldredge, REDA International
Jian Zhu, REDA International
Michelle Cantave, REDA International
N. Craig Gilmore, REDA International

The Do Not Call Registry: Friend or Foe? The Effect of Do Not Call Lists on Survey Response
Berwood Yost, Franklin & Marshall College
Jennifer Harding, Franklin & Marshall College
Christina Abbott, Franklin & Marshall College
Angela Knittle, Franklin & Marshall College

Listed Numbers and Do Not Call Registration in a National Telephone Survey
Jim Lepkowski, University of Michigan
Howard Schuman, University of Michigan
Rui Wang, University of Michigan
Richard Curtin, University of Michigan

Changing Political Minds - Burgundy
Chair: Dan Merkle, ABC News
Discussant: Rob Daves, Star Tribuine

Handling the Truth: An Evaluation of Voter Capability to Discern Political Advertisement Accuracy
Chris Borick, Muhlenberg College

The Swift Boat Ads in Comparative Context: An Empirical Examination of Advertisement Effectiveness
Chris Borick, Muhlenberg College

Changing Minds in the 2004 Election?
Jeffrey Jones, Gallup Poll
Joseph Carroll, Gallup Poll

Using the Internet to Learn About Presidential Candidates and Issue Positions in the 2004 Presidential Primary and General Election Campaigns
Kenneth Winneg, University of Pennsylvania

Religion and Public Opinion - Pasteur/Lafeyette
Chair: Ariela Keysar, CUNY
Discussant: Hugh Gladwin, Florida International Institute

Trends in Public Opinion among Religious Fundamentalists, 1972-2002
Anne Driscoll, University of California, Davis
Katherine Heck, University of California, Davis

The National Spiritual Transformations Study: An Overview
Tom W. Smith, NORC, University of Chicago

The Public Acceptance of Evolution and the Big Bang
Jon Miller, Northwestern University
Linda Kimmel, Northwestern University
Rafael Pardo, Foundation BBVA

The Impact of "Church Night" on Southern Polls with Short Field Time
Richard Clark, Carl Vinson Institute of Government
Craig Smith, School of Public and International Affairs, UGA

A Space for "Place" in Racial Attitudes Research: Exploring Context Effects - Voltaire
Chair: Matthew Hunt, Northeastern University
Discussant: Tyrone Forman, Univ. of Illinois at Chicago

Congregational Context and Racial Attitudes
Michael Emerson, University of Notre Dame

Neighborhood Racial Composition and Perceived Racial Discrimination in the Black Women's Health Study
Matthew Hunt, Northeastern University
Lauren Wise, Boston University
Marie-Claude Jipguep-Akhtar, Howard University
Lynn Rosenberg, Boston University

Educational Attainment and Perceived Discrimination among African Americans: Testing the Racial Awareness and Racial Context Hypotheses
Mosi Ifatunji, University of Illinois at Chicago

"Trust in People" Among Black and White Americans -- A Multi-Level Analysis
Marylee Taylor, Pennsylvania State University

3:45 p.m. – 4:00 p.m. REFRESHMENT BREAK
In the Exhibit Hall
Fontainebleau C-D

Sponsored by: Abt Associates

AAPOR CONCURRENT SESSIONS
Friday 4:00 pm-5:30 pm

Political Issues I - Fontainebleau A
Chair: David Wilson, Gallup Organization
Discussant: David Dutwin, International Communications Research

After the California Recall: Governor's Ratings, 2004 Election Choices, and the Implications for Direct Democracy
Mark Baldassare, Public Policy Institute of California

Overlapping Decisions in a Presidential Campaign: Battleground Wisconsin in 2004
G. Donald Ferree Jr., University of Wisconsin

As Easy as 1-2-3? An Assessment of Ranked-Choice Voting in the San Francisco 2004 Election
Lisel Blash, San Francisco State University
Francis Neely, San Francisco State University

Nonresponse in Exit Poll Methodology: A Case Study in Mexico
Rene Bautista, University of Nebraska-Lincoln
Mario Callegaro, University of Nebraska-Lincoln
Francisco Abundis, Parametria SA de CV
Jose A. Vera, Parametria SA de CV

Incentives I - Champagne
Chair: Tracy Tuten Ryan, Virginia Commonwealth Univ.
Discussant: Eleanor Singer, Univ. of Michigan Survey Research Center

The Effect of Incentives on Two Physician Mail Surveys: A Response Rate Comparison
Jeanine Christian, Battelle Centers for Public Health Research and Evaluation

Individualized Treatments within a Household: Can Targeted Incentives Raise Young Male Response?
Christina Frederick, Arbitron

Value and Timing Strategies in Prize Draws: A Further Examination of the Immediacy Effect in Web Surveys
Tracy Tuten, Virginia Commonwealth University
Mirta Galesic, Joint Program in Survey Methodology
Michael Bosnjak, University of Mannheim

Comparative Analyses of Parallel Paper, Phone, and Web Surveys With and Without Incentives: What Differences Do Incentive and Mode Make?
Danny Olsen, Brigham Young University
Vaughn Call, Brigham Young University
Steve Wygant, Brigham Young University

Translation - Brittany
Chair: Janet Harkness, ZUMA
Discussant: Karen Bogen, UMass Boston

How Survey Questions Mean
Janet A. Harkness, ZUMA

Lost in Translation: Quality Issues in Multilingual Research
Lauren Doerr, NORC - University of Chicago

Translation Effects in the Demographic Health Survey
Adam Gluck, ORC Macro

Managing Survey Translation: Knowledge Base for Researchers Who are Not Multi-Lingual
Man-chi Mandy Sha, NORC

Cellular Phones I - Fontainebleau B
Chair: Linda Dimitropoulos, RTI International
Discussant: Peyton Craighill, Pew Research Center

Characteristics Related to Cell Phone Status: Why Generation Y Should be Targeted
Laura Hancock

Comparison of Cell-Only and Landline Telephone Users Among 18-24 Year-Olds
Robin Albee, University of Missouri-Columbia
James Cole, University of Missouri

Can You Hear Me Now?: Differences in Vote Behavior in the Cell and Landline Populations
Tarek Albaghal, University of Maryland

Ownership & Usage Patterns of Cell Phones: 2000 - 2005
Peter Tuckel, Hunter College
Harry O'Neill, NOP World

Panel Surveys - Burgundy
Chair: Todd Robbins, Abt Associates
Discussant: Janet Streicher, CitiBank

Looking for Mr. Right: Effective Locating Techniques for Panel Surveys
Angela Herrmann, NORC
Karen Grigorian, NORC
Syed Ahsan, NORC

Ensuring High Cooperation Rates in a 4-year Longitudinal Study of People with Multiple Sclerosis
Todd Robbins, Abt Associates
Michele Laramie, Abt Associates
Sarah Minden, Abt Associates
Debra Frankel, Abt Associates
Carin Cartwright-Chunga, Abt Associates
Brander Sieber, Abt Associates
Paul Howard, Abt Associates

Using Process Data to Predict Attrition From a Panel Survey: A Case Study
Femke De Keulenaer, University of Antwerp

Increasing Response to Mailed Follow-up Surveys: A Test of Two Motivators
Sue Mann, Fred Hutchinson Cancer Research Ctr.
Diana Lynn, Fred Hutchinson Cancer Research Ctr.
Arthur Peterson, Jr., Fred Hutchinson Cancer Research Ctr.

Pre-election Surveys I - Lemans/Bordeaux
Chair: Peter Miller, Northwestern University

Media Coverage & Nonresponse: A Time Series Examination of the 2004 Election Cycle
Natalie Jomini Stroud, University of Pennsylvania
Kate Kenski, University of Pennsylvania

Estimating Non-Response in Presidential Polls
Eugene Ericksen, Temple University
Sarah Butler, Temple University

Methodological Issues in Pre-Election Polling: Lessons from ABC News' 32-Night Tracking Poll
Daniel Merkle, ABC News
Gary Langer, ABC News
David Lambert, TNS

How SurveyUSA Weights Its Election Poll Data
Joseph Shipman, SurveyUSA
Jay Leve, SurveyUSA

Partisan Identification in the Pre-election Environment
Frank Newport, Gallup Organization

Media Use - Pasteur/Lafeyette
Chair: Scott Keeter, Pew Research Center
Discussant: Kim Downing, Institute for Policy Research

Media Use, Political Talk and Social Capital
Hernando Rojas, University of Wisconsin-Madison
Dhavan Shah, University of Wisconsin-Madison

The Pitfalls of Self-Reported News Exposure
Markus Prior, Princeton University

Putting Survey Response Data into Context of Everyday Life: Data Modeling With an Electronic Measurement System of Media Exposure
Beth Webb, Arbitron
Jay Reid, Arbitron

The PPM/Diarykeeper Study: Same-Sample Comparisons of Radio Listening Behavior
Fran Harmon, Arbitron
Beth Webb, Arbitron

The European Social Survey: Improving Survey Quality In a Cross-National Context - Voltaire
Chair: Caroline Roberts, City University
Discussant: Peter Mohler, ZUMA

From Design to Implementation: Methodological Innovation on the ESS
Roger Jowell, City University
Rory Fitzgerald, City University
Gillian Eva, City University

Refusal Conversion and the Estimation of Non-Response Bias in The European Social Survey (Round 1): An Analysis of Contact Forms Combined With Substantive Data
Jaak Billiet, University of Leuven
Inneke Stoop, Social & Cultural Planning Office
Rory Fitzgerald, City University

Methodological Advances on the ESS: A Mixed Mode Future?
Caroline Roberts, City University
Peter Lynn, University of Essex
Robert Machin, Gallup Europe
Agnes Illyes, Gallup Europe

A New Source of US-European Comparisons on Citizen Engagement
Marc M. Howard, Georgetown University

5:45 p.m. – 6:30 p.m. AAPOR Honorary Life Members' Reception – Club Atlantic

This is a great time to recognize our Honorary Life Members These folks represent some of the biggest names in public opinion research, the ones who did the first polls, conducted the first research, wrote the classics in our field. Stop by and say hello – Honorary Life Members will be wearing a ribbon so you'll be able to identify the face behind that famous name.

6:30 p.m. DINNER ON YOUR OWN – Gather with a group of friends for dinner and hit the Miami hotspots.

9:30 p.m. PUB CRAWL – Details will be posted on the AAPOR message board.

SATURDAY, MAY 14

7:00 a.m. – 3:00 p.m.	AAPOR REGISTRATION DESK OPEN – Grand Gallerie
7:00 a.m.	FUN RUN/WALK – Sponsored by Gallup
7:00 a.m. – 9:00 a.m.	Breakfast - Cash & Carry in the Grand Gallerie
7:00 a.m. – 9:00 a.m.	Heritage Interviewing Meeting – Imperial I
7:30 a.m. – 9:00 a.m.	Small Business/Consultants' Meeting – Imperial II
8:15a.m. – 3:00 p.m.	EXHIBITS OPEN: Software, Technology & Books Fontainebleau C-D

AAPOR CONCURRENT SESSIONS
Saturday 8:15am-9:45 am

Difficult and Rare Populations - Fontainebleau A
Chair: Beth-Ellen Pennell, University of Michigan

Interviewing Persons with Hearing Impairments
Elisha Smith, Mathematica Policy Research
Susan Mitchell, Mathematica Policy Research
Debra Wright, Mathematica Policy Research

Hospice, HIPAA, and Hope: Survey Research with the Terminally Ill
Terri Kovach, Wayne State University

Amish Women's Health Survey: Methodological Considerations with Unique Populations on Sensitive Topics
Berwood Yost, Franklin & Marshall College
Christina Abbott, Franklin & Marshall College
Jennifer Harding, Franklin & Marshall College
Angela Knittle, Franklin & Marshall College

A Brain is Terrible Thing to Lose: Locating U.S.-Educated Foreign Nationals Intending to Live Abroad
Lauren Seward, NORC
Angela Herrmann, NORC
Tom Hoffer, NORC

Gaining Cooperation - Fontainebleau B
Chair: Brad Edwards, Westat
Discussant: Murray Edelman, CBS News

Analysis of Behavioral Risk Factor Surveillance System Partial Completes and Terminations
Herbert Stackhouse, CDC
Ziya Gizlice, NC State Center for Health Statistics

Avoiding Refusal Training: Comparative Work and Indication of Long Term Impact
Mark McConaghy, Office of National Statistics, UK
Siobhan Carey, Office of National Statistics, UK

A Successful Conversion or Double Refusal: A Study of the Process of Refusal Conversions in Telephone Survey Research
Kana Fuse, Ohio State University
Dong Xie, University of Central Arkansas

Evaluating Methods for Increasing Physician Survey Cooperation
Murrey Olmsted, RTI International
Emily McFarlane, RTI International
Joseph Murphy, RTI International
Craig Hill, RTI International

Internet and Web Surveys I -Brittany
Chair: Scott Crawford, Survey Sciences Group
Discussant: Bill McCready, Knowledge Networks

Using the Web to Survey College Students: Institutional Characteristics That Influence Survey Quality
Scott D. Crawford, Survey Sciences Group
Sean Esteban McCabe, University of Michigan
Karen Kurotsuchi Inkelas, University of Maryland

What They See Is Not What We Intend-Gricean Effects in Web Surveys
Ting Yan, Joint Program in Survey Methodology

Visual Context Effects in Web Surveys
Mick Couper, University of Michigan
Frederick Conrad, University of Michigan
Roger Tourangeau, Joint Program in Survey Methodology

Interactive Feedback Can Improve Quality of Responses in Web Surveys
Frederick Conrad, University of Michigan
Mick Couper, University of Michigan
Roger Tourangeau, Joint Program in Survey Methodology
Mirta Galesic, Joint Program in Survey Methodology

Multiple and Mixed Modes - Champagne
Chair: Darby Steiger, Gallup Organization
Discussant: Mario Callegaro, University of Nebraska

Data Quality Issues in a Multimode Survey
Claire Wilson, Mathematica Policy Research
Debra Wright, Mathematica Policy Research
Tom Barton, Mathematica Policy Research
Paul Guerino, American Institutes for Research

Dual-Frame, Dual-Mode Designs to Improve RDD Survey Response Rates: An Initial Feasibility Study
Jill Montaquila, Westat
Mary Hagedorn, Westat
Michael Brick, Westat
Shelley Brock Roth, Westat
Priscilla Carver, Westat
Chris Chapman, National Center for Education Statistics

Web Survey Methodologies: A Comparison of Survey Accuracy
Jon Krosnick, Stanford University
Norman Nie, Stanford University
Douglas Rivers, Stanford University and Polimetrix

Cellular Phones II - Lemans/Bordeaux
Chair: Clyde Tucker, Bureau of Labor Statistics
Discussant: Linda Piekarski, Survey Sampling

The Prevalence and Impact of Wireless Substitution: Updated Data from the 2004 National Health Interview Survey
Stephen Blumberg, NCHS/CDC
Julian Luke, NCHS/CDC
Marcie Cynamon, NCHS/CDC

Non-interviews in Mobile Phone Surveys
Vasja Vehovar, University of Ljubljana

Quality Assessed: Cellular Phone Surveys versus Traditional Telephone Surveys
Charlotte Steeh, Georgia State University

Pre-election Surveys II - Pasteur/Lafeyette
Chair: Michael Traugott, University of Michigan
Discussant: Lowndes F. (Rick) Stephens, University of South Carolina

Re-examining Our Assumptions in Pre-Election Polls
Randall Thomas, Harris Interactive
David Krane, Harris Interactive
John Bremer, Harris Interactive
Humphrey Taylor, Harris Interactive
Renee Smith, Harris Interactive

A Dispositional Model of Election Polls
David Dutwin, ICR/International Communications Research
Melissa Herrmann, ICR/International Communications Research

A New "Interval" Measure of Election Poll Accuracy
Jay Leve, SurveyUSA
Joseph Shipman, SurveyUSA

The Performance of the Polls in the 2004 Presidential Election
Michael Traugott, University of Michigan

Health Surveys II - Burgundy
Chair: Mary Losch, University of Northern Iowa
Discussant: Sherman Edwards, Westat

Quality of Children's Height, Weight, and BMI information in the MEPS
Frances Chevarley, AHRQ

Comparing Administrative and Survey Data to Identify Adults with Chronic Conditions
Jennifer Dunne, University of MA Boston
Patricia Gallagher, University of MA Boston
Vickie Stringfellow, University of MA Boston

An Assessment of the Quality of Health Data Collected from Multiple Sources: Registries, Provider Records, and Household Reports
Meena Khare, NCHS/CDC
Linda Piccinino, Abt Associates

Michael Battaglia, Abt Associates
Assessing the Impact of Medicaid Bias on Counts of the Uninsured
Robert Goidel, Louisiana State University
Dek Terrell, Louisiana State University
Steven Procopio, Louisiana State University
Douglas Schwalm, Louisiana State University

Religion and Politics: Values and Divisions in the 2004 Election - Voltaire
Chair: Ariela Keysar, Brooklyn College

Religious Affiliations and the 2004 Presidential Election
Marjorie Connelly, New York Times

Evangelical Voters: Bedrock of the Republican Party
Scott Keeter, Pew Research Center

Religion, Ideology, Partisanship and "Moral Values" in the 2004 Election
Gary Langer, ABC News
Jon Cohen, ABC News

Religious Identification, Congregational Membership and Political Party Preference
Ariela Keysar, Brooklyn College
Barry Kosmin, Institute for Jewish Policy Research

9:45 a.m. – 10:00 a.m. REFRESHMENT BREAK
In the Exhibit Hall
Fontainebleau C-D

Sponsored by: SRBI

AAPOR CONCURRENT SESSIONS
Saturday 10:00 am-11:30 am

Data Quality II - Fontainebleau A
Chair: Nicolaos Synodinos, Univeristy of Hawaii
Discussant: Polly Phipps, Bureau of Labor Statistics

Improving Survey Quality in the Measurement of Social Capital
Abigail Dewar, Office for National Statistics
Amanda Wilmot, Office for National Statistics

Development of the UK Continuous Population Survey From a Data Collection Perspective: Improving Survey Questions and the Quality of Data Output
Alison Blackwell, Office for National Statistics
Abigail Dewar, Office for National Statistics

Data Quality and the Use of Standardized Child Assessments in Survey Research
Michelle Ernst, NORC - University of Chicago
Mike Pergamit, NORC - University of Chicago

Measuring Quality in Observational Data Collection
Jaana Myllyluoma, Battelle CPHRE
Diana Buck, Battelle CPHRE

Telephone Surveys - Fontainebleau B
Chair: Mike Brick, Westat
Discussant: Karol Krotki, RTI International

Study of Ported Telephone Numbers Using the NeuStar Database
Michael Yang, University of Chicago
Manas Chattopadhyay, University of Chicago
Ned English, NORC, University of Chicago

Purging Out-Of-Scope and Cellular Telephone Numbers From RDD Samples
Michael Battaglia, Abt Associates
Meg Ryan, Marketing Systems Group
Marcie Cynamon, NCHS/CDC

Testing the Impact of Caller ID Technology on Response Rates in a Mixed Mode Survey
Norman Trussell, Nielsen Media Research
Paul Lavrakas, Nielsen Media Research

Using Multiple Sources of Contact Data to Improve Quality in Telephone Surveys
Jason Markesich, Mathematica Policy Research
Anne Ciemnecki, Mathematica Policy Research
Nuria Diaz-Tena, Mathematica Policy Research
Anne Hower, Mathematica Policy Research

School and Youth Surveys - Brittany
Chair: Maria Krysan, University of Illinois, Chicago
Discussant: Roger Tourangeau, University of Maryland

Re-examining Approaches to Achieving High Response Rates on Web-based Surveys of Post-secondary Students: Results from a Follow-up Survey
Bronwyn Nichols Lodato, NORC-University of Chicago
Rashna Ghadialy, NORC-University of Chicago

Measurement Errors in Student's Proxy Reports and Their Consequences for SES-Effects on School Achievement
Frauke Kreuter, Joint Program in Survey Methodology
Kai Maaz, MPI
Rainer Watermann, MPI

Assessing Environmental Education: The Development and Validation of an Upper Elementary Environmental Education Instrument
Pamela Jull, Applied Research Northwest
Aaron Ignac, Applied Research Northwest

Indicators of Adolescent Well-Being in the NLSY97
Tiffany R King, NORC/Loyola University

The Web and Other Modes II -Champagne
Chair: Linda Dimitropoulos, RTI International
Discussant: Don Dillman, Washington State Univ.

Qualitative Comparison of Paper and Online Self-Administered Modes
Karen Grigorian, NORC - University of Chicago
Scott Sederstrom, NORC - University of Chicago

Evaluating Nonsampling Errors in a Study Comparing Data Collected by Mail and Using the Web
Virginia Lesser, Oregon State University
Lydia Newton, Oregon State University

Effectiveness of E-mail and Paper Mail Notifications for Internet Surveys
Paul Ruggiere, University of North Texas
D'Arlene Ver Duin, University of North Texas

Transforming a Paper Survey into a Web-based Survey: Respondent Experiences
Annette Luyegu, Mathematica Policy Research

Incentives II - Burgundy
Chair: Colleen Porter, University of Florida
Discussant: Polly Armbsy, Coda Research

A Test of Monetary Incentives for a Large-Scale Establishment Survey
Paul Biemer, RTI International
Christopher Ellis, RTI International
Kimberly Robbins, RTI International
Angela Pitts, RTI International

Exploring a New Establishment Survey Incentive to Improve Response Rates
Amy Luo, Ernst &Young
Glenn White, Ernst &Young

An Incentive Experiment for a Survey of School Districts and Schools
David Cantor, Westat
Thomas A. Fiore, Westat
Beth Sinclair, Westat
James Demery, Westat
Ellen Schiller, Abt Associates

Offering Cash Incentives to Students in a School Setting: Effects of an Incentive on Student Participation Rates
Debbie Herget, RTI International
Tiffany Lytle, RTI International
Ellen Stutts, RTI International
Peter Siegel, RTI International
Kimberly Ault, RTI International
Mani Medarametla, RTI International
James Rogers, RTI International
Daniel Pratt, RTI International

Predicting Elections - Lemans/Bordeaux
Chair: Mark Schulman, Schulman, Ronca & Bucuvalas

Polls Apart?: Factors Affecting Vote Choice for President
Randall Thomas, Harris Interactive
Mitchell Sanders, Harris Interactive
Renee Smith, Harris Interactive
Robin Repass, Harris Interactive
Alyssa Johnson, Harris Interactive

A Probability Theory-Based Test of the Reliability of 2004 Election Polls
Joel Bloom, University of Oregon
Jennie Pearson, University of Oregon

Survey of Political Surveys [SOPS]
Nat Ehrlich, Michigan State Univ.
Larry Hembroff, Michigan State Univ.

Outcomes and Lessons Learned from Polling Voters about the Voting Experience
Lonna Atkeson, VoteWatch
Rich Clark, VoteWatch
Julio Cesar Campero, Aguirre International
Joan Crowley, VoteWatch
Joan Decker, VoteWatch
Susan Gabbard, Aguirre International
Marie Hansberry, VoteWatch
Trish Hernandez, Aguirre International
Steven Hertzberg, Votewatch
Kristin Juffer, VoteWatch
Kerry Krisman, VoteWatch
Carlos Leon, VoteWatch
Kate Mulqueen, VoteWatch
Robert Santos, NuStats
Fritz Scheuren, NORC
Diane Stoner, VoteWatch
Carmen Sum, Aguirre International
Nicole Vicinanza, Aguirre International
Cheryl Wiese, VoteWatch
Jen Ziemke, VoteWatch

Vote Over-Reporting: Testing the Social Desirability Hypothesis in Telephone and Internet Surveys
Allyson Holbrook, University of Illinois, Chicago
Jon Krosnick, Stanford University

War and Foreign Policy - Pasteur/Lafeyette
Chair: G. Donald Ferree, University of Wisconsin
Discussant: Barbara Bardes, University of Cincinnati

Combat and Confidence: The Foreign Policy Attitudes of U.S. Soldiers Engaged in the Global War on Terror
Jason Dempsey, United States Military Academy & Columbia University

American Leaders and the Public: Areas of Convergence and Discord in How They View Foreign Affairs
Alvin Richman, U.S. Department of State

The Hall of Mirrors: How the US Public and Elite Misperceive Each Other on Foreign Policy
Steven Kull, Program on International Policy Attitudes

Modeling the Support for the U.S.-Led war on Terrorism Among Egyptian Citizens: Implications for International Public Diplomacy
Michael Elasmar, Boston University

Nonresponse - Voltaire
Chair: Virginia Lesser, Oregon State University
Discussant: Young Chun, AIR

The Effect of Survey Follow-up on Nonresponse Bias: Joint Canada/United States Survey of Health, 2002-03
Stephen Blumberg, NCHS/CDC
Karen Davis, NCHS/CDC
Meena Khare, NCHS/CDC
Michael Martinez, NCHS

Analyzing Marginal Response Rates in the CAHPS Medicare Fee-for-Service Survey
Larry Campbell, RTI International
G. Gordon Brown, RTI International
Lisa Carpenter, RTI International
Linda Dimitropoulos, RTI International

Determinants and Outcomes of Initial Contact in the National Health Interview Survey, 2004
James Dahlhamer, National Center for Health Statistics
Catherine Simile, National Center for Health Statistics
Barbara Stussman, National Center for Health Statistics
Beth Taylor, National Center for Health Statistics

NHANES Converted Refusals: Are They Different from Willing Respondents in Cardiovascular Risk Factors?
Yinong Chong, Centers for Disease Control & Prevention
Margaret Carroll, Centers for Disease Control & Prevention
Vicki Burt, Centers for Disease Control & Prevention
Pat Montalvan, Westat

MEET THE AAPOR BOOK AWARD WINNER
Saturday - 11:30 a.m. – 12:15 p.m.

Meet the Author Session – Fontainebleau C-D

Author: Howard Schuman, Charlotte Steeh, Lawrence Bobo, Maria Krysan

Book: *Racial Attitudes in America: Trends and Interpretations (Harvard Univ. Press, 1997)*

AAPOR POSTER SESSION
Saturday 11:30 am-12:15 pm

Poster Session – Fontainebleau C-D

39. When a Stranger Calls: The Impact of Caller ID on Telephone Surveys
Kathy Krey, Baylor University
Jodien Matos, Baylor University

40. Who's Calling?: The Impact of Caller-ID Displays on Telephone Survey Response
Allan McCutcheon, UNL-Gallup Research Center
Jack Ludwig, Gallup Research Center

41. Strengthening Higher Education Through Gridiron Success? Public Perceptions of the Impact of National Football Championships on Academic Quality
Robert Goidel, Louisiana State University
John Maxwell Hamilton, Louisiana State University

42. Improving In-Person Data Collection in a Random Household Survey in Low- and Moderate Income Census Tracts
Ronald Bass, ORC Macro
Jaceey Sebastian, ORC Macro

43. Seeking Quality With Secondary Survey Data
Thomas R. Marshall, University of Texas, Arlington

44. Identical Study Protocol, Substantially Different Outcome Rates: Determining the Reasons
Karen Foote Retzer, University of Illinois

45. "The Art of Associating": The Central Role of Peers in Civic Life
Casey Klofstad, Harvard University

46. From Inches to Centimeters: The Uncritical Use of Available Measurements
Pablo Paras, University of Connecticut
Luis Estrada, UCSD

47. Modeling Question-Answer Processes of Structured Interviews using Interactive Hidden Markov Models
Adriaan Hoogendoorn, Vrije Universiteit
Cees Elzinga, Vrije Universiteit
Wil Dijkstra, Vrije Universiteit

48. Adapting to a Changing Data Collection Environment in an Ongoing Survey
Mark Pierzchala, Mathematica Policy Research
Tom Barton, Mathematica Policy Research
Paul Guerino, American Institutes for Research
Cheryl Hills, Mathematica Policy Research

49. When Census Data Don't Add Up in the Field: Two Case Studies
Kathleen Tiefenwerth, Battelle Centers for Public Health Research and Evaluation
Karen Tucker, Battelle Centers for Public Health Research and Evaluation

50. The Survey Organization-Respondent Relationship: A Taxonomy of Conceptualizations
James Chesire, NORC-University of Chicago
Jake Bartolone, NORC-University of Chicago

51. A Comparison of Within Household Random Selection Methods for Random Digit Dial Surveys
Molly Longstreth, University of Arkansas
Todd Shields, University of Arkansas

52. Sixteen Tons (of Paperwork) and What Do You Get? Self-Reported Workweek Length of Doctoral Scientists and Engineers
Tom Hoffer, NORC
Karen Grigorian, NORC

53. The Impact of Follow-up Contacts on Survey Data and Response Rates
Elizabeth Westin, Westat
Michele Harmon, Westat
Kerry Levin, Westat

54. You Don't Have to Accept Low Survey Response Rates - How We Achieved Record Survey Cooperation Rates
Jane Traub
Kathy Pilhuj
Daniel Mallett

55. Developing a Prescription for Physician Surveys
Janice Ballou, Mathematica Policy Research
Brian Roff, Mathematica Policy Research
Julita Milliner-Waddell, Mathematica Policy Research
Frank Potter, Mathematica Policy Research

56. Examining the Use and Effectiveness of a Help Desk in Providing Support for Self Administered Web Interviews
David Roe, RTI International
Lisa Carley-Baxter, RTI International
Jeff Franklin, RTI International
Rodney Baxter, RTI International

57. The More You Know the Less You Chat: General Political Knowledge as a Predictor of Online Chat Room Use
Masaki Hidaka, University of Pennsylvania

58. Using Proxy Reports of Quality of Life for Nursing Home Residents
Annette Totten, University of Minnesota

59. A Cumulative Count, General Approach to Item-NonResponse
Michael Wood, Hunter College, CUNY

60. Incidence and Impact of Controlled Access Situations on Unit Nonresponse
David Cunningham, RTI International
Laura Flicker, RTI International
Joseph Murphy, RTI International
Jeremy Aldworth, RTI International
Susan Myers, RTI International
Joel Kennet, Substance Abuse and Mental Health Services Administration

61. Mixed Mode Data Collection Using Paper and Web Questionnaires. A Cost and Response Rate Comparison in a Survey of Students Housing Conditions
Peter Werner, Linkping University
Gosta Forsman, Swedish Road Administration

62. Assessing the Accuracy of Respondent Reports of the Location of their Home Relative to Geographic Boundaries and Other Characteristics
John Baldridge, University of Montana
James Sylvester, University of Montana
William Borrie, University of Montana

63. Accuracy of Reports of Behavioral Health Service Use Among Public Assistance HMO Members: Results From a Record Check Study
Scott Beach, University of Pittsburgh
Janet Schlarb, University of Pittsburgh
Donald Musa, University of Pittsburgh
James Schuster, Community Care Behavioral Health

64. What's the Benefit of a Mail Supplement to a Telephone Survey for Gathering Data From "Hard To Reach" Populations?
Kerry Levin, Westat
Cynthia Helba, Westat
Barbara Forsyth, Westat
Karen Masken, IRS

65. Spanish-Speaking Households: Reaching Them Faster
Lorayn Olson, Abt Associates
Martin Frankel, Abt Associates and Baruch College, CUNY
Larry Osborn, Abt Associates
Sergei Rodkin, Abt Associates

66. Cognitive Interviewing en Espanola -- Extending the Realm of Cross-Cultural Pretesting
Martha Hunt, National Cancer Institute
Gordon Willis, National Cancer Institute

67. Modeling Components of Response Propensity in Centralized Telephone Surveys
Barbara C. O'Hare, Arbitron
Sheila Cross, Arbitron
Sonja Ziniel, University of Michigan
Robert M. Groves, University of Michigan

68. Survey Outcomes and the Characteristics of Telephone Number Exchanges
John Kennedy, Indiana Univ. Center for Survey Research
Nancy Bannister, Indiana Univ. Center for Survey Research

69. Partial Interviews and Data Quality in a Large Telephone Screening Survey
Kristen Olson, University of Michigan
Mary Cay Murray, Abt Associates
Jessica Cardoni, Abt Associates
Michael Battaglia, Abt Associates
Jim Lepkowski, University of Michigan

70. Presence of Others: Korean General Social Survey (KGSS) and General Social Survey (GSS)
Youngshil Park, Survey Research Center, SKKU

71. Making Sense of the Third-Person Effect: A Personality Approach
Yanjun Zhao, Southern Illinois University, Carbondale

72. Testing a Shorter Interview for Refusal Conversion in a Telephone Survey
Vasudha Narayanan, Westat
Sherman Edwards, Westat
Vickie Mays, UCLA
Stephanie Fry, Westat
Susan Cochran, UCLA

73. What We Think Others Think: A Motivated Reasoning Model of Public Opinion Perception
Lilach Nir, University of Pennsylvania

74. Socioeconomic Status, Race, and Ethnicity on Information Gain From Public Information Campaigns
Dianne Rucinski, University of Illinois, Chicago

75. Voter Turnout and Likely Voter Models
Joanne Miller, University of Minnesota
Lawrence Jacobs, University of Minnesota
Samuel Best, University of Connecticut
Chase Harrison, University of Connecticut

76. Are Web Options Making a Difference?
Geri Mooney, Mathematica Policy Research
Barbara Rogers, Mathematica Policy Research
Melissa Wood, Mathematica Policy Research
Deborah Trunzo, Mathematica Policy Research

12:15 p.m. – 2:00 p.m. PLATED LUNCH SESSION
- Fontaine / Fleur De Lis
(Core Meal)

2004 Exit Polls
Chair: Robert Daves, Minneapolis Star Tribune
Warren Mitofsky, Mitofsky International
Kathy Frankovic, CBS News
Fritz Scheuren, NORC

12:15 p.m. – 2:00 p.m. Chapter Representatives Meeting - Imperial I
Racial Attitudes Researchers Mtg. - Imperial IV
POQ Advisory Meeting - Imperial III
Telephone Conference Committee Mtg. - Imperial IV

AAPOR CONCURRENT SESSIONS
Saturday 2:15 pm-3:45 pm

Cognition - Fontainebleau A
Chair: Tom Guterbock, University of Virginia

Using Cognitive Interviews to Improve the Quality of Questions on Working Time in the Spanish Labor Force Survey
José Luis Padilla, University of Granada
Miguel Angel Martinez, National Statistical Institute
Teresa Moratilla, National Statistic Institute
Andres Gonzalez, University of Granada

Evaluating Argument Repertoire: Opinion Quality, Political Sophistication, and Ambivalence
Nam-Jin Lee, University of Wisconsin-Madison

Establishment Surveys from the Perspective of Socially Distributed Cognition
Boris Lorenc, Stockholm University

When Do Respondent Misconceptions Lead to Survey Response Error?
Anna Suessbrick, New School for Social Research
Michael Schober, New School for Social Research
Frederick Conrad, University of Michigan

Who is Being Heard? Item Non-Response Bias and Response Quality Bias in Open Ended Comments
Mark Andrews, ORC Macro

Consumer Expenditures/Finances - Brittany
Chair: Vince Price, Univ. of Pennsylvania
Discussant: Mark Hertzog, Hertzog Research

Relationships among Expenditure Reporting Rates, Household Characteristics and Interview Processes Variables in the U.S. Consumer Expenditure Interview Survey
John Eltinge, U.S. Bureau of Labor Statistics
Moon Jung Cho, U.S. Bureau of Labor Statistics

Learning to Say No: Conditioned Underreporting in an Expenditure Survey
Nhien To, Bureau of Labor Statistics
Jennifer Shields, Bureau of Labor Statistics

Consumer Sentiment and Consumer Spending: The Gambler's Fallacy
Nat Ehrlich, Michigan State University

Bilingual Approaches and Influences on Measurement and Data Collection: Evidence from the Survey of Consumer Finances
Man-chi Mandy Sha, NORC
Leslie Athey, NORC

Importance of Response Rates - Fontainebleau B
Chair: Robert Oldendick, Univ. of South Carolina
Discussant: Gregg Murray, University of Texas - Brownsville

How Different are Hard-To-Reach Respondents? Is it Worth the Effort to Track Them Down?
Mary Outwater, University of Oklahoma
Jeanette A. Norris, University of Oklahoma

The Relation Between Likelihood of Response and Data Quality In The American Time Use Survey
Scott Fricker, BLS

Are Lower Response Rates Hazardous for Your Health? Do Higher Response Rates Translate Into Better Estimates of Health Insurance Coverage and Access to Care?
Michael Davern, University of Minnesota
Kathleen Thiede Call, University of Minnesota
Meg Brown Good, University of Minnesota
Jeanette Ziegenfuss, University of Minnesota

High Response Rate or Better Data Quality? Examining the Trade-offs for an Establishment Survey
Lauren Harris-Kojetin, Institute for the Future of Aging Services/AAHSA
Kristen Kiefer, The National Council on the Aging

Designing Questionnaires -Champagne
Chair: Kenneth Steve, Nielsen Media Research
Discussant: Bob Belli, Univ. of Nebraska

Fostering Storytelling: A Qualitative Approach to Enhance Questionnaire Development
William Mickelson, Chamberlain Research Consultants
Jon Harrington, Chamberlain Research Consultants

12:15 p.m. – 2:00 p.m. PLATED LUNCH SESSION
- Fontaine / Fleur De Lis
(Core Meal)

2004 Exit Polls
Chair: Robert Daves, Minneapolis Star Tribune
Warren Mitofsky, Mitofsky International
Kathy Frankovic, CBS News
Fritz Scheuren, NORC

12:15 p.m. – 2:00 p.m. Chapter Representatives Meeting - Imperial I
Racial Attitudes Researchers Mtg. - Imperial IV
POQ Advisory Meeting - Imperial III
Telephone Conference Committee Mtg. - Imperial IV

AAPOR CONCURRENT SESSIONS
Saturday 2:15 pm-3:45 pm

Cognition - Fontainebleau A
Chair: Tom Guterbock, University of Virginia

Using Cognitive Interviews to Improve the Quality of Questions on Working Time in the Spanish Labor Force Survey
José Luis Padilla, University of Granada
Miguel Angel Martinez, National Statistical Institute
Teresa Moratilla, National Statistic Institute
Andres Gonzalez, University of Granada

Evaluating Argument Repertoire: Opinion Quality, Political Sophistication, and Ambivalence
Nam-Jin Lee, University of Wisconsin-Madison

Establishment Surveys from the Perspective of Socially Distributed Cognition
Boris Lorenc, Stockholm University

When Do Respondent Misconceptions Lead to Survey Response Error?
Anna Suessbrick, New School for Social Research
Michael Schober, New School for Social Research
Frederick Conrad, University of Michigan

Who is Being Heard? Item Non-Response Bias and Response Quality Bias in Open Ended Comments
Mark Andrews, ORC Macro

Consumer Expenditures/Finances - Brittany
Chair: Vince Price, Univ. of Pennsylvania
Discussant: Mark Hertzog, Hertzog Research

Relationships among Expenditure Reporting Rates, Household Characteristics and Interview Processes Variables in the U.S. Consumer Expenditure Interview Survey
John Eltinge, U.S. Bureau of Labor Statistics
Moon Jung Cho, U.S. Bureau of Labor Statistics

Learning to Say No: Conditioned Underreporting in an Expenditure Survey
Nhien To, Bureau of Labor Statistics
Jennifer Shields, Bureau of Labor Statistics

Consumer Sentiment and Consumer Spending: The Gambler's Fallacy
Nat Ehrlich, Michigan State University

Bilingual Approaches and Influences on Measurement and Data Collection: Evidence from the Survey of Consumer Finances
Man-chi Mandy Sha, NORC
Leslie Athey, NORC

Importance of Response Rates - Fontainebleau B
Chair: Robert Oldendick, Univ. of South Carolina
Discussant: Gregg Murray, University of Texas - Brownsville

How Different are Hard-To-Reach Respondents? Is it Worth the Effort to Track Them Down?
Mary Outwater, University of Oklahoma
Jeanette A. Norris, University of Oklahoma

The Relation Between Likelihood of Response and Data Quality In The American Time Use Survey
Scott Fricker, BLS

Are Lower Response Rates Hazardous for Your Health? Do Higher Response Rates Translate Into Better Estimates of Health Insurance Coverage and Access to Care?
Michael Davern, University of Minnesota
Kathleen Thiede Call, University of Minnesota
Meg Brown Good, University of Minnesota
Jeanette Ziegenfuss, University of Minnesota

High Response Rate or Better Data Quality? Examining the Trade-offs for an Establishment Survey
Lauren Harris-Kojetin, Institute for the Future of Aging Services/AAHSA
Kristen Kiefer, The National Council on the Aging

Designing Questionnaires -Champagne
Chair: Kenneth Steve, Nielsen Media Research
Discussant: Bob Belli, Univ. of Nebraska

Fostering Storytelling: A Qualitative Approach to Enhance Questionnaire Development
William Mickelson, Chamberlain Research Consultants
Jon Harrington, Chamberlain Research Consultants

A Picture is Worth a Thousand Words: Results from Cognitive Testing of Graphic Show Cards
Anne Hartman, National Cancer Institute
Jennifer Crafts, Westat
Suzanne McNutt, Westat
Gordon Willis, National Cancer Institute
Tracey Summerall, Westat
Amy Yaroch, National Cancer Institute

Using Cognitive Procedures and Traditional Pilot Studies to Improve the Quality of Questions Concerning Earnings in the Spanish Labor Force Survey
José Luis Padilla, University of Granada
Miguel Angel Martinez, National Statistical Institute
Teresa Moratilla, National Statistical Institute
Andres Gonzalez, University of Granada
Isabel Esteban, National Statistic Institute

The Importance of Properly Measuring Importance
Amy Gershkoff, Princeton University
(AAPOR Student Paper Award Winner)

Voter Intent and Choice - Lemans/Bordeaux
Chair: Jon Miller, Northwestern Univ. Medical School
Discussant: Jay Mattlin, MediaMark Research

To Vote or Not to Vote?: A Comparison of Vote Intention Measures
Randall Thomas, Harris Interactive
Mitchell Sanders, Harris Interactive
Renee Smith, Harris Interactive
Susan Behnke, Harris Interactive

Voter Intent, Voting Technology and Measurement Error
Frederick Conrad, University of Michigan
Emilia Peytcheva, University of Michigan
Michael Traugott, University of Michigan
MIchael Hanmer, Georgetown University
Paul Herrnson, University of Maryland
Ben Bederson, University of Maryland
Richard Niemi, University of Rochester

Measures of difficulty in election polling
Joseph Shipman, SurveyUSA
Jay Leve, SurveyUSA

Moment by Moment: Analyzing the Presidential Debates
Anthony Salvanto, CBS News

Cross-National Research II - Burgundy
Chair: Caroline Roberts, London School of Economics & Political Science
Discussant: Martha Hunt, National Cancer Institute

The Changing Patterns of Religious Practice and Belief in Western Europe: A Cross-National Cohort Analysis
Allan McCutcheon, UNL-Gallup Research Center

Equivalence of the Schwartz Value System across 20 European countries - Results from the ESS 2002
Peter Ph. Mohler, ZUMA
Katrin Wohn, ZUMA

Cross-National Measurement of the Religion Effects and Its Application on Public Attitude Toward Science and Technology in European New Member Countries
Dan Liao, University of Nebraska, Lincoln
David Palmer, Gallup Research Center
Lingrui Jiang, Gallup Research Center
Xiaoming Liu, Gallup Research Center
Allan McCutcheon, Gallup Research Center

Attitude towards Science and Technology In European New Member Countries: A Cross-National Study
Dan Liao, University of Nebraska, Lincoln
Lingrui Jiang, Gallup Research Center
David Palmer, Gallup Research Center
Xiaoming Liu, Gallup Research Center
Allan McCutcheon, Gallup Research Center

Media Research - Pasteur/Lafeyette
Chair: Robert Lee, University of California - Berkeley

Using Public Opinion Polling to Evaluate Mass Media Health Promotion Campaigns: The Good, the Bad and The Not-so-Good
Corinne Hodgson, Corinne S. Hodgson & Associates

Who Learns From The News? The Moderating Role of Cognitive Processing Styles in Learning From The News
Clarissa David, University of Pennsylvania

News discrepancy Perception and News Credibility Judgment: The Role of the Self as a Comparison Anchor in Judgmental process of News Credibility
Hyunseo Hwang, University of Wisconsin-Madison
Gun Hyuk Lee, Changwon National University
Sung Gwan Park, Seoul National University

News Discrepancy and Information Search: The Effects of News Slants on Audiences' Information Search Patterns Over the Internet
Hyunseo Hwang, University of Wisconsin-Madison
Kwangjun Heo, University of Wisconsin-Madison
Sun-Young Lee, University of Wisconsin-Madison

Does Television Change Us? An Analysis of Three Experiments
Bethany Albertson, University of Chicago
(AAPOR Student Paper Award Honorable Mention)

How AAPOR Can Help Meet the Needs of Mid-Career Members - Voltaire
Chair: Sandra Berry, RAND

AAPOR CONCURRENT SESSIONS
Saturday 4:00 pm-5:00 pm

Standards - Fontainebleau A
Chair: Sandy Berry, RAND

ISO Standards for Market, Opinion, and Social Research: A Preview
Tom W. Smith, NORC, University of Chicago

Professional Standards: New Challenges, Changing Ethics?
Diane Bowers, CASRO

Reporting Standards for Internet Surveys and Polls
Richard Tychansky, University of Toronto

War in Iraq - Brittany
Chair: Karl Feld, RTI International

Support for the War in Iraq: American Casualties as a Survey Item
Mark West, U. North Carolina at Asheville
Donald Diefenbach, U. North Carolina at Asheville

Foreign Policy and America's Place in the World: Generational Divisions on U.S. Priorities and Iraq
Nicole Speulda, Pew Research Center

Satisficing - Fontainebleau B
Chair: Rob Eisinger, Lewis & Clark College

Effect of Respondent Motivation and Task Difficulty on Nondifferentiation in Ratings: A Test of Satisficing Theory Predictions
Sowmya Anand, Ohio State University
Jon Krosnick, Stanford University
Kenneth Mulligan, University of Georgia
Wendy Smith, Research by Design
Melanie Green, University of Pennsylvania
George Bizer, Eastern Illinois University

Comparing Major Survey Firms in Terms of Survey Satisficing: Telephone and Internet Data Collection
Jon Krosnick, Stanford University
Norman Nie, Stanford University
Douglas Rivers, Stanford University and Polimetrix

Effect of Interview Pace and Items Position on Satisficing
Michael Lemay, Joint Program in Survey Methodology
Scott Fricker, Joint Program in Survey Methodology
Mirta Galesic, Joint Program in Survey Methodology
Roger Tourangeau, Joint Program in Survey Methodology
Ting Yan, Joint Program in Survey Methodology

Genes and Biotech - Champagne
Chair: Mike Dennis, Knowledge Networks

Knowledge, Communication and Opinion Formation: Comparing Attitudes Regarding Genetically Modified Foods
Gregg Murray, University of Texas at Brownsville
Susan Banducci, Texas Tech University
Jeff Karp, Texas Tech University

Biotechnology and Trusts: What Drives Citizens to Support Biotechnology?
Zuoming Wang, Cornell University
Dietram Scheufele, University of Wisconsin, Madison
Dominique Brossard, University of Wisconsin, Madison

How Clean is the Air You Breathe?: Predicting Beliefs about Regional Air Quality
Donald Diefenbach, U. North Carolina at Asheville
Mark West, U. North Carolina at Asheville

Terrorism - Lemans/Bordeaux
Chair: Stanley Feldman, SUNY Stony Brook

Psychological and Political Responses to the Madrid Terrorist Attacks
Kenneth Rasinski, NORC, University of Chicago
Tom W. Smith, NORC, University of Chicago
Juan Diez-Nicolas, Analisis Sociologicos Economicos Y Politicos

Willing to Believe: Explaining the Belief that Saddam Hussein Aided with the September 11th Attacks
Russ Tisinger, University of Pennsylvania

Public Prudence, the Policy Salience of Terrorism and Presidential Approval following Terrorist Incidents
William Josiger, Georgetown University
George Shambaugh, Georgetown University

What is Terrorism? Challenges to Improving the Construct Validity of a Complex Concept
Stephanie Willson, National Center for Health Statistics
Kristen Miller, National Center for Health Statistics

Order and Context - Pasteur/Lafeyette
Chair: John Kennedy, Indiana University

Response Order Effects in Online Surveys
Randall Thomas, Harris Interactive
Susan Behnke, Harris Interactive
Alyssa Johnson, Harris Interactive

Causes of Context Effects: How Questionnaire Layout Induces Measurement Error
Andy Peytchev, University of Michigan
Roger Tourangeau, Joint Program in Survey Methodology

Context, Race, Gender and Policy Considerations
David Wilson, Gallup/Michigan State University

Detroit Area Survey - Burgundy
Elizabeth Martin, U.S. Census Bureau
Diane Colasanto, Consultant
Mick Couper, University of Michigan

The Detroit Area Study was established in 1952 as a survey research practicum associated with the Department of Sociology's graduate program at the University of Michigan. For the next 50 years, graduate students and their faculty advisors carried out annual surveys of metropolitan Detroit. The DAS was abolished in 2004. This session commemorates the DAS with brief presentations on its contributions to the study of racial attitudes, survey methodology, student training, and other topics, as well as reminiscences from former DAS students, teaching fellows, and directors.

Blogs and Bloggers 2004: Extending the Reach, Influence, Understanding and Misperceptions of Pre-Election and Exit Polls - Voltaire
Chair: Nancy Belden, Belden, Russonello & Stewart

Panelists:
Mark Blumenthal, Mystery Pollster
Steve Soto, The Left Coaster
Joel Bloom, Oregon Survey Research Lab
Tom Bevan, RealClear Politics
Tom McIntyre, RealClear Politics

5:00 p.m. – 6:30 p.m. AAPOR MEMBERSHIP & BUSINESS MEETING - Voltaire

6:45 p.m. – 7:30 p.m. PRESIDENT'S RECEPTION: Meet and mingle with President Nancy Belden and the AAPOR Executive Council. – Club Atlantic

Hors d'oeuvres sponsored by: Marketing Systems Group

7:30 p.m. – 9:15 p.m. AWARDS BANQUET (Core Meal) - Fontaine/Fleur De Lis

Wine sponsored by: Pulse Train
Awards Program Booklet sponsored by: ORC Macro
PowerPoint Presentation jointly sponsored by: SPSS & RTI
Centerpieces sponsored by: CfMC
Head table Centerpieces sponsored by: Oxford Press

9:30 p.m. – 11:30 p.m. ANNUAL BOOK SALE - Fontaniebleau C-D

10:00 p.m. - ? APPLIED PROBABILITY - Imperial I

SUNDAY, MAY 15

7:00 a.m. – 11:30 a.m. AAPOR REGISTRATION DESK OPEN - Grand Gallerie

7:00 a.m. – 9:00 a.m. Breakfast - Cash & Carry in the Grand Gallerie

7:00 a.m. – 9:00 a.m. Survey Faculty Meeting – Imperial I

AAPOR SHORT COURSE
Sunday - 8:30 a.m. – 12:00 noon.

It Takes More Than Words to Write a Question: How Visual Layout Affects Answers to Mail and Internet Surveys – Imperial III
Don Dillman

AAPOR CONCURRENT SESSION
Sunday 8:15 am-9:45 am

Ideology and Issues - Voltaire
Chair: David Moore, Gallup Organization
Discussant: Nat Ehrlich, Michigan State University

Is Ideological Self-Identification in Internet Speech Silenced by Opinion Opposing the Speaker?
David Fan, University of Minnesota

Ideology, Issue Preferences, and Political Choice: The Paradox of "Conflicted Conservatives"
Christopher Ellis, UNC-Chapel Hill

The Future of the First Amendment: A National Survey of High School Students
Kenneth Dautrich, University of Connecticut
John Bare, Blank Family Foundation
David Yalof, University of Connecticut

Surveys of Children and Youth - Fontainebleau B
Chair: Lois Timms-Ferrara, Roper Center
Discussant: Kristin Stettler, U.S. Census Bureau

Children and Juveniles as Respondents - Results from Field Experiments on Measurement Error
Marek Fuchs, University of Kassel

Evaluating Sources of Variation In Survey Estimates of Children with Disabilities Using Federal Education Data
Stacey Bielick, ESSI/AIR
Jennifer Park, National Center for Education Statistics
Elvira Hausken, National Center for Education Statistics

Factors Contributing to Assessment Burden in Preschoolers
Cassandra Rowand, Mathematica Policy Research
Susan Sprachman, Mathematica Policy Research
Holly Rhodes, RTI International
Ina Wallace, RTI International

Evaluating Differences in Health Care Quality Ratings in Children Based on the Level of Proxy Respondent Involvement with Child's Care
Timothy Beebe, Mayo Clinic
Walter Suarez, Midwest Center for HIPAA Education
Kathleen Thiede Call, University of Minnesota

Internet and Web Surveys II - Brittany
Chair: James Newswanger, IBM
Discussant: Reg Baker, Market Strategies

The Market Value Survey: Ensuring Quality on a Government Web Based Survey
John Flatley, Office for National Statistics
Dave Ruston, Office for National Statistics

Using Portal Technology to Improve Quality at the U.S. Census Bureau-Presentation and Demonstration of Portal Technology
Cheryl Landman, U.S. Census Bureau
Deborah Stempowski, U.S. Census Bureau

A Comparison of an Online Card Sorting Task to a Rating Task
Randall Thomas, Harris Interactive
Leonard Bayer, Harris Interactive
Alyssa Johnson, Harris Interactive
Susan Behnke, Harris Interactive

Cognitive Interviewing - Champagne
Chair: Ashley Landreth, U.S. Census Bureau

Cognitive Interview Usability Testing of the Redesigned Sudden Unexplained Infant Death Reporting Form
Danna L. Moore, Washington State University

Cognitive Interviewing: the case of verbal reports
Erik Hjermstad, University of Nebraska-Lincoln
Monica Sanchez, University of Nebraska-Lincoln

A Cognitive Tool to Predict How Well Respondents Perform Survey-Related Tasks
Patricia A. Gwartney, University of Oregon

Designing Items to Measure Healthy Marriage: Lessons Learned from Cognitive Interviews
Lina Guzman, Child Trends
Kristin Moore, Child Trends
Greg Matthews, Child Trends
Zakia Redd, Child Trends

Does Cognitive Interviewing Really Improve Survey Quality?
Stanley Freedman, Energy Information Administration
Robert Rutchik, Energy Information Administration
Kara Norman, Energy Information Administration

Incentives III - Lemans/Bordeaux
Chair: Cliff Young, IPSOS-Opinion
Discussant: Allan McCutheon, Univ. of Nebraska

Unintended Consequences of Incentive Induced Response Rate Differences
Duston Pope, Market Strategies
Scott D. Crawford, Survey Sciences Group
Eric O. Johnson, Henry Ford Health Science Center
Sean Esteban McCabe, University of Michigan

The Use of Monetary Incentives in the Survey of Income and Program Participation
Denise Lewis, U.S. Census Bureau
Kathleen Creighton, U.S. Census Bureau

The Impact of Providing Incentives to Initial Refusal Cases on Sample Composition and Data Quality in a Telephone Survey
Douglas Currivan, RTI International

Incentives: Do You Get What You Pay For?
Ronald Broach, NORC
Karen Grigorian, NORC

Voting: Foreign and Domestic - Burgundy
Chair: Volker Stocke, University of Mannheim
Discussant: Patrick Murray, Rutgers University

Economic Perceptions and Voting Behaviors in the Second Swedish Election to the EU-Parliament
Lingrui Jiang, UNL-Gallup Research Center
Yu Feng, UNL-Gallup Research Center
Allan McCutcheon, UNL-Gallup Research Center

Response Privacy and Elapsed Time Since Election Day as Determinants for Vote Overreporting
Volker Stocké, University of Mannheim

A Comparison of Presidential Candidate Vote Intention Measures in U.S. Elections
Randall Thomas, Harris Interactive
David Krane, Harris Interactive
Mitchell Sanders, Harris Interactive
Susan Behnke, Harris Interactive

A Failing Tool or Changing Society? The Validity of the "Strength-of-Personality Scale" and the Role of Opinion Leaders in the 2002 German Federal Election
Thomas Petersen, Institut für Demoskopie Allensbach

Self-reports and Health - Pasteur/Lafeyette
Chair: Cecilia Gaziano, Research Solutions
Discussant: Meg Good, University of Michigan

Assessing the Quality of Height and Weight Data Collected During Telephone Interviews
Keith Smith, Abt Associates
Marjorie Morrissey, Abt Associates

Designing Self-Administered Questionnaires for Health Center Personnel to Evaluate a Quality Improvement Initiative
Jessica Graber, University of Chicago
Anne Kirchhoff, University of Chicago

Reliability of Self-Reported Conditions in Health Surveys: Findings From the Army Chemical Corps Health Survey
John Boyle, Schulman, Ronca & Bucuvalas

Aging and Errors of Measurement in Surveys - The Case of Self-Reports of Health
Duane Alwin, Pennsylvania State University
Ryan McCammon, University of Michigan
Willard Rodgers, University of Michigan
Halimah Hassan, University of Michigan

Laying the Groundwork for Including Cell Phone Numbers in Telephone Survey Samples - Fontainebleau A

Chair: Brian J. Meekins, Bureau of Labor Statistics

Final Results from the 2004 CPS Supplement on Telephone Usage
Clyde Tucker, Bureau of Labor Statistics
Michael Brick, Westat
Brian J. Meekins, Bureau of Labor Statistics

Surveying Cell Phone Households - Results and Lessons?
Y. Angela Yuan, Westat
Bruce Allen, Westat
Michael Brick, Westat
Sarah Dipko, Westat
Stanley Presser, Joint Program in Survey Methodology
Clyde Tucker, Bureau of Labor Statistics
Daifeng Han, Joint Program in Survey Methodology
Laura Burns, Joint Program in Survey Methodology
Mirta Galesic, Joint Program in Survey Methodology

An Investigation of Response Differences between Cell Phone and Landline Interviews
Sarah Dipko, Westat
Pat Dean Brick, Westat
Michael Brick, Westat
Stanley Presser, Joint Program in Survey Methodology

AAPOR CONCURRENT SESSIONS
Sunday 10:00 am-11:30 am

Mail Surveys and Advance Contact - Fontainebleau A

Chair: Jane Traub, Scarborough Research
Discussant: Keith Neuman, Environics Research Group

Does Envelope Size Affect Respondent Participation?
Kerry Levin, Westat
Martha Kudela, Westat
Karen Masken, IRS

Mail Survey Non-Contact Identification: Who Did We Not Get Through To?
Benjamin Healey, Massey University
Philip Gendall, Massey University

Testing an Advance Contact Targeted-Awareness Campaign to Raise Response Rates
Ana Patricia Melgar, Nielsen Media Research
Paul J. Lavrakas, Nielsen Media Research
Agnieszka Flizik, Nielsen Media Research
Rosemary Holden, Nielsen Media Research
Maria Anatro, Nielsen Media Research

Is it Worth the Effort: RDD Telephone Surveys and Advance Survey Notification
Leslyn Hall, ORC Macro
Randall ZuWallack, ORC Macro
Kirsten Ivie, ORC Macro

Modes and Methods - Fontainebleau B

Chair: Lars Kaczmirek, ZUMA

Web Survey Response Rates Compared to Other Modes - A Meta-Analysis
Katja Lozar-Manfreda, University of Ljubljana
Michael Bosnjak, University of Mannheim
Iris Haas, University of Mannheim
Vasja Vehovar, University of Ljubljana

Mode Effects in Customer Satisfaction Measurement
Darby Steiger, Gallup Organization
Linda Keil, Gallup Organization
Greg Gaertner, Gallup Organization

Prompting Efforts to Raise Response Rates for a Web-Based Survey
Lekha Venkataraman, NORC
Maggie Parker, NORC

An Investigation of the Impact of Departures of Standardized Interviewing on Response Errors in Self-Reports about Child Support and Other Family-Related Variables
Jennifer L. Dykema, University of Wisconsin, Madison
(AAPOR Student Paper Award Winner)

Mobilizing Information Online: The Effects of Primary-Source and Secondary-Source Website Use on Political Engagement
Bruce Hardy, Cornell University
(MAPOR Student Paper Award Winner)

Interviewers - Brittany

Chair: Dawn Nelson, U.S. Bureau of Census
Discussant: Jim Wolf, Indiana Univ. Public Opinion Lab

Person-Based Data Collection in Practice: An Evaluation of Interviewer/Respondent Interactions
Jennifer Hunter, U.S. Census Bureau
Ashley Landreth, U.S. Census Bureau

Impact of Interviewer Effects on Survey Inference
Trivellore Raghunathan, University of Michigan

Interviewer Judgments about the Quality of Telephone Interviews
John Tarnai, Washington State University
M. Chris Paxson, Washington State University

Interviewer Intervention for Data Quality in the 2004 Survey of Consumer Finances
Yongyi Wang, NORC - University of Chicago
Steven Pedlow, NORC - University of Chicago

Measuring Race and Ethnicity - Champagne
Chair: Diane Willimack, U.S. Census Bureau

Can a Survey Change One's Race? An Experiment on Context Effects and Racial Self-Classification
Thomas Craemer, Stony Brook University

A Comparison of Single and Multiple Race Category Choices in NLSY97, Round 6
Parvati Krishnamurty, NORC
Lidan Luo, NORC

Measuring Ethnic Identity Within a Community-Based Sample of Black American Adults
Rachel E. Davis, University of Michigan
Abdul R. Shaikh, University of Michigan
Guangyu Zhang, University of Michigan
Ken Resnicow, University of Michigan

Black, White, Other? Why Racial Identifications within Hispanic Ethnicity Matter; A Look at Professional and Social Outcomes Among Research Doctorate Recipients
Vincent Welch, Jr., NORC
Kim Williams, NORC

New Research on the Differences Between Hispanic "Origin" and Hispanic "Identity" and Their Implications
Paul J. Lavrakas, Nielsen Media Research
Matt Courser, PIRE
Lillian Diaz-Castillo, Westat

Political Issues II - Lemans/Bordeaux
Chair: Adam Gluck, ORC Marco International

Interpreting the Meaning of the Presidential Approval Question
George Bishop, University of Cincinnati

Thinking and Feeling the Partisan Gender Gap
Barry Burden, Harvard University

Behavioral Self-report Measures of Non-political and Political Contributions
Randall Thomas, Harris Interactive
Jonathan Klein, University of Rochester
Mitchell Sanders, Harris Interactive
Lisa Wilding, Harris Interactive

Political Knowledge Under Respondent-Friendly Conditions
Markus Prior, Princeton University
Arthur Lupia, University of Michigan

The Role of Opinion Polls in the 2004 Philippine Elections
Mahar Mangahas, Social Weather Stations
Linda Luz Guerrero, Social Weather Stations
Gerardo Sandoval, Social Weather Stations

Political Party Identification - Burgundy
Chair: Nick Moon, NOP Social & Political
Discussant: Sheldon Gawiser, NBC News

Projection and Partisanship: The Effects of Party Identification on Opinion Processing
Kate Kenski, University of Pennsylvania
Dannagal Young, University of Pennsylvania

Attitude or Demographic? Measurement Effect on Party Identification Estimates in Preelection and Public Policy Polls
Robert Daves, Star Tribune

Party Identification and the 2004 Election
Larry Hugick, Princeton Survey Research Associates
Stacy DiAngelo, Princeton Survey Research Associates

Weighting by Party ID: When It Comes to Elections, Can It Really Be Tossed Out the Door?
Tatiana Koudinova, TIPP Poll/Technometrica Market Intelligence
Raghavan Mayur, TIPP

Cognition and Health - Pasteur/Lafeyette
Chair: David Johnson, Penn. State University
Discussant: Stephen Cohen, Bureau of Labor Statistics

Using Screening Questions to Identify Persons with Mobility Impairment: Field Test Results
Vickie Stringfellow, University of MA Boston
Patricia Gallagher, University of MA Boston

Qualified Answers and Other Doubt Expressions as Indicators of Cognitive Problems in a Health Survey
Stasja Draisma, Vrije Universiteit
Wil Dijkstra, Vrije Universiteit
Yfke Ongena, Vrije Universiteit

Identifying Mismatches Between Common Sense and Technical Definitions on an HIV Risk Behavior Survey
Lisa Moses, NCHS
Paul Beatty, Centers for Disease Control and Prevention

Applying Cognitive Psychological Principles to the Improvement of Survey Data: A Case Study from the National Survey on Drug Use and Health
Joel Kennet, Substance Abuse and Mental Health Services Administration
Dicy Painter, Substance Abuse and Mental Health Services Administration
Peggy Barker, Substance Abuse and Mental Health Services Administration
Jeremy Aldworth, RTI International

The Government and Surveys - Voltaire
Chair: Jennifer Rothgeb, U.S. Census Bureau
Discussant: Jack Fowler, UMass

Making Health Policy at the Ballot-Box: Public Attitudes Toward Five Health Propositions and the Future of Health Policy
Kristy Michaud, Public Policy Institute of California
Renatta DeFever, Public Policy Institute of California
Jon Cohen, ABC News

A Quantitative and Qualitative Approach to Gauging Public Confidence in Official Statistics in a British Environment
Amanda Wilmot, Office for National Statistics, UK
Jacqui Jones, Office for National Statistics, UK

Elements Behind Trust in Government, an Explanatory Model
Rene Bautista, Gallup Research Center
David Palmer, Gallup Research Center

Neighborhoods Matter: Fixes and Random Effects in Police Satisfaction Surveys in the UK
Rachel Dinkes, AIR
(DC/AAPOR Student Paper Award Winner)

Thank you for attending the 2005 Annual Conference. We hope to see you next year in Montreal!

SAVE THE DATES

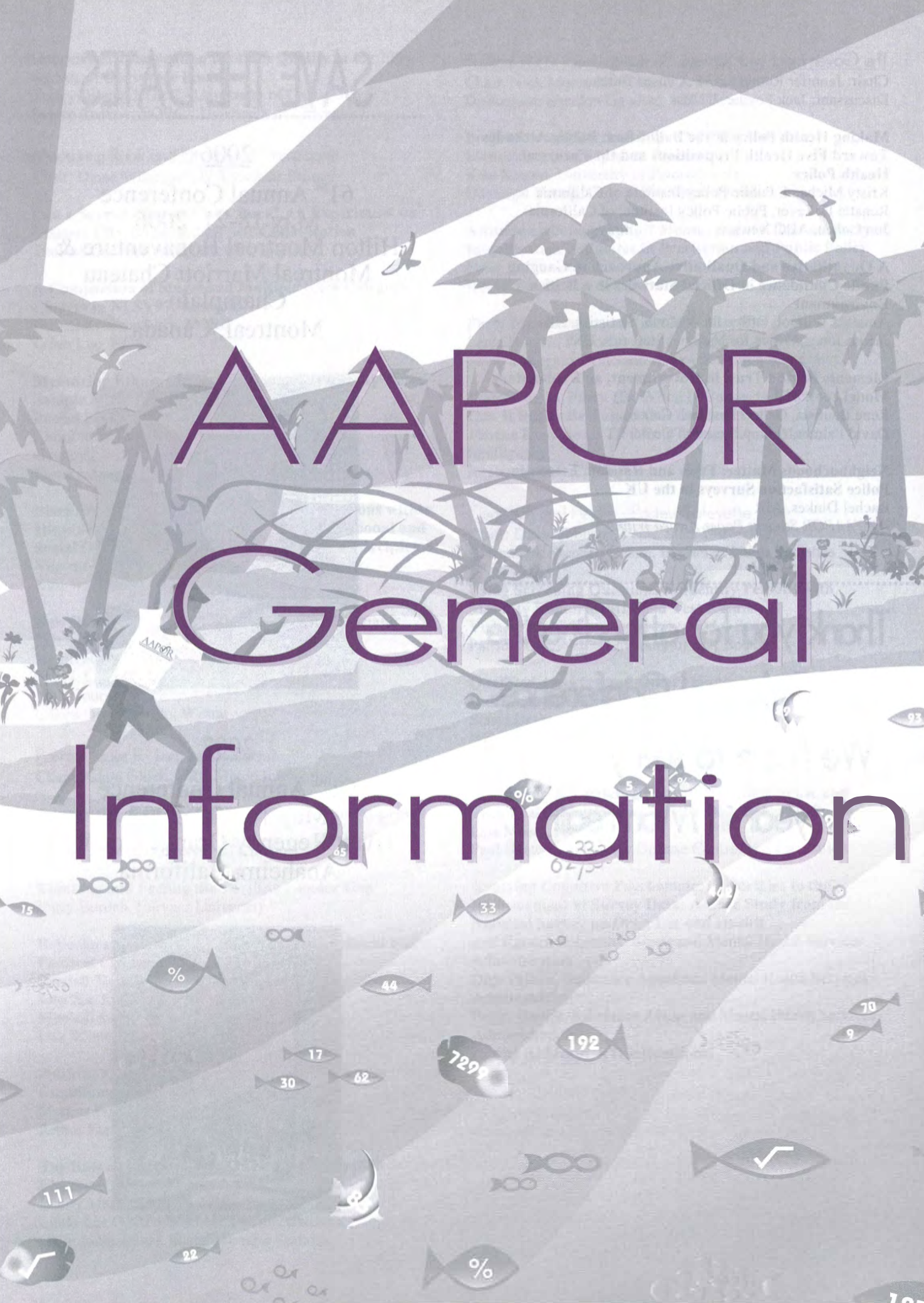
2006

61st Annual Conference
May 18-21, 2006
Hilton Montreal Bonaventure &
Montreal Marriott Chateau Champlain
Montreal, Canada

2007

62nd Annual Conference
May 16 - 20, 2007
Hyatt Regency Orange County
Anaheim, California

Hyatt Regency Orange County



AAPOR General Information

AAPOR Executive Council: 2004/2005

President
Nancy Belden, Belden Russonello & Stewart

Vice President
Cliff Zukin, Rutgers University

Past President
Elizabeth Martin, U.S. Bureau of the Census

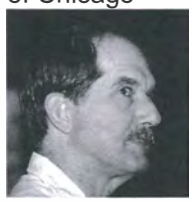
Secretary - Treasurer
Sandra H. Berry, RAND Survey Research Group

Associate Secretary- Treasurer
Jennifer M. Rothgeb, U.S. Bureau of the Census

Standards Chair
Mick P. Couper, University of Michigan

Associate Standards Chair
Nancy Mathiowetz, University of Wisconsin - Milwaukee

Conference Chair
Tom W. Smith, NORC, University of Chicago



Associate Conference Chair
David Moore, Gallup Organization

Membership & Chapter Relations Chair
Mollyann Brodie, Henry J. Kaiser Family Foundation

Associate Membership & Chapter Relations Chair
Brad Edwards, Westat

Publications & Information Chair
Daniel M. Merkle, ABC News

Associate Publications & Information Chair
Shapard Wolf, Arizona State University

Councilors-at-Large
Robert M. Groves, University of Michigan

Susan H. Pinkus, Los Angeles Times

Photos Compliments of Steve Everett

AAPOR Executive Council: 2005/2006

President
Cliff Zukin, Rutgers University

Vice President
Robert P. Daves, Minneapolis Star Tribune

Secretary - Treasurer
Jennifer M. Rothgeb, U.S. Census Bureau

Associate Secretary-Treasurer
Paul Beatty, Centers for Disease Control & Prevention

Past President
Nancy Belden, Belden Russonello & Stewart

Standards Chair
Nancy Mathiowetz, Univ. of Wisconsin - Milwaukee

Associate Standards Chair
Tom Guterbock, University of Virginia

Conference Chair
David Moore, Gallup Organization

Associate Conference Chair
Patricia Moy, University of Washington

Membership & Chapter Relations Chair
Brad Edwards, Westat

Associate Membership & Chapter Relations Chair
Katherine "Kat" Lind Draughon, Draughon Research

Publications & Information Chair
Shapard Wolf, Arizona State Univ.

Associate Publications & Information Chair
Steve Everett, Everett Group

Councilors-at-Large
Susan H. Pinkus, L.A. Times

Robert Y. Shapiro, Columbia University

AAPOR Executive Office Staff

Michael P. Flanagan, CAE
Executive Coordinator
mflanagan@goAMP.com

Monica A. Frihart
Association Assistant
mfrihart@goAMP.com

Missy M. Johnson, CMP
Meeting Manager
mjohnson@goAMP.com

Tara Nelson
Data Administrator
tnelson@goAMP.com
NO PICTURE AVILABLE

Linda Collins
Manager of Accounting Services
Lcollins@goAMP.com

Kathy Ramseier
Accounting Associate
Kramseier@goAMP.com

AAPOR Past Presidents

1947-48 Clyde W. Hart
1948-49 Elmo C. Wilson
1949-50 Paul F. Lazarsfeld
1950-51 Julian L. Woodward
1951-52 Bernard Berelson
1952-53 Archibald M. Crossley
1953-54 Samuel A. Stouffer
1954-55 George H. Gallup
1955-56 Harry Alpert
1956-57 Gerhart D. Wiebe
1957-58 Frederick F. Stephan
1958-59 David Wallace
1959-60 Herbert H. Hyman
1960-61 Robert O. Carlson
1961-62 John W. Riley, Jr.
1962-63 Joseph T. Klapper
1963-64 Charles Y. Glock
1964-65 Herbert E. Krugman
1965-66 Raymond A. Bauer
1966-67 Leo Bogart
1967-68 Paul B. Sheatsley
1968-69 Robert N. Ford
1969-70 Robert T. Bower
1970-71 Richard H. Baxter
1971-72 W. Phillips Davison
1972-73 Sidney Hollander
1973-74 Harold Mendelsohn
1974-75 Herbert I. Abelson
1975-76 Richard Maisel
1976-77 Irving Crespi
1977-78 Hope Lunin Klapper
1978-79 Reuben Cohen
1979-80 Jack Elinson
1980-81 Helen J. Kaufmann
1981-82 Seymour Sudman
1982-83 Burns W. Roper
1983-84 Laure M. Sharp
1984-85 Albert E. Gollin
1985-86 Howard Schuman
1986-87 J. Ronald Milavsky
1987-88 Eleanor Singer
1988-89 Warren J. Mitofsky
1989-90 Philip Meyer
1990-91 Joan S. Black
1991-92 Norman M. Bradburn
1992-93 Kathleen A. Frankovic
1993-94 Stanley Presser
1994-95 Andrew Kohut
1995-96 Robert M. Groves
1996-97 Diane Colasanto
1997-98 James R. Beniger
1998-99 Michael R. Kagay
1999-00 Michael W. Traugott
2000-01 Murray Edelman
2001-02 Don A. Dillman
2002-03 Mark A. Schulman
2003-04 Elizabeth (Betsy) Martin
2004-05 Nancy Belden

AAPOR Honorary Life Members

Dr. Herbert I. Abelson
Dr. Valentine Appel
Prof. Morris Axelrod
Ms. Rena Bartos
Dr. Richard H. Baxter
Mr. Joe Belden
Dr. Leo Bogart
Dr. Ann F. Brunswick
Mr. Peter Case
Mr. Reuben Cohen
Ms. Helen M. Crossley
Mr. Morton A. David
Mrs. Diana Druker
Mr. Henry L. Dursin
Dr. Alex Edelstein
Dr. Jack S. Elinson
Mr. Kendrick S. Few
Mr. Mervin Field
Mr. Raymond Fink
Mr. Robert N. Ford
Dr. G. Ray Funkhouser
Dr. Bertram Gold
Dr. Melvin A. Goldberg
Mrs. Toby Goldberg
Mr. Fred H. Goldner
Mr. Harold F. Gosnell
Dr. Lester P. Guest
Dr. Pearl H. Guest
Mr. Roger Karr Harter
Ms. Babette Kass Himler
Dr. Joseph R. Hochstim
Mr. Sidney Hollander
Mr. K. David Inouye
Prof. Daniel Katz
Ms. Helen J. Kaufmann
Dr. Hope Lunin Klapper
Dr. Herbert E. Krugman
Prof. Gladys Engel Lang
Dr. Kurt Lang
Dr. Gove P. Laybourn, Jr.
Ms. Barbara Lee
Dr. Seymour Lieberman
Prof. Robert G. Mason
Dr. Harold Mendelsohn
Dr. J. Ronald Milavsky
Ms. Selma F. Monsky
Prof. Paul Neurath
Mr. William Nicholls, II
Dr. Robert C. Nuckols
Mr. Harry W.O'Neill
Ms. Lorna Opatow
Mr. Frank E. Orenstein
Dr. C. Robert Pace
Dr. Alice Padawer-Singer
Mr. John Shelton Reed
Mr. Henry W. Riecken
Dr. John W. Riley, Jr.
Mr. Glenn H. Roberts
Dr. Theresa F. Rogers
Dr. Bernard Roshco
Dr. Sydney Roslow
Mr. James M. Sears
Dr. David L. Sills
Mr. Edwin H. Sonnecken
Dr. Joe L. Spaeth
Mrs. Mary A. Spaeth
Mr. Larry Ulin
Ms. Marilyn Watts
Ms. Margaret Weidenhamer
Dr. Tibor Weiss
Mr. Kenneth A. Wells
Dr. Gerhart D. Wiebe
Mrs. Isabelle N. Woodrow
Ms. Anne Schuetz Zanes
Ms. Ilse Zeisel
Mrs. Pearl R. Zinner

Thank you for your continued support!

AAPOR Committees

AAPOR thanks the many volunteers who have committed countless hours of behalf of our organization!

Conference Operations Committee

Katherine "Kat" Lind Draughon, Draughon Research
Operations Committee Chair

Nancy L. Whelchel, North Carolina State University
Associate Operations Committee Chair

Linda L. Dimitropoulos, Conference Social Activities Chair	RTI International
Karl Feld, Development Chair	RTI International
Patricia Moy, Book & Technology Exhibit Chair	University of Washington
Carl Ramirez, Audio Visual Chair	U.S. General Accounting Office

2004-2005 Conference Committee

Scott Althaus	University of Illinois	Lew Horner	Ohio State University
Paul Beatty	CDC	Carla Jackson	Schulman, Ronca & Bucuvalas
John Benson	Harvard School of Public Health	Timothy Johnson	Univ. of Illinois at Chicago
Paul Biemer	RTI International	Scott Keeter	Pew Research Center
George Bishop	University of Cincinnati	Courtney Kennedy	University of Michigan
Johnny Blair	Abt Associates	Jibum Kim	NORC, Univ. of Chicago
Michael Brick	Westat	Gary Langer	ABC News
Donald Camburn	RTI International	Ronald E. Langley	University of Kentucky
Young Chun	American Institutes for Research	Paul Lavrakas	Nielsen Media Research
Mick Couper	University of Michigan	John Loft	RTI International
Michael Davern	University of Minnesota	Larry McGill	Princeton University
Theresa Demaio	U.S. Census Bureau	Daniel Merkle	ABC News
Dennis Dew	NORC	Joann Miller	University of Minnesota
Mark DiCamillo	Field Research Corp	Pama Mitchell	Mitchell Research Design
Karen Donelan	Partners Health Care	Warren Mitofsky	Mitofsky International
Kim Downing	University of Cincinnati	David Moore	Gallup
Katherine Lind Draughon	Draughon Research	Dawn V. Nelson	U.S. Census Bureau
Kathy Dykeman	Knowledge Networks	Diane O'Rourke	Univ. of Illinois at Chicago
Stephanie Eckman	NORC, Univ. of Chicago	Colm O'Muircheartaigh	NORC, Univ. of Chicago
Murray Edelman	CBS News	Vince Price	University of Pennsylvania
Robert Eisinger	Lewis & Clark College	Ken Rasinski	NORC, Univ. of Chicago
Carol-Ann Emmons	NORC	Dietram Scheufele	Univ. of Wisconsin, Madison
Steve Everett	Everett Group	Alisu Schoua-Glusberg	Research Support Services
David Fan	University of Minnesota	Andy Smith	University of New Hampshire
Barry Feinberg	Custom Research, Inc.	Tom W. Smith, Chair	NORC- Univ. of Chicago
Philip Gendall	Massey University	Fred Solop	Northern Arizona Univ.
Karen Goldenberg	U.S. Bureau of Labor Statistics	Kristin Stettler	U. S. Census Bureau
Anna Greenberg	Greenberg Quinlan Rosner Research	Kent Tedin	University of Houston
		Michael Traugott	University of Michigan
Robert Groves	University of Michigan	Norman Trussell	Nielsen Media
Tom Guterbock	University of Virginia	Clyde Tucker	Bureau of Labor Statistics
Patricia A. Gwartney	University of Oregon	Bob Wyatt	Middle Tennessee State Univ.
Janet Hoek	Massey University	Cliff Zukin	Rutgers University

Innovators Award Committee

Bob Groves, Chair	University of Michigan	Frank M. Newport	Gallup Organization
E. Deborah Jay	Field Research Corp.	Susan H. Pinkus	Los Angeles Times
James M. Lepkowski	University of Michigan	Robert Y. Shapiro	Columbia University
Peter V. Miller	Northwestern University	Cliff Zukin	Rutgers University

AAPOR Award Committee

Nancy Belden, Chair	Belden, Russonello & Stewart
Deborah Jay	Field Research Corporation
Vincent Price	University of Pennsylvania
Lydia Saad	Gallup Organization
Howard Schuman	University of Michigan

AAPOR Book Award Committee

Paul Biemer	RTI International
Lawrence Bobo	Stanford University
Henry Brady	University of California-Berkeley
Scott Keeter	Pew Research Center
Paul Lavrakas	Nielsen Media
Warren Mitofsky	Mitofsky International
Patricia Moy	University of Washington
Vincent Price	University of Pennsylvania
Nora Cate Schaeffer	University of Wisconsin
Howard Schuman	University of Michigan
Bob Shapiro	Columbia University
Tom W. Smith, Chair	NORC - Univ. of Chicago

AAPOR Policy Impact Committee

David Cantor	Westat
Diane Colasanto	Consultant
Elizabeth Martin, Chair	U.S. Census Bureau
Fritz Scheuren	NORC
Judy Tanur	SUNY Stony Brook

AAPOR Council Nominations Committee

Lawrence Bobo	Harvard University
Henry Brady	Univ. of California – Berkeley
Mollyann Brodie	The Henry J. Kesier Family Fond.
Elizabeth Martin, Chair	U. S. Census Bureau
Roger Tourangeau	University of Maryland
Evans Witt	Princeton Survey Research Assoc.

AAPOR Endowment Fund Committee

Sandra H. Berry	RAND Survey Research Group
Edward A. Blair	University of Houston
Mario Callegaro	University of Nebraska-Lincoln
Helen Crossley	Research Consultant
Brad Edwards, Chair	Westat
Jack Elinson	Mailman School of Public Health
Harry W. O'Neill	NOP World
Diane O'Rourke	University of Illinois
Jennifer M. Rothgeb	U. S. Census Bureau

Membership & Chapter Relations Committee

Janice Ballou	Mathematica Policy Research
Mollyann Brodie, Chair	The Henry J. Kaiser Family
Jonathan Cohen	ABC News Polling Unit
Brad Edwards, Assoc. Chair	Westat
Douglas Blanks Hindman	Washington State Univ.
Graham Hueber	Opinion Research Corp.
Randy Keesling	RTI International
Carl Ramirez	U. S. General Accounting Office
Doug Schwartz	Quinnipiac Univ. Policy Institute

Education Committee

Allan L. McCutcheon	University of Nebraska-Lincoln
Geraldine Mooney, Chair	Mathematica Policy Research
Lydia Saad	Gallup Organization
Darby Miller Steiger	Gallup Organization
Roger Tourangeau	University of Maryland

Official Photographer

Steve Everett	Everett Group

Heritage Interview Committee

Herb Abelson	Princeton, NJ
Nancy Belden,	Belden Russonello & Stewart
Barbara Bryant	University of Michigan
Peyton M. Craighill-Co-Chair	Rutgers University
Don Dillman	Washington State Univ.
Brad Edwards	Westat
Jack Elinson	Mailman School of Public Health
Kathleen A. Frankovic	CBS News
Dawn Nelson	U.S. Bureau of the Census
Harry O'Neill	RoperASW
Susan Pinkus	Los Angeles Times
Tom W. Smith	NORC
Janet L. Streicher-Co-Chair	Citibank
Mark Schulman	Schulman, Ronca & Bucuvalas
Barbara Foley Wilson	National Center for Health Stats.
Shapard Wolf	Arizona State University

AAPOR Seymour Sudman Student Paper Award Cmte.

MollyAnn Brodie	Henry J. Kaiser Family Fond.
Maria Krysan	Univ. of Illinois at Chicago
David Moore, Chair	Gallup Organization
Clyde Tucker	U.S. Bureau of Labor Statistics

Public Opinion Quarterly Advisory Committee

Bob Groves	University of Michigan
Richard Kulka	RTI International
Frank Newport	Gallup Organization
Susan Pinkus	Los Angeles Times
Vincent Price, Chair	University of Pennsylvania
Tom W. Smith	NORC

Public Opinion Quarterly Editor

Peter V. Miller	Northwestern University

AAPOR Standard Definitions Committee

Mick Couper	University of Michigan
Robert P. Daves	Minneapolis Star Tribune
Paul J. Lavrakas	Nielsen Media Research
Tom W. Smith, Chair	NORC, Univ. of Chicago
Nancy Mathiowetz	Univ. of Wisconsin-Milwaukee
Shapard Wolf	Arizona State University

AAPOR Standards Committee

Nancy Bates	U.S. Census Bureau
Mick Couper, Chair	University of Michigan
Thomas Guterbock	University of Virginia
Paul J. Lavrakas	Nielsen Media Research
Mary Losch	University of Northern Iowa
Trevor Tompson	Associated Press

Newsletter Editor

Daniel M. Merkle	ABC News

Long Range Planning Committee

Nancy Belden	Belden Russonello & Stewart
Diane Colasanto	Consultant
Katherine Lind Draughon	Draughon Research
Brad Edwards	Westat
Scott Keeter	Pew Research Center
Nancy Mathiowetz	Univ. of Wisconsin-Milwaukee

Finance Committee

Sandra H. Berry, Chair	RAND Survey Research Center
Michael Flanagan	AMP
Daniel Merkle	ABC News
Jennifer Rothgeb	U. S. Census Bureau
Mark Schulman	Schulman, Ronca & Bucuvalas

Award Winners

Innovators Award Winners

2000	•Andrew Kohut
	•Robert M. Groves
	•Thomas Piazza
	•J. Merrill Shanks
	•Charlie Thomas
	•Richard Rockwell
	•William Lefes
	•Tom W. Smith
2001	•Norman Nie
	•Douglas Rivers
2002	•David Celentano
	•Jerome Johnston
	•Judy Lessler
	•James O'Reilley
	•Charles Turner
2003	•Robert P. Daves
	•Paul Lavrakas
	•Tom W. Smith
2004	•Robert Casady
	•Robert Lepkowski
	•Clyde Tucker
2005	•Thomas B. Jabime
	•Miron L. Straf
	•Judith M. Tanur
	•Roger Tourangeau

Julian Woodward Award Recipients

1955	Paul F. Lazarsfeld
1956	Herbert H. Hyman
1957	Public Opinion Quarterly
1958	Samuel A. Stouffer
1959	Elmo Roper
1960	Clyde W. Hart
1961	Fund for the Republic

AAPOR Award Recipients

1962	Angus Campbell
1963	George H. Gallup
1964	Harold D. Lasswell
1965	Harry H. Field
1966	Hadley Cantril
1967	Hans Zeisel
1968	Elmo C. Wilson
1969	Roper Public Opinion Research Center
1970	Archibald M. Crossley
1971	Walter Lippmann
1972	Jean Stoetzel
1973	Rensis Likert
1974	Bernard Berelson
1975	Raymond A. Bauer
1976	Joseph T. Klapper
1977	Leo Bogart
1978	W. Phillips Davison
1979	Mervin D. Field
1980	Shirley A. Starr
1981	Lester R. Frankel
1982	Paul B. Sheatsley
1983	Paul K. Perry
	Matilda White Riley
	John R. Riley, Jr.
	Wilbur Schramm
1984	Ithiel de Sola Pool
1985	Daniel Katz
1986	Philip Converse
1987	Norman Bradburn
	Seymour Sudman
1988	Burns W. Roper
1989	Gladys Engel Lang
	Kurt Lang
1990	Herbert E. Krugman
1991	Joe Belden
1992	James Davis
1993	Jack Elinson
1994	Howard Schuman
1995	Herbert I. Abelson
1996	Eleanor Singer
1997	Irving Crespi
1998	Albert E. Gollin
1999	Charles F. Cannell
	Warren J. Mitofsky
2000	Philip Meyer
2001	Robert M. Groves
2002	Tom W. Smith
2003	Don A. Dillman
	Frank Stanton
2004	Benjamin J. Page
2005	TBA at the Awards Banquet

AAPOR Book Award

2004
John R. Zaller, *The Nature and Origins of Mass Opinion* (Cambridge University Press, 1991)

2005
Howard Schuman, Charlotte Steeh, Lawrence Bobo, Maria Krysan, *Racial Attitudes in America: Trends and Interpretations* (Harvard University Press, 1997)

Policy Impact Award

2004
The Urban Institute
Assessing the New Federalism Project

2005
The HIV Cost and Services Utilization Study (HCSUS) Conducted by Rand Corporation

AAPOR T-shirt Slogan Contest Winner

Doug Lonnstrom, Siena College, ***"If you Don't Like the Estimate, Just Weight"***

AAPOR Student Paper Award

1967
First
Kay K. Deaux, University of Texas
Second
Thomas D. Cook, Stanford University
G. Ray Funkhouser, Stanford University
Paul Hirsch, University of Michigan
John A. Michael, Columbia University
Sanci Michael, Columbia University

1968
First
Lawrence W. Green, University of California, Berkeley
Second
John S. Reed, Jr., Columbia University
Third
Marshall Childs, Columbia University
Honorable Mentions
Jonathan Kelley, University of California, Berkeley
Richard Young, Stanford University

1969
First
David Knoke, University of Michigan
Second
Charles Atkin, University of Wisconsin
Drury Sherrod, Stanford University
Honorable Mentions
Kent Anderson, Stanford University
Douglas Hall, Stanford University
Jonathan Kelley, Univ. of CA, Berkeley

1970
First
Charles K. Atkin, Univ. of Wisconsin
Second
Gary A. Mauser, Univ. of California, Irvine
Third
Philip Palmgreen, Univ. of Kentucky

1971
Marcus Felson, University of Michigan

1972
Gwen Bellisfield, New York University

1973
Paul J. Placek, Vanderbilt

1974
D. Garth Taylor, University of Chicago

1975
First
James R. Beniger, Univ. of CA, Berkeley
Second
Martin I. Horn, University of Connecticut
Third
Victoria L. Swigert, SUNY-Albany
Honorable Mention
Bonnie J. Kay, Northwestern University

1976
Winner
Robert Navazio, University of North Carolina
Honorable Mentions
Stanley Presser, University of Michigan
Kevin Lang, Oxford University
Alicia J. Welch, University of Massachusetts

1977
Winners
Twila Foster, Univ. of California, Berkeley
(Helen S. Dinnerman Prize)
Michael Goldstein, Univ. of CA, Berkeley
Trudy Martin, Univ. of CA, Berkeley
Mark J. Rogers, Univ. of CA, Berkeley
Honorable Mentions
Josephine Holz, Philadelphia, Pennsylvania
Claire B. McCullough, Univ. of Maryland
Michael J. O'Neil, Northwestern University

1978
Winner
Marie Crane, University of Michigan
Second
Michael Carozzo, University of Kentucky
Honorable Mention
Jeff Sobal, University of Pennsylvania

1979
Eric R.A.N. Smith, Univ. of CA, Berkeley

1980
Honorable Mention
Marianne Berry, University of Michigan

1981
Lawrence D. Bobo, University of Michigan

1982
Richard Bagger, Princeton University
Lynda Clarizio, Princeton University
Earl Cook, Princeton University
Linda Curtis, Princeton University
Robert Giuffra, Princeton University
Keating Holland, Princeton University
David Heubner, Princeton University
Claire Laporte, Princeton University
Warren Lazarow, Princeton University
Demetria Martinez, Princeton Univ.
Burns Stanfield, Princeton University
Terri Steinhaus, Princeton University
Michele Warman, Princeton University

1983
Winner
John Zaller, Univ. of CA, Berkeley
Honorable Mentions
Cynthia Fletcher, Iowa State University
John G. Geer, Princeton University
Jon A. Krosnick, University of Michigan
Bruce Peterson, University of Chicago

1984
Winners
Jon A. Krosnick, University of Michigan
Robert W. Kubey, Univ. of Chicago
Honorable Mentions
Phyllis M. Endreny, Columbia Univ.
Carrol J. Glynn, University of Georgia
John Zeglarski, Rutgers University

1985
Winner
Donald P. Green, Univ. of CA, Berkeley
Honorable Mentions
Gerald M. Kosicki, Univ. of Wisconsin
Gary R. Pettey, University of Wisconsin
Martin I. Gilen, Univ. of CA, Berkeley

1986
Winner
Dorothy Watson, Univ. of Wisconsin
Honorable Mentions
Jacqueline Scott, Univ. of Michigan
Ken Dautrich, Eagleton Institute

1987
Winner
Jacqueline Scott, Univ. of Michigan
Honorable Mentions
Stephen Ayidiya, University of Akron
Pamela Campanelli, Univ. of Michigan

1988
Winner
James Dearling, Univ. of Southern CA
Honorable Mentions
William Axinn, University of Michigan
Glenn Dempsey, University of Chicago
Donna Wasserman, Univ. of Michigan

1989
Winner
Diana Mutz, University of Wisconsin

1990
Winners
Jonathan Cowden, Yale University
Shoon Murray, Yale University
Honorable Mention
Barbara Bickart, University of Florida

1991
Martin I. Gilen, Univ. of CA, Berkeley

1992
Winner
Jeffery Mondak, Univ. of Pittsburgh
Honorable Mentions
Jian-Hua Zhu, Univ. of Connecticut
Anne S. Welch, Northwestern Univ.

AAPOR Student Paper Award Continued...

1993
Winner
Michael Wänke, Univ. Mannheim
Honorable Mentions
E. Clement Brooks, Univ. of CA, Berkeley
E. Marla Fletcher, Northwestern Univ.
J. Frederick-Collins, Univ. of North Carolina

1994
Winner
Scott L. Althaus, Northwestern Univ.
Honorable Mentions
Robert M. Eisinger, Univ. of Chicago
Michael W. Link, Univ. of South Carolina
Richard J. Timpone, SUNY at Stony Brook

1995
Mollyann Brodie, Harvard University
Damarys Canache, Univ. of Pittsburgh

1996
Winner
Maria Krysan, Pennsylvania State Univ.
Honorable Mentions
Daniel Dowd, Yale University
Julie Press, Univ. of California, LA
Eleanor Townsley, Univ. of CA, LA

1997
Winner
Paul Goren, University of Pittsburgh
Honorable Mention
Robert D. Woodberry, Univ. of North Carolina

1998
Winner
Cindy T. Christen, Univ. of Wisconsin, Madison
Prathana Kannaovakun, Univ. of Wisconsin, Madison
Honorable Mentions
J. Tobin Grant, Ohio State University
Stephen T. Mockabee, Ohio State Univ.
Devon Johnson, Harvard University

1999
Bo Zhou, University of Akron

2000
Alexander Todorov, New York University

AAPOR Seymour Sudman Student Paper Award

2001
Joshua D. Clinton, Stanford University

2002
David Dutwin, University of Pennsylvania

2003
Winners
Devon Johnson, Harvard University
Markus Prior, Princeton University
Honorable Mention
Clarissa Davis, Univ. of Pennsylvania

2004
Kate Kenski, University of Pennsylvania
Dannagal Goldthwaite Young, Univ. of Pennsylvania

2005
Winners
Amy R. Gershkoff, Princeton University
Jennifer Dykema, University of Wisconsin - Madison
Honorable Mention
Adria Lawrence, University of Chicago
Bethany Albertson, University of Chicago

Burns "Bud" Roper Fellows

2005 Steven Barkan, University of Maine
Kelly N. Foster, University of Georgia
HeyKyung Koh, Institute of International Education
Neil Malhotra, Stanford University
David Wilson, Gallup Organization

AAPOR Regional Chapter Presidents 2004-2005

Washington, DC/Baltimore - DCAAPOR
Carl Ramirez – U. S. General Accounting Office

Midwest – MAPOR
Douglas Blanks Hindman – Washington State University

New England – NEAAPOR
Doug Schwartz – Quinnipiac Univ. Polling Institute

New York – NYAAPOR
Graham Hueber – Opinion Research Corp.

New Jersey – NJAAPOR
Janice M. Ballou - Mathematica Policy Research

Pacific – PAPOR
Jonathan Cohen – ABC News Polling Unit

Southern – SAPOR
Randy Keesling – RTI International

Very Special Thanks from the Conference Chair

Nancy Belden	Belden Russonello & Stewart
Katherine "Kat" Lind Draughon	Draughon Research
Ned English	NORC
Steve Everett	Everett Group
Michael P. Flanagan, CAE	AAPOR Executive Coordinator
Monica A. Frihart	AAPOR Association Assistant
Missy M. Johnson, CMP	AAPOR Meeting Planner
Jibum Kim	NORC
Seokho Kim	NORC
David Moore	Gallup
Tara Nelson	AAPOR Data Administrator
Nancy Whelchel	N.C. State University
Cliff Zukin	Rutgers University

Conference Program Design, Editing, Mailings, Advertising

Michael P. Flanagan, CAE	AAPOR Executive Coordinator
Monica A. Frihart	AAPOR Association Assistant
Missy M. Johnson, CMP	AAPOR Meeting Manager
Tara Houston-Sawyers	AMP Graphics Artist

Contributions from April 2004 - February 2005

General Endowment

Janice M. Ballou
Nancy Belden
Sandra H. Berry
Jacques B. Billiet
Robert J. Blendon
Michael Butterworth
Floyd Ciruli
Mick P. Couper
Joye Dillman
Brad Edwards
Thomas Guterbock
Ronald Hinckley
Syed Arabi Idid
David A. Jodice
David Krane
Richard A. Kulka
Paul J. Lavrakas
Allan L. McCutcheon
Judith Mopsik
Diane O'Rourke
Orjan O.V. Olsen
Alice Robbin
Jennifer M. Rothgeb
Mark A. Schulman
J. Merrill Shanks
Barbara Simon
Paul A. Talmey
John Tarnai
Andrew Thibault
Nicholas J. Tortorello
Clyde Tucker
Cliff Zukin

Seymour Sudman Fund

Joan S Black
Edward A. Blair
Jacqueline Bruskin
Michael Butterworth
Skip Camp
Karen Donelan,
James L. Esposito
J. E. Godard
Miguel Gomez-Barrantes
Robert M. Groves
Timothy Johnson
Ariela Keysar
Paul J. Lavrakas
Janice J. Leonard
John Loft
Elizabeth Nichols
Diane O'Rourke
Linda K. Owens
Alvin Richman
Jennifer M. Rothgeb
John Tarnai
Clyde Tucker
Richard B. Warnecke
Diane K. Willimack
Cliff Zukin

Heritage Interview Series Fund

Nancy Belden
Barbara Everitt Bryant
Don A. Dillman
Brad Edwards
James M. Ellis
Eugene P. Ericksen
David Krane
Janice J. Leonard
Warren J. Mitofsky
Geraldine Mooney
John E. Nienstedt
Alice Robbin
Robert O. Simmons, Jr.
Janet L. Streicher
John Tarnai

CONFERENCE SPONSORS

AAPOR acknowledges and gratefully expresses our appreciation to the following organizations. The support provided by these corporate colleagues enables AAPOR to continue to hold quality conferences.

PLATINUM-LEVEL SPONSORS

NORC
AT THE UNIVERSITY OF CHICAGO
A national organization for research and computing

NORC - *For sponsoring Thursday evening's dessert reception and Friday morning's break*

PULSE TRAIN
the future for survey software

Pulse Train - *For sponsoring Saturday evening's wine*

GOLD-LEVEL SPONSORS

Knowledge NETWORKS

Knowledge Networks - *For sponsoring Thursday evening's wine*

Abt
Abt Associates Inc.

Abt Associates - *For sponsoring Friday afternoon's break*

FOR PUBLIC OPINION RESEARCH
Roper
CENTER

Roper Center - *For sponsoring the Totebags*

Marketing Systems Group

Marketing Systems Group- *For sponsoring the hors d'oeuvres at Saturday evening's President's Reception*

SRBI
Schulman, Ronca & Bucuvalas, Inc.

Schulman, Ronca, & Bucuvalas, Inc. - *For sponsoring Saturday morning's break*

SPSS

SPSS - *For sponsoring the hors d'oeuvres at Thursday evening's reception and jointly sponsoring the Saturday evening banquet's PowerPoint presentation*

ORC MACRO[SM]

ORC Macro - *For sponsoring the Saturday Evening Banquet Program Book*

WESTAT
AN EMPLOYEE-OWNED RESEARCH CORPORATION

Westat - *For sponsoring the Building a Career in Survey Research Session and a Commemorative Item – carabiner with compass*

SILVER-LEVEL SPONSORS

CfMC - *For sponsoring the Saturday Evening Banquet Centerpieces*

Greenfield Online – *For sponsoring a drink ticket for all new members at the Thursday Evening New Member/ All Chapter Reception*

Rand Survey Research Group - *For sponsoring the A/V for three breakout rooms*

RTI – *For jointly sponsoring the Saturday Evening Banquet PowerPoint Presentation*

Western Wats - *For sponsoring a Commemorative Item - Paper/Book Clips*

www.e-tabs.com
E-Tabs - *For sponsoring a Commemorative Item –CD Holder*

Oxford University Press - *For sponsoring the Saturday Evening Banquet Head Table Centerpieces*

The Gallup Organization - *For sponsoring the Fun Run/Walk and the Golf Outing*

Survey Sampling International - *For sponsoring the break for the Survey Translation: Issues and Procedures Short Course*

AAPOR Technology, Software and Book Exhibits

Technology & Software Exhibitors

AcWire Technologies – Wayne, NJ
ASDE Survey Sampler – Hull, Quebec
ATP Canada – Richmond Hill, Ontario
CfMC – San Francisco, CA
E-Tabs – Sandwich, IL
Gazelle Global Research Services, New York, NY
GMI Inc. – Mercer Island, WA
Greenfield Online – Wilton, CT
iModerate – Denver, CO
Marketing Systems Group – Fort Washington, PA
NORC – Chicago, IL
Opinion Access Corp. – Long Island City, NY
ORC Macro – New York, NY
Oxford University Press – Cary, NC
Pulse Train Limited – Guildford, Surrey UK
Roper Center – Storrs, CT
RTI – Research Triangle Park, NC
Sawtooth Technologies, Inc. – Northbrook, IL
Scientific Telephone Samples – Foothill Ranch, CA
Snap Survey Software – Portsmouth, NH
SPSS – New York, NY
SRBI – New York, NY
Survey Sampling International – Fairfield, CT
Westat – Rockville, MD
Western Wats – Orem, UT

Publishers/Book Exhibitors

Cambridge University Press – New York, NY
University of Chicago Press – Chicago, IL
Columbia University Press – New York, NY
Greenwood – Westport, CT
Guilford Press – New York, NY
Harvard University Press – Cambridge, MA
Lawrence Erlbaum Associates, Inc. – Mahwah, NJ
Oxford University Press – New York, NY
Princeton University Press – Princeton, NJ
Rowman & Littlefield – Lanham, MD
SAGE Publications – Thousand Oaks, CA
John Wiley & Sons – Somerset, NJ

Agencies That Have Helped to Ensure AAPOR's Financial Health Through Contributions During the Past Year

A

A.S.E.P., INC.	San Juan, Puerto Rico
Affordable Samples, Inc.	Old Greenwich, CT
Aguirre International	Burlingame, CA
American Institutes For Research	Washington, DC
Anderson, Niebuhr & Associates, Inc.	Arden Hills, MN
ASDE Survey Sampler, Inc.	Gatineau (Hull), Quebec, Canada
Aspen Systems Corporation	Rockville, MD

B

Baseline Market Research, Ltd.	Fredericton, New Brunswick, Canada
Bauman Research & Consulting, LLC	Glen Rock, NJ
Belden Russonello & Stewart	Washington, DC
Bisconti Research, Inc.	Washington, DC
Bureau for Social Research	Stillwater, OK
Buro De Investigacion De Mercados, S.A. De C.V.	Mexico, DF CP Mexico

C

California Survey Research Services, Inc.	Van Nuys, CA
CALLC	Bethesda, MD
Center For Governmental Services	Auburn University, AL
Center For Social And Behavioral Research	Cedar Falls, IA
Center For Survey Research	Charlottesville, VA
Center For Survey Research	Boston, MA
Center For Survey Research	Bloomington, IN
Center For Survey Statistics & Methodology	Ames, IA
Charlton Research Company	Pleasanton, CA
Clearwater Research, Inc.	Boise, ID
Critical Insights, Inc.	Portland, ME

D

D3 Systems, Inc.	Vienna, VA
Data Recognition Corporation	Maple Grove, MN
Dittman Research Corporation of Alaska	Anchorage, AK

E

Eastern Research Services	Springfield, PA
Elway Research, Inc.	Seattle, WA
Ernst & Young Quantitative Economics & Statistics Group	Washington, DC

F

Facts Worldwide India Private Limited	Prabhadevi, Mumbai India
Fairbank, Maslin, Maullin & Associates	Santa Monica, CA
Field Research Corporation	San Francisco, CA
Freeman, Sullivan & Co./Population Research Systems, LLC	San Francisco, CA

G

Gilmore Research Group	Seattle, WA
Global Strategy Group	Chevy Chase, MD

H

Health Survey Research Unit	San Francisco, CA
Hollander, Cohen & Mcbride, Inc.	Baltimore, MD

I

ICR International Communications Research	Media, PA
Ilres S.A.	Luxembourg
Indiana Univ. Public Opinion Laboratory	Indianapolis, IN
Institut Fur Demoskopie Allensbach	Allensbach, Germany
Institute for Policy Research	Cincinnati, OH
Intermedia Survey Institute	Washington, DC
Investigaciones Sociales Aplicadas, S.C.	Mexico, DF Mexico

J

J.D. Franz Research, Inc.	Sacramento, CA
Juarez & Associates, Inc.	Los Angeles, CA

K

Keleman Associates, Inc.	New York, NY
Knowledge Networks, Inc.	Menlo Park, CA
Kochevar Research Associates	Charlestown, MA

L

Leger Marketing	Montreal, Quebec, Canada
LHK Partners Inc.	Newtown Square, PA

M

Macro International Inc. (ORC Macro)	Beltsville, MD
Marketing Systems Group	Fort Washington, PA
Mathematica Policy Research, Inc.	Princeton, NJ
Mitofsky International	New York, NY
Monroe Mendelsohn Research, Inc.	New York, NY
MORI UK	London, United Kingdom
Mund Americas	Mexico City, DF Mexico

N

National Opinion Research Center (NORC)	Chicago, IL
National Opinion Research Svs. (NORS)	Miami, FL
NHK Public Opinion Research Division	Tokyo, Japan

O

O'Neil Associates, Inc.	Tempe, AZ
Olivares Plata Consultores, S.A. DE C.V.	Mexico, DF, Mexico
Opinion Access Corp.	Long Island City, NY
Oregon Survey Research Laboratory	Eugene, OR
Orima Research	Campbell Act, Australia

P

Pearson S.A. DE C.V.	Mexico, DF Mexico
Public Opinion Laboratory	Dekalb, IL

Q

QSA Research And Strategy	Alexandria, VA
Questar	Eagan, MN

R

Rand Survey Research Group	Santa Monica, CA
Reda International, Inc.	Wheaton, MD
Renaissance Research & Consulting	New York, NY
Research Support Services	Evanston, IL
RLS Associates	Ann Arbor, MI
Roper Center	Storrs, CT
RTI International	Research Triangle Park, NC
Ruth Diamond Market Research Services, Inc.	Buffalo, NY

S

Schulman, Ronca & Bucuvalas, Inc.	New York, NY
Social & Economic Sciences Research Center	Pullman, WA
Social Science Research Laboratory	San Diego, CA
Social Science Survey Center	Santa Barbara, CA
Social Weather Stations	Quezon City 1101, Philippines
SPSS Inc.	Chicago, IL
Stennis Institute of Government	Mississippi State, MS
Survey Research Center	Fayetteville, AR
Survey Research Center	Ann Arbor, MI
Survey Research Center	University Park, PA
Survey Research Center	Berkeley, CA
Survey Research Institute	Ithaca, NY
Survey Research Laboratory	Chicago, IL
Survey Research Laboratory	Columbia, SC
Survey Research Unit	Chapel Hill, NC
Survey Sampling International	Fairfield, CT
SurveyUSA	Verona, NJ

T

The Gary Siegel Organization, Inc.	Chicago, IL
The Michael Cohen Group LLC	New York, NY
The Polling Company, Inc.	Washington, DC
TNS	Horsham, PA

U

U.S. Census Bureau	Washington, DC
University Center For Social & Urban Research	Pittsburgh, PA
University of Wisconsin Survey Center	Madison, WI
UNL Gallup Research Center	Lincoln, NE

V

Virginia Tech Center for Survey Research	Blacksburg, VA

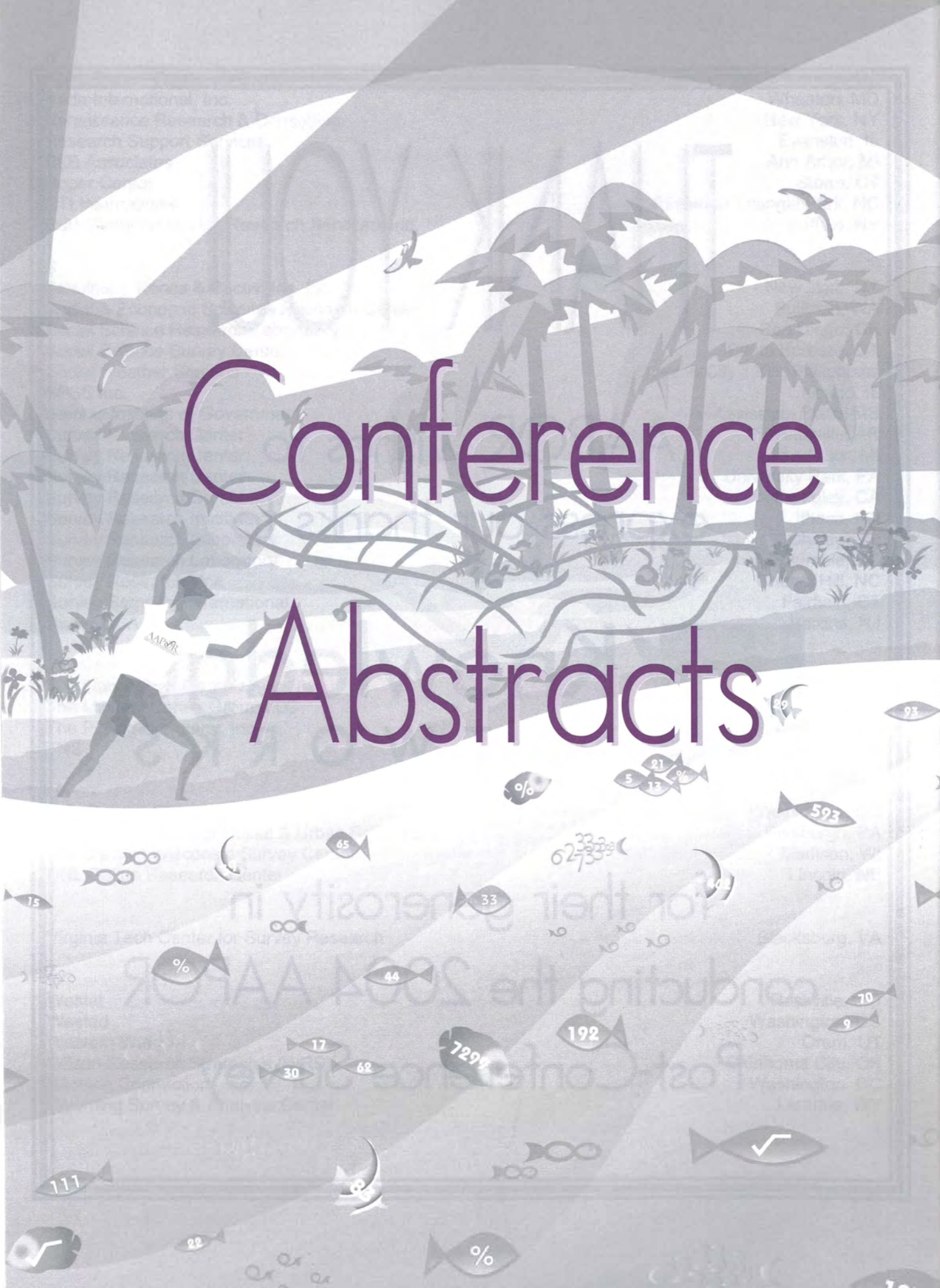
W

Westat	Rockville, MD
Wested	Washington, DC
Western Wats	Orem, UT
Wilson Research Strategies (WRS)	Oklahoma City, OK
Winning Connections Inc.	Washington, DC
Wyoming Survey & Analysis Center	Laramie, WY

THANK YOU!

AAPOR wishes to
express our thanks to

for their generosity in
conducting the 2004 AAPOR
Post-Conference Survey



Conference Abstracts

WEIGHTING

***Propensity Models Versus Weighting Cell Approaches to Nonresponse Adjustment: A Methodological Comparison*, Peter Siegel, RTI International; siegel@rti.org, James Chromy, RTI International; jrc@rti.org, and Elizabeth Copello, RTI International; copello@rti.org**

Statistical adjustment of nonresponse is a deep and pervasive issue for sample surveys. Contemporary statistical methods offer two broad classes of approach to nonresponse adjustment. One is the use of a traditional weighting cell approach. More recently, response propensity modeling, using, typically, logistic regression, has been developed as a further approach to nonresponse adjustment. Additionally, RTI's General Exponential Model (GEM) is a generalization of raking type weight adjustments, and in addition to nonresponse adjustment it can optionally include features such as poststratification and weight trimming. We use the data from the Education Longitudinal Study of 2002 (ELS:2002) to compare the results of the weighting class method, raking, a logistic regression propensity model, and GEM. We focus on nonresponse adjustment but also look at extreme weight adjustment and poststratification. For the one-dimensional case where each unit is in one unique cell, it can be shown that the four methods will produce the same results. Expanding to many variables and multiple dimensions, marginal totals, variances, and weight distributions can be compared. We also conduct a limited nonresponse bias analysis and examine the effects each method has on reducing nonresponse bias.

***Dealing with wide Weight Variation in Polls*, Richard Griffin, U.S. Census Bureau; richard.a.griffin@census.gov**

It is well known by survey practitioners that wide weight variation is not good. If a respondent with a relatively large weight is an outlier the resulting estimate may be inappropriately skewed. However, weight variation is not always bad. Unequal selection probabilities result in unbiased estimates that have weight variation. Consider optimal allocation in stratified sampling when the variance of the characteristic of interest is different from one strata to the next. The optimal allocation results in an unbiased minimum variance estimate that has weight variation. Why do we consider weight variation bad? The problem is that most surveys have multiple purposes and many estimates are computed other than those for which the sample design was targeted. Statistically the problem is to choose the estimation procedure that produces the lowest mean square error (the variance of an estimate plus the square of its bias). A biased estimate can have lower mean square error than an unbiased estimate if its variance is enough lower than the variance of the unbiased estimate to compensate for the bias. Polls often have equal probability samples with unequal balancing weights to compensate for differential nonresponse. This paper looks at the effect of unequal weights for equal probability polls on the mean square error of estimates. Balancing weights using known population proportions as well as shrinkage weights under a model are examined.

***Party ID Weighting - Experiments to Improve Survey Quality*, Olena Kaminska, MSR in Univ. of Connecticut,; olena.kaminska@uconn.edu, and Christopher Barnes, Univ. of Connecticut - CSRA; cebuconn@yahoo.com**

Weighting the pre-election polls on party ID has become one of the most discussed and controversial issues in survey research field. While most studies on the topic examine if party identification is stable and on this basis attempt to prove or disprove the validity of weighting on party ID, it may be reasonable to weight on PID even if the distribution of party ID is not stable. A portion of change in party ID may be an indicator of the change in response willingness rather than actual change in party ID. If party ID is correlated with response willingness, weighting can improve the response rate error. The paper contains three experiments regarding the possibility of improving survey quality by weighting on party ID. First, we examine if the prediction of the election results of 2000 improves after adding the weight on party ID to original weights. Second, assuming that weighted dataset on original weights is more accurate as compare to non-weighted, we examine if weighting only on party ID shifts the results in the direction to those weighted on initial weights. Third we compare the change in demographic questions after weighting on party ID with those taken from external datasets (such as Census and NES). The study is based on 2000 pre-election polls sponsored by next organizations: ABC News/Washington Post, Gallup/CNN/USA Today, NBC/WSJ, PSRA/Pew and Newsweek.

***Real-Time Propensity Models for Responsive Survey Design and Post-Survey Adjustment Through Propensity Models: Comparisons of Model Fit and Model Specification*, Robert M. Groves, University of Michigan; bgroves@isr.umich.edu, Jim Lepkowski, University of Michigan; jimlep@umich.edu, John VanHoewyk, U of Michigan; JVH@isr.umich.edu, and Paul Schulz, U of Michigan; pschulz@umich.edu**

The increasing uncertainty of the US public's reaction to survey requests has led to heightened awareness of the interplay of costs, nonresponse rates, and nonresponse errors in large scale household surveys. Groves and Heeringa (2004) outline an approach to survey design that directs orderly changes to key features of the recruitment protocol of a survey, based on real-time analysis of the incoming survey data. These so-called "responsive designs" identify a set of alternative key statistics, callback alternatives, and incentive options prior to the start of the data collection. Cost and error-related models are measured during the early phases of a survey, and then used to determine final design features, more nearly cost-optimal. In short, these designs adapt or respond to real-time information about the performance of the survey. A key tool in this responsive process is the use of propensity models on the sample case level, estimating the probability that an active case will be interviewed. Propensity models are also used after the survey data collection period has been completed. These propensity models are often used to form weighting class adjustments in an attempt to reduce unit nonresponse error (Little, 1982). These models identify groups that have higher or lower likelihood of being measured, based on all knowledge available at the end of the data collection. This paper addresses whether the predictors of propensity during data collection heavily overlap those available at the end of the data collection period. It relates this analysis to efforts during the data collection to attempt to achieve the most cost efficient acquisition of completed cases. It then studies how the estimated propensities of cases correlate with key statistics in the survey, among respondents. Conclusions are drawn about the relative utility of observational and process data predictors for response propensity in adjustment models.

SCALING AND RESPONSE SCALES

***Survey Question Length and the Reliability of Measurement*, Duane Alwin, Pennsylvania State University; dfa2@psu.edu, and Ryan McCammon, University of Michigan; mccammon@isr.umich.edu**

One element of survey question development on which there is a diversity of opinion is the issue of question length. Many subscribe to the view that questions should be as short as possible, although there is still some debate on this issue. In general, this topic is characterized by a variety perspectives and very little empirical evidence. Where there is evidence, the support seems to favor longer questions. However, on the basis of theories of the survey response process, we argue that the use of an excessive number of words may get in the way of the respondent's comprehension, and it is hypothesized that due to the cognitive burden of questions of greater length, there may be increased risk of measurement errors with longer questions. This paper addresses this problem, suggesting that one way to evaluate the issue of question length is in terms of its effects on measurement precision. We report the results of a recent study of the reliability of self-reports in face-to-face interviews using more than 330 questions spanning a range of content from 6 large-scale panel surveys conducted by the University of Michigan's Survey Research Center. The bulk of these results present a relatively convincing case for the conclusion that, at least with respect to some types of survey questions, there are declining levels of reliability associated with greater numbers of words in questions and further support for the typical advice given to survey researchers that their questions should be as short as possible. Findings reinforce the conclusions of some previous studies that verbiage in survey questions - either in the question text or in the introduction to the question - has negative consequences for the quality of measurement.

***Rating versus Comparative Trade-off Measures: Trending Changes in Political Issues across Time and Predictive Validity*, Randall Thomas, Harris Interactive; rthomas@harrisinteractive.com, Alyssa Johnson, Harris Interactive; amjohnson@harrisinteractive.com, Susan Behnke, Harris Interactive; sbehnke@harrisinteractive.com, and Mitchell Sanders, Harris Interactive; msanders@harrisinteractive.com**

We investigated the use of rating, ranking, constant sum, and paired comparison tasks for evaluating both the importance and support for a wide range of political issues. In 8 waves of fielding (running from October, 2003 to November, 2003), we had U.S. respondents participate in a web-based survey. For respondents of the 7 pre-election waves, we recontacted them following the November elections to complete another evaluative task. This allowed us to determine both the test-retest reliability and predictive validity of the initial evaluations. We replicated many of our earlier findings concerning the differences between rating and comparative trade-off tasks and further extended our understanding of the relative comparability and utility of the different measures across time.

***Measurement and Scaling Methods Independent of Response Style*, Jon Pinnell, MarketVision Research, Inc.; jpinnell@mv-research.com, and Lisa Fridley, MarketVision Research, Inc.; lfridley@mv-research.com**

Researchers are often forced to deal with differences between respondents' response style in analyzing survey data. Response style is often thought to include the subset of response alternatives (scale points) that a particular respondent is more likely to use. However, the researcher must also deal with differences in respondents' variation of responses. That is, research style reflects both a location effect and a dispersion effect. Both of which can cause substantial difficulty in analysis and interpretation. In practice, researchers commonly use ratings to measure attribute importance, preference, perception, satisfaction, and attitudes. Attempted solutions to the response style problem have included standardizing (or maybe just centering) data, using ranking rather than rating methods, and constant sum allocation. For the practicing researcher, each can have a detrimental effect, either on the audience of the research or on the respondent. This paper will explore two techniques - one old and one new - that we have found to better scale respondents and eliminate the impact of response style. The method of paired comparisons seeks only ordinal responses from respondents, but produces metric output. Historically, this metric output was available only for a population as a whole, rather than for each individual. However, by introducing Bayesian methods we can provide metric scaling separately for each individual. The second approach, which is growing in popularity, is Maximum Difference Scaling, or MaxDiff. MaxDiff can also rely on Bayesian methods to provide individual level estimation. This paper will provide background on paired comparison and MaxDiff, as well as the Bayesian estimation. Empirical results from commercial studies comparing these two methods to traditional rating scales will be presented.

INTERVIEWER RACE AND TRAINING

***Race-Matching: Interviewers' reactions to the race-matching process*, Cindy Boland-Perez, UALR- IOG Survey Research Center; cbboland@ualr.edu, P. Denise Cobb, SIUE; pcobb@siue.edu, and Patricia Lebaron, UALR-IOG Survey Research Center; palebaron@ualr.edu**

Research shows that responses to survey questions can be influenced by whether respondents perceive themselves and their interviewers to be of the same race. Several studies have shown that African-American respondents who perceive they are being interviewed by white interviewers are likely to give substantively different responses from those of his/her counterparts who perceive the interviewer to be African-American. These findings are especially relevant when the content of questions directly relates to racial attitudes, perceptions of prejudice, and experiences related to race and ethnicity. With awareness of this potential interviewer effect on measurement for surveys of this kind, a university-based survey center developed an interviewer/respondent race-matching methodology for an annual survey of racial attitudes that began in 2003. While there are clear benefits to using this matching protocol, such as encouraging respondent honesty and frankness, there may be latent pitfalls for such designs. A survey design that incorporates interviewer/respondent race-matching protocol may interject unique problems into the processes of a call center. Thus to explore this issue further, we developed an exploratory qualitative inquiry to focus on over thirty interviewers' reactions to collecting data on race-related questions and especially their interpretations and sensitivity to the race-matching process. We were particularly interested in instances in which the race-matching design breaks down and respondents perceive interviewers to be of a different race. We argue that while survey researchers have made clear strides in understanding the effects of race-matching on measurement, there is much to learn about the potential for marginalization of interviewers. Moreover, given that effective interviewing is one key to achieving quality data in telephone survey operations, it is critical to understand the potential positive and deleterious outcomes that such survey procedures may have on the interviewing staff.

***Respondent Views on Perceived Racial Context and Race of Interviewer Effects*, David Wilson, Gallup/Michigan State University; david_wilson@gallup.com, and Jack Ludwig, Gallup; jack_ludwig@gallup.com**

At their most basic level, race of interviewer (ROI) effects consist of two parts: a question, and a response. The question is said to elicit a need to consider the context of the survey interview, and the response is the result of the consideration. The explanation for the altering of responses based on the interaction between the question and context, has been social desirability. However, the social desirability model also assumes a level of "social conformity pressure" on the part of the respondent. We argue that in order for respondents to rationally consider the different racial expectations of the interviewer, there should be some perceived understanding of the history, nature, or extent to which differences in race matter. Put simply, the more respondents perceive racial differences, the more respondents should consider race in their responses to items. We consider a negative racial context to be one such measure of racial awareness. A negative racial context is an environment characterized by perceived unfair treatment based on race. In the analysis of race relations, perceived context provides an indication of the extent to which racial groups are aware of either real or perceived racial differences in society. We expect that both Blacks and Whites, who perceive a more negative racial environment, should also be more likely to be sensitive to the racial context of certain survey items. The underlying assumption is that respondents actually consider these perceptions of the context to be relevant to any survey topic that shows ROI differences. Using three years of data collected by the Gallup Organization, we observe the effects of racial context and race of interviewer on perceptions race relations in America.

***Out of the lab and into the field: Developing a protocol for responding to diversified recruitment needs*, Karen Whitaker, National Center for Health Statistics; KWhitaker@cdc.gov, Miller Kristen, NCHS; KMiller@cdc.gov, and Willson Stephanie, NCHS; SWillson@cdc.gov**

Poster Over the past decade, cognitive testing has become a standard pre-testing mechanism for surveys conducted by federal agencies. Indeed, there is a growing trend among federal agencies to require cognitive evaluation of questions prior to their inclusion on federal surveys. However, depending on the subject matter of the questions, adequate pre-testing may preclude cognitive interviews from being conducted in typical fashion-that is, from within the confines of the question-design laboratories of the federal agencies with respondents local to the DC metropolitan area. For projects that necessitate specific target groups which either cannot be brought to the lab or accessed in the DC metropolitan area, special considerations must be made. This poster will describe a project conducted by the Questionnaire Design Research Laboratory staff of the National Center for Health Statistics which cognitively evaluated a disaster needs assessment questionnaire for the National Institute of Mental Health. To successfully evaluate the questions, interviews were conducted in Orlando, Florida with citizens who were impacted by Hurricanes Charlie, Francis, and Jeanne. In describing this project, the poster will present the advantages for off-site cognitive evaluation, and characterize dimensions of projects that necessitate field interviews. Additionally, it will articulate the special considerations necessary for conducting off-site cognitive interviews such as insider/grass roots contacts, recruiting strategies, remuneration, interviewing location(s) i.e., in-home vs. private room in a public facility, recording equipment, computer equipment, printing and copying needs, telephone service for making calls to potential respondents and communicating with sponsor, and the safekeeping and transport of identifiable data, i.e., audio and video tapes and consent forms. Finally, the poster will summarize the primary findings that were discovered through the advantaged position of an off-site location.

***An inquiry into the efficacy of a short training for interviewers*, Claire Durand, Universite de Montreal; Claire.Durand@umontreal.ca, Marie-Eve Gagnon, Universite de Montreal; Marie-Eve.Gagnon@umontreal.ca, Christine Doucet, Universite de Montreal; Christine.Doucet@umontreal.ca, and Eric Lacourse, Universite de Montreal; Eric.Lacourse@umontreal.ca**

Interviewers are at the heart of survey quality and therefore we should be paying considerable attention to their work. In devising the training presented in this paper, we postulated that a better knowledge of the survey process would help empower interviewers and allow them to find, by themselves, better ways in which to perform their job. This paper will present a training experiment conducted at a private firm located in Montreal, Canada. This firm was awarded the contract to conduct the 2004 Canadian Addiction Survey of 14,000 persons throughout Canada. The survey lasted four months and required that the firm hire many new interviewers; it was therefore an ideal setting to experiment with complementary training of interviewers. The process consisted of a) an observation of the standard training sessions provided by the firm; b) a pre-training questionnaire to all interviewers after about 20 hours of work on the project; c) the monitoring of the evolution of the daily performance of the interviewers using an index of the net contribution to performance (NCI); d) one-hour training sessions to three groups of interviewers (n=18) identified as either low-performers or newly hired interviewers; e) a post training questionnaire two weeks after the training; and f) an assessment of the impact of the training in three areas : knowledge acquisition and attitudes, retention, and performance. Previous analysis (Durand et al., 2004) showed that the training had a positive impact in these three areas. This paper will further examine the impact of training on performance, focusing on the link between pre training attitudes, performance, and impact of training. The question asked is: who profited from training? More specifically, the paper will compare the attitudes of low versus high performers and the relationship between these attitudes and receptivity to training.

MORALITY, RELIGION AND POLITICS

***The basis of voting preferences among evangelical Christians in the 2004 election*, Kenneth Blake, Middle Tennesee State University; kblake@mtsu.edu, Robert Wyatt, Middle Tennessee State University; rwyatt@mtsu.edu, and Holly Warf, Middle Tennessee State University; hkw2b@mtsu.edu**

It is by now a widely circulated thesis that evangelical Christian voters helped propel Republican George W. Bush's win over Democratic challenger John Kerry in the 2004 presidential race. The aid Bush received from evangelical "'values' voters" may be the latest example of the increasingly close ties between committed, white evangelical Protestants and the Republican Party. This paper explored whether evangelicals' voting preferences appear based on the sort of ideologically constrained opinions that Converse (1964) famously found so rare in the electorate, or, instead, on what Popkin (1991) termed "'gut' reasoning." Drawing upon CHAID analyses of data from two statewide RDD telephone polls conducted in America's evangelical heartland during fall 2004 and spring 2005, the study finds little evidence to exempt evangelicals from Converse's (1964) critique of voters as generally lacking a constrained political belief system. Asked about five domestic issues on which Bush and Kerry had articulated divergent views, evangelicals exhibited significantly less ideological constraint and issue knowledge compared to non-evangelicals, although neither group's average scores turned up all that far from the zero mark. And while evangelicals did, indeed, differ from non-evangelicals on the issues of restricting abortion and allowing gay marriages or gay civil

unions, indicators of the sort of "gut" rationality described by Popkin dominated as predictors of voting preference among white evangelicals. Only among white non-evangelicals did ideological constraint play something of a supporting role in voting preference. Finally, differences were noted between the heuristics likely employed by evangelicals and non-evangelicals. Specifically, while perceived candidate intelligence was an important consideration for many white evangelicals, most characterized the two candidates as equally intelligent and relied more on perceived candidate morality as a guide to voting choices. Relatively speaking, perceived candidate intelligence was more important to non-evangelicals, and perceived candidate morality simply was not a distinguishing factor.

***Understanding the Catholic Vote in 2004*, Kate Stewart, Belden Russonello & Stewart; katestewart@brspoll.com, John Russonello, Belden Russonello & Stewart; Johnrussonello@brspoll.com, and Rachel Sternfeld, Belden Russonello & Stewart; rachelsternfeld@brspoll.com**

Catholic voters comprise 27% of the electorate. They have been a key voting group since the early 1970's. In every presidential election since then, the candidate that wins the Catholic vote has won the popular vote. The 2004 election is of particular interest to look at the Catholic vote because the Catholic clergy became much more involved in politics than ever before to energize social conservatives. While Kerry lost the Catholic vote to Bush by five percentage points (47% to 52%), the message the Bishops were preaching about morality did not inspire all Catholic voters. In the exit poll data, moral issues ranked third, behind terrorism and jobs as the reason Catholics chose one candidate over the other. This paper will examine the Catholic vote in the 2004 Presidential election. We will look at subgroups of Catholic voters - women, men, conservatives, liberals, moderates, church attenders and many other groups - to provide a profile of the Catholic voter. The paper will also explore why Catholics voted the way they did and the potential impact the Bishops may have had in this election. The paper will use the 2004 exit poll data as well as a survey among Catholic voters that BRS conducted in July 2004 among 2,223 likely Catholic voters. This survey asked not only intentions of voting but the importance of issues to voters and the influence of the Bishops on voting decisions and contained an oversample of Hispanic Catholics.

***The Impact of Attitudes toward Stem Cell Research on the Vote for President*, Jon Miller, Northwestern University; j-miller8@northwestern.edu, Rafael Pardo, Foundation BBVA; rpardo@fbbva.es, and Linda Kimmel, Northwestern University; l-kimmel@northwestern.edu**

The extensive debate about stem cell research in the 2004 election marks the first time that a science issue has become a major issue in a presidential election in the United States. This paper will report on the results of a panel study of U.S. adults that began with a national sample of approximately 2000 adults in December, 2003, asking an extensive set of items about biotechnology and environmental issues. In October, 2004, each respondent was asked to complete a second online questionnaire (Knowledge Networks conducted all three waves of this study) concerning their involvement in the campaign and their assessment of the candidates and issues. Immediately following the election, each respondent was asked to complete a third survey which asked whether they voted in the election, for which candidate they voted, and the reasons for their choice. The three waves of this study provide an unparalleled opportunity to map the influence of a science issue in the context of more traditional voting models. This analysis begins with a background or stability model (using LISREL), building on the work of Bereleson, Lazarsfeld, and McPhee. A second model includes a set of issue variables, building on the work of Campbell, Converse, Miller, Stokes, and others. These issues may be described broadly as domestic issues, foreign policy issues, and moral issues (abortion, same-sex marriage). A third model inserts a variable that incorporates both the importance of the stem cell issue and the direction of each respondent's attitude (for or against). Supporters of stem cell research generally rated the issue as less important in their final candidate decision; opponents of stem cell research tended to rate the issue as being relatively more important. The analysis shows that stem cell attitudes changed few votes, but generally reinforced previous attitudes toward each candidate.

SURVEYING RACIAL ATTITUDES IN AMERICA: NEW DATA SOURCES AND SELECTED RESULTS

***Understanding Perceptions of Racial Discrimination*, Stanley Feldman, SUNY; Stanley.Feldman@sunysb.edu, and Leonie Huddy, SUNY; lhuddy@notes.cc.sunysb.edu**

White Americans vary enormously in the extent to which they perceive current racial discrimination against African-Americans. These differences in perception are extremely important politically, wielding powerful influence on support for government assistance to blacks. Whites who perceive widespread racial discrimination support government policies to assist blacks whereas those who disavow contemporary racial discrimination do not. Researchers are divided about how to interpret white denial of discrimination. From the perspective of symbolic or modern racism, it is seen as a clear indicator of prejudice; but others, critical of the modern racism approach, suggest that it reflects an awareness of increased opportunities for blacks and a substantial reduction in prejudice in contemporary American society. We use a new, two-wave national telephone (CATI) survey to examine white Americans' beliefs about black discrimination. In addition to standard questions on current levels of racial discrimination in American society, we include a number of new measures that allow us to carefully assess different aspects of whites' perceptions of discrimination, including: (1) respondent assessments of the validity of a series of factual statements about evidence both for and against the existence of contemporary discrimination in various facets of American life, allowing us to identify whether whites accurately assimilated or argued against evidence for the existence or absence of racial discrimination; (2) measures of factual knowledge concerning evidence of past racial discrimination in the United States; and (3) a series of experimentally altered policy questions that allow for more fine-grained tests of the effects of racial attitudes on policy support. We will address key questions about the nuances and predictors of perceptions of discrimination, thereby extending existing research, which has failed to examine closely the underlying meaning of perceived levels of contemporary racial discrimination.

***The National Politics Study: Ethnic Pluralism and Politics in the 21st Century*, Vincent Hutchings, University of Michigan; vincenth@umich.edu, Cara Wong, University of Michigan; cjwong@umich.edu, Ron E. Brown, Wayne State University; aa4723@wayne.edu, and James S. Jackson, University of Michigan; jamessj@umich.edu**

Immigration and differential birth rates are producing major changes in the ethnic and racial composition of the population and, eventually, of the American electorate. Over the next several decades, individuals from Asian American, Caribbean American, Hispanic, and African American groups will constitute an increasingly substantial share of the population until mid-century, at which point non-Hispanic whites are projected to comprise a numerical minority. These groups differ greatly in their ethnic and racial identity, political consciousness, and attachment to American ideological and political institutions. The National Politics Study (NPS) is an ambitious project undertaken by the University of Michigan's Program for Research on Black Americans and Center for Political Studies. The primary goal is to gather

comparative data about political attitudes, beliefs, aspirations, and behaviors in order to advance the study and knowledge of race and ethnic involvement in politics. This study is unique in its scope and sampling plan, with nationally representative samples of African Americans, non-Hispanic whites, Caribbean Americans, Hispanics, and Asian Americans. Topics of the survey include voting preferences, party affiliation, organizational membership, immigration, racial consciousness, acculturation, and views of government policies. In this paper, we report preliminary analyses of this extensive new data set, in particular, focusing on the several new within and between group comparisons that are possible given the sample size and composition. The research objectives are to examine how identity, consciousness, ideological beliefs, socio-demographic, social and economic status factors influence the political lives of people that differ in ethnic and racial background, immigration and citizenship status, and geographic dispersion. Existing theories and models of political behavior have not taken such considerations into account, so that the results of this study will do much toward modifying and extending existing theories that have been developed without benefit of data from such a racially diverse sample.

***A Multi-Ethnic Measure of Racial Residential Preferences: New Results from Chicago and Detroit*, Maria Krysan, University of Illinois at Chicago; krysan@uic.edu, Reynolds Farley, University of Michigan; renf@umich.edu, and Mick Couper, University of Michigan; mcouper@umich.edu**

Residential segregation has been called the "structural lynchpin" of American society, creating and perpetuating racial inequalities in such areas as education, employment, health, income, and exposure to crime. Since the 1970s there has been interest in the role of people's attitudes about living with people of different races in shaping patterns of segregation. The traditional measure of racial residential preferences was limited to techniques asking about just two groups: blacks and whites. In the 1990s, meeting the challenge posed by the increasing diversity of the U.S., Charles (2000) developed a new technique in which respondents "drew their own" neighborhood representing their ideal racial/ethnic mix. By including Whites, Blacks, Asian Americans, and Hispanics among the possible groups to consider including in the neighborhood, the measure thus tapped preferences in a multi-ethnic context. In 2004, parallel computer-assisted face-to-face surveys in Detroit (n=734) and Chicago (n=800) replicated and expanded this technique by (1) including Arab Americans among the groups respondents were asked to consider; (2) asking respondents to draw both the most and the least desirable neighborhood composition; (3) using open-ended questions to gain insight into the reasons behind the preferences; and (4) including over-samples of respondents in racially mixed neighborhoods, so as to gather racial residential preferences among those who are not themselves currently segregated. This paper analyzes the neighborhoods that respondents drew, with the following key objectives: (1) to identify which racial/ethnic groups they include and exclude-and in what numbers; (2) to establish the key demographic and social predictors of racial residential preferences; (3) to identify the social psychological underpinnings of racial residential preferences; (4) to compare the responses between those in integrated and segregated neighborhoods.

***On Race and Orientation of Juvenile Court Workers: Perspective on the Difference that Difference Might Make*, Geoff Ward, Northeastern University; g.ward@neu.edu**

Friday 8:15 am -9:45 am - AAPOR ABSTRACTS

INFORMED CONSENT

***Development and Testing of Informed Consent Questions to Link Survey Data with Administrative Records*, Nancy Bates, U.S. Census Bureau; nancy.a.bates@census.gov**

When permission is granted from respondents, social security numbers (SSNs) are commonly used to link federal agency survey data with administrative records for methodological, policy, and other statistical purposes. Over the decade, respondents' willingness to provide SSNs has declined significantly. This is presumably related to increased concerns about identify theft, loss of control over private information, and concerns about confidentiality violations (computer "hackers") (Singer; 2001; Gerber; 2001). The Census Bureau and other federal agencies have begun research, development, and evaluation of new record linkage and search methodologies that do not require direct collection of SSN (or in some cases, only the last four digits of SSN). Adoption of these new methods could mean adding new informed consent requests or notifications about planned data linkages. In 2003, the Census Bureau conducted cognitive interviews to test and develop alternative informed consent questions that do not explicitly ask for SSNs. In 2004, a new "opt-out" informed consent question was tested in the Questionnaire Design and Evaluation Research Survey (QDERS). The QDERS was a split-panel random-digit-dial survey that yielded 4,000 completed interviews. The new informed consent question was tested against traditional requests for SSN and a request for the last four digits of SSN. Survey results indicated wide variation in refusal rates to the requests for informed consent among four treatments (request for SSN with an explanation; request for SSN in the traditional fashion; request for last four digits of SSN; and the new opt-out informed consent question). We also discovered significant subgroup differences among those who refused to provide informed consent within and across treatments. The paper concludes with a discussion of how these findings may impact informed consent policies in the federal statistical system.

***The Effect of Introductory Consent Information on Response Rates and Respondent Understanding in a Telephone Interview*, Mary Losch, University of Northern Iowa; mary.losch@uni.edu, Shelly Campo, University of Iowa; shelly-campo@uiowa.edu, and Gene Lutz, University of Northern Iowa; gene.lutz@uni.edu**

Consent information provided in survey research has come under increasing scrutiny by Institutional Review Boards (IRBs) in recent years. Researchers often argue that the requirement for extensive and detailed consent information at the outset of a minimal risk telephone interview leads to reduced response rates and does not increase participant protections. IRBs, on the other hand, often argue that such information is needed to ensure participant understanding. Little empirical research has addressed the effects of varying consent information on response rates and participant understanding in survey research. In this experiment, 1,000 participants in a telephone survey of cancer screening knowledge and behavior were randomly assigned to either a standard consent condition or an expanded consent condition. Analyses of the effects of consent type on response rates, break-offs, measures of understanding of consent elements, and preferences for standard versus expanded consent will be presented.

***Do Characteristics of Consent Forms Affect the Response Rate?*, Dragana Bolcic-Jankovic, Center for Survey Research, UMass Boston; dragana.bjankovic@umb.edu, Brian Clarridge, Center for Survey Research, Umass Boston; brian.clarridge@umb.edu, and Floyd J. Fowler, Jr., Center for Survey Research, Umass Boston; fjfowler@fimdm.org**

Researchers are required to obtain authorization from patients to gain access to protected health information for research purposes under the HIPPA Privacy Rule. The challenges faced in obtaining signed approval from individual patients include: 1) the forms are idiosyncratic by hospital, 2) the forms are long, and 3) the forms are often written in legal language. So although obtaining signed authorization forms may help protect hospitals, it is often challenging for researchers collecting the forms and it creates a potential for lower response rates. This paper reports on a recent telephone survey of previously hospitalized patients in which respondents were asked if they would be willing to allow researchers to review their medical records to get more complete information about their hospital stay. The study included patients from 16 Massachusetts hospitals. Nine hospitals required the use of their own distinct forms, which they designed themselves, while the other 7 were variations of these distinct forms. Only those respondents who agreed over the phone to be mailed authorization forms are included in the analysis. The goal of this paper is to learn whether characteristics of the consent form, rather than some patient or hospital characteristics, are associated with the rates at which a signed consent forms are returned. Forms varied with respect to characteristics such as length, complexity of the form, distinctness of the institution's name on the form, confidential information requested (SSN), witness signature required, and whether a copy of the consent form was included for patients' records. The analysis focuses on whether or not the characteristics relate to the receipt of a signed consent form. Preliminary results suggest that the format of the consent form matters.

***When silence is not an option: finding a better balance between confidentiality and respondent protection in ethical codes*, Iain Noble, Department for Education and Skills; iain.noble@dfes.gsi.gov.uk, and Nick Moon, NOP World; nmoon@nopworld.com**

In most social research ethical codes (including AAPOR's) one key principle is that respondents should not be endangered by participating in the research. Confidentiality is taken as a cornerstone of this protection: release of the names of respondents to third parties is not allowed unless the respondent specifically waives confidentiality. We present the case that researchers should not only be ready to ignore confidentiality but also, in certain circumstances, make preparations before a survey to implement such disclosure as rapidly as possible. We believe that relevant codes, such as AAPOR's should be amended to cover circumstances where disclosure can, indeed, must take place. We discuss an incident on a survey of young people in England where information came to light about specific ways in which a 14 year old respondent might be at risk. UK law provides that those with privileged access to information in the case of children have a right to contact appropriate authorities in such cases without the permission of the respondent. The sponsoring authority, a government department with responsibility for children's and young people's welfare had a specific duty to act. It might be argued that these are rare and specific circumstances that can be resolved as and when they occur in an ad hoc manner. We disagree and believe both that they are more common than assumed and also that preparation is crucial if prompt action may be required. Further, researchers and interviewers need to be sensitised to the existence of such problems, be ready to recognise them and, crucially, know what needs to be done in such circumstances. The current silence of codes of ethics on what would justify breaches of confidentiality militates against this. It is time this changed because as citizens we should take a wider view of our responsibilities to respondents.

MAKING CONTACT

***Gaining Efficiencies in Scheduling Callbacks in Large RDD National Surveys*, Jeffery Stec, InteCap, Inc.; jstec@intecap.com, Paul J Lavrakas, Nielsen Media research; paul.lavrakas@nielsenmedia.com, and Charles D. Shuttles, Nielsen Media Research; Chuck.Shuttles@NielsenMedia.com**

In the past 20 years, since RDD telephone surveys first achieved their now recognized stature as a primary means to survey a random sample of the U.S. population, the effort needed to properly and fully dial and dispose of the telephone numbers in the sampling pool has grown extensively. As such, telephone survey researchers are constantly seeking cost-beneficial ways to schedule their calls efficiently so as to minimize costs without adding to nonresponse (and possible nonresponse bias). Four times each year (January/February, March/April, June/July, and October/November), Nielsen Media Research (NMR) uses a list-assisted RDD frame to sample respondent households for its national (all 50 states) dual-stage mixed-mode diary surveys of television viewing in the United States. In each of these surveys more than 1,200,000 RDD telephone numbers are used for as the initial sampling pool. Following fixed calling rules, which have existed since the mid-1990s, and which space up to 15 call attempts per telephone number at various times of day over a 14-16 day field period, NMR makes approximately 8,000,000 dialings during each of these seasonal surveys. This effort leads to far better than average response rates compared to most other commercial RDD surveys. This paper will present analyses using NMR's 2003 and 2004 calling data from four diary surveys to investigate where efficiencies are likely to be gained by changing the fixed calling rules now used. Past research along these lines suggested that efficiencies may be gained, while keeping the total number of dialings fairly constant, if fewer call attempts were made to certain call history patterns and more call attempts were made to certain other call history patterns.

***Calling strategies, Response rates, Efficiency and Representativeness*, G Donald Ferree Jr, University of Wisconsin; gferree@ssc.wisc.edu**

There is an ongoing concern with efficiency in telephone survey operations, with regard both to the impact of overall survey cost and the payoff in representativeness of various strategies, focusing especially on response rates and their significance. CATI systems often provide ways to fine tune allocation of calls based on the prior history of a given telephone number. As part of its ongoing efforts to improve the efficiency of its survey operations, the University of Wisconsin Survey Center examined a number of both statewide and national surveys examining the impact of calling strategies on achieving completed interviews. Surveys varied by length of field period - approximately one week, one month, and one quarter year, and included both statewide and national RDD studies. Among the factors examined are time period of call (e.g. weekday daytime, weekend evening), number of previous calls, whether or not earlier attempts resulted in a human contact, and so on. Data are stored for each telephone number in the original sample on its call history (day, time, and result of each call, as well as such derived characteristics as how many previous calls involved contact with either an informant or designated respondent, whether there had been a refusal by an informant or respondent). These data permit not only examination of patterns of success that might increase efficiency, but allow one to "simulate" the impact of making fewer callbacks, giving up on refusal conversion, and the like by examining the characteristics of cases which would have been lost, for instance, if callbacks had stopped at a lower number. The paper finally makes recommendations on the kind of information which should be retained so that future decisions on calling strategies can take advantage of each organization's experience with different types of surveys.

***An Experimental Test of Answering Machine Message Content to Improve Response Rates*, Charles D. Shuttles, Nielsen Media Research; Chuck.Shuttles@NielsenMedia.com, Paul J Lavrakas, Nielsen Media research; paul.lavrakas@nielsenmedia.com**

One of the most ubiquitous forms of telephone technology, other than the actual telephone itself, is the answering machine/voice mail. A majority of households utilize this type of technology for recording messages from callers and/or screening calls. Yet, encountering an answering machine is positively correlated with eventually completing a call with a household (Stec, Lavrakas, & Shuttles. 2004). We reasoned that leaving a message on the answering machine or voice mail, depending on the content of the message, could raise the subsequent propensity for a respondent's to participate in a survey. This paper will report on a large-scale design experiment testing different answering message contents. Within this design was embedded a 2 x 2 factorial experiment. This study was performed as a part of the September - October 2004 Nielsen TV Ratings Diary study - a multi-stage survey conducted in 25 metropolitan areas throughout the United States. Six different message contents were delivered on the last calling attempt (i.e., 15-attempt structure) of the RDD-stage call to those households that had not previously been reached and had a matched mailing address. Each message version was left on approximately 900 answering machines/voice mail from an original RDD sample of 249,000 numbers. The message versions included testing the standard message that has been left by Nielsen for many years (Control condition), a very short version with minimal information, a longer explanatory/persuasive version (including a mention of cash incentive that is included with the mail-stage diary survey), and other variations. The dependent variables examined will include the RDD-stage contact and response rates and the mail-stage response rate.

***Adminstrative Interventions and Their Impact on Response Rates in Military Attitude Surveys*, James Caplan, Defense Manpower Data Center; caplanjr@osd.pentagon.mil, and Barbara Quigley, Defense Manpower Data Center; quiglebm@osd.pentagon.mil**

In spite of declining national survey return rates the survey program at the Defense Manpower Data Center has successfully increased response rates by using a variety of administrative techniques. Response rates to frequent, lengthy surveys have either increased or remained the same over the past three years. This paper will provide details and show the effects of these techniques. Web-based surveys were first administered to active-duty military in July 2002. Since then five additional cross-sectional surveys have been fielded to that population. One-page follow-up surveys to nonrespondents identified reasons for nonresponse, including eligibility, computer/Internet access, training or temporary duty. Follow-up survey results and e-mail comments showed nonrespondents were neither negative toward surveys nor different in attitudes but rather encountered technical problems in responding or were unavailable because of training, moving or deployment. The results were later used to further improve response rates. Sample members were asked to update personal information prior to the fielding period. E-mail addresses including purchased ones were used for reminders to sample members. E-mails contained hyperlinks to the survey. This paper will outline effects of e-mail reminders, which produced better results than postal reminders. Between 40% and 59% of completed returns resulted from valid e-mail addresses gleaned from personnel records. E-mail reminders increased the likelihood of getting a response. For example, the use of purchased e-mail addresses increased response rates by 11.6% at a cost of less than 37 cents for each completed survey. Demographic comparisons using service, paygrade, minority status and gender showed changes in response rates for the six Web surveys. The effect of these changes on response rates for the different demographic comparisons will be discussed. Discussion will center on utilizing feedback from one-page follow-up surveys, comments and e-mail messages to improve response rates.

THE WEB AND OTHER MODES I

***Mode Effects for Hybrid Telephone/Internet Surveys and Reaching Cellphone-Only Households*, Dale Kulp, Marketing Systems Group; DKulp@m-s-g.com, Melissa Herrmann, ICR/International Communications Research; mherrmann@icrsurvey.com, David Dutwin, ICR/International Communications Research; ddutwin@icrsurvey.com, and Steve Lavine, Common Knowledge; steve@your2cents.com**

The evolution of our national telecommunications network along with advances in telephony are creating new telephone service options in the consumer marketplace. Although these new service options increase consumer choice, convenience and are often less expensive, they also engender challenges for researchers. Where the telephone researcher's traditional focus was limited to telephone versus non-telephone household biases, coverage issues now include cellular-only households and the likely undercoverage of households with non-traditional telephone landline equivalents (e.g., households subscribing to local phone service offered by cable-TV companies). International Communications Research has managed an ongoing daily telephone omnibus service (CENTRIS) for seven years. This survey focuses on and tracks many aspects of consumers' communications, entertainment, and most importantly, in-home PC/internet activity. MSG/GENESYS designed this survey and currently oversees the sampling and survey implementation. With the help of Common Knowledge, this survey vehicle is currently migrating from a strictly RDD sample to a dual frame/dual mode telephone/internet design. The primary purpose of this paper will be to communicate design issues, implementation experience and specific results, including mode variations in the standard set of tracked behaviors, attitudinal items and demographics. The paper will also review weighting and estimation issues and final procedures, specifically focusing on the overlapped portion of the sample frame (households with in-home internet access). Finally, the authors will review findings relating to samples of cellular-only and out-of-frame respondents developed through the internet portion of the sample.

***The Transition From Telephone to Online Data Collection in Time Series Measurement: The estimation of mode effects*, Jon Miller, Northwestern University; j-miller8@northwestern.edu, Linda Kimmel, Northwestern University; l-kimmel@northwestern.edu, william mccready, Knowledge Networks, Inc.; bmccready@knowledgenetworks.com, and Mike Dennis, Knowledge Networks; mdennis@knowledgenetworks.com**

With the growth of serious response problems in cold random-digit-dialed national samples, numerous analysts have turned to online samples to improve response rates and facilitate timely data collection. This paper reports on the use of Knowledge Networks' national adult sample to continue some of the time series items previously collected from telephone RDD samples. Attention is focused on responses asking a respondent to rate his or her level of interest in selected public policy issues and his or her level of understanding of the same issues. These items have been reported extensively in the NSF's Science and Engineering Indicators series and have been used to estimate the proportion of adults attentive to science and technology and other issue clusters. Recent replications of this work using Knowledge Networks' adult panel indicate that respondents tend to rate their interest in and knowledge of the same public policy issues significantly lower in online data collections than in earlier telephone interviews. It is hypothesized that there is a mode effect. The almost 20 years of telephone data are remarkably stable, with almost all variations following major events such as the Persian Gulf War in 1991.

The across-the-board reduction in self-assessed knowledge levels in 2003 and 2004 suggest that the absence of an interviewer may have removed a source of implicit social pressure to inflate the level of interest and that the presentation of a "menu" of policy issues may have provided a different context than the sequential presentation of issues in telephone interviews. A series of descriptive and analytic techniques (including structural equation models) are used to assess the mode hypothesis.

***In Search of Equivalency Across Modes: Experimental Results Comparing Alternative Question Formats for Eliciting Dates in Telephone and Web Modes*, Leah Christian, Washington State University; lmchristian@wsu.edu, Don Dillman, Washington State University; dillman@wsu.edu, and Jolene Smyth, Washington State University; jsmyth@wsu.edu**

Words convey meaning in interview and self-administered surveys, but respondents to paper and web surveys infer additional meaning from the symbols and numbers used in, and the graphical design of questionnaires. We report the results of several experiments to compare different uses of words, symbols and graphics designed to instruct respondents to report date responses in a desired format. These experiments were embedded in a series of surveys, three web and one telephone, of randomly sampled Washington State University undergraduate students. We find that the manner in which respondents were verbally instructed to report the month and year in the query had a powerful effect on the phone survey; however, on the web survey, the stem change had less effect since respondents were provided additional visual instructions located with the response space, making them more accessible at the time of response. The most dramatic effect from the web surveys is that using symbols (e.g. MM\YYYY) to convey the number of digits respondents should use to report the month and year strongly increased the use of the desired format. In addition, placing the symbols in close proximity to each answer space increased their visibility at the time of response and thereby resulted in more respondents using a four-digit year response format. Our research suggests that verbal language changes have powerful effects in phone surveys, where aural communication is the primary mode of communication. In web surveys where other visual languages can be used to support verbal language, symbols and graphical location can effectively be used to instruct respondents to report answers in a particular format.

***Survey Mode Effects; Comparison between Telephone and Web*, Howard Speizer, Market Strategies; howard_speizer@marketstrategies.com, Reg Baker, Market Strategies; reg_baker@marketstrategies.com, and Karin Schneider, Market Strategies; Karin_Schneider@Marketstrategies.com**

Web surveys are frequently faster than other modes, less expensive, and with modern software they can accommodate even the most complex questionnaires. However, problems abound, including limited Internet penetration, no email address analogue to the national telephone number sampling frame, and reasons to suspect that people may respond differently on the Web than they do on the telephone. This last issue-potential mode differences between Web and telephone-is the main focus of this paper. The survey data for this research are drawn from a monthly customer satisfaction survey for an energy distribution company. Representatives from a sample of large business customers are asked to rate the company on a number of satisfaction criteria. Each month a new sample is drawn and an invitation to complete a Web survey is mailed to representatives at each company selected. Non-responders are followed up by telephone. On average, each month, we achieve a 50% response rate with roughly 40% of the respondents reporting by Web and the remainder by telephone. Our analysis has found significant differences across modes for a number of items, many of which are consistent with mode effect hypotheses that have been posited between self-, and interviewer-, administered modes. Self-selection of mode has partially biased our results, and in the spring of 2005 we will test these differences by randomly assigning sample members to either Web or telephone treatments. This research provides valuable insight into the impact of mode on response to a number of commonly asked customer satisfaction questions. The research expands our knowledge about respondent preference for survey mode and assists efforts to improve participation rates for surveys that employ a multi-mode design. The paper also provides some practical experience on the problems associated with transitioning respondents from one mode (telephone) to another (Web).

***How Does Social Desirability Affect Responses?: Differences in Telephone and Online Surveys*, Humphrey Taylor, Harris Interactive; htaylor@harrisinteractive.com, David Krane, Harris Interactive; dkrane@harrisinteractive.com, and Randall Thomas, Harris Interactive; rthomas@harrisinteractive.com**

Many have suggested that human interviewers can invoke social norms with regard to propriety in responding. We conducted a nationwide phone survey in parallel with a nationwide online survey. We selected a series of items that we expected would vary with regard to socially desirable responding (e.g. belief in God, driving over the speed limit, giving money to charity, gambling, etc.). We found that for items separately scaled as undesirable, online respondents were more likely to endorse them as being true for them. For items scaled as desirable, phone respondents were more likely to admit to them being true of them. As an example, online respondents admitted to often driving over the speed limit at a higher rate. In contrast, online respondents were less likely to admit going to church, mosque, or synagogue. We will also present results comparing the ratings of these items with the endorsements of these items in both modalities.

SAMPLING I

***A Test of a Combined RDD/Registration-Based Sampling Model in Oregon's NEP Survey*, Joel Bloom, University of Oregon; jbloom@uoregon.edu, Warren Mitofsky, Mitofsky International; mitofsky@mindspring.com, and Joe Lenski, Edison Media Research; lenski@edisonresearch.com**

The 2004 election provided a unique opportunity to test a dual frame sample design using random-digit-dialing (RDD) and registration-based-sampling (RBS). With RDD, one can reach all households with a working phone. Self-reports of registration are, however, overstated. With RBS, everyone is a registered voter, and every name is accompanied by a wealth of information about the individual and past voting. Unfortunately, RBS lists are incomplete. While RBS lists typically include 60-70% of registered voters, another 10% are lost for various reasons. If any of those covered differ from those excluded, RBS could produce biased results. In this paper we analyze the results of a telephone survey conducted in lieu of an exit poll for the National Election Pool by the Oregon Survey Research Laboratory at the University of Oregon. The survey of 907 voters was conducted October 22-31, with a combination of RBS and RDD sampling. After de-duplication of the two frames, every voter reached in the RDD sample represented someone not covered in the RBS list. We analyze the performance of the samples from the two frames, and determine whether this dual frame methodology offers an efficient compromise for researchers concerned about the drawbacks of both methodologies.

Assessment of Address Frame Replacements for RDD Sampling Frames, **Michael Link, Centers for Disease Control and Prevention; MLink@cdc.gov, Michael Battaglia, Abt Associates Inc.; Mike_Battaglia@abtassoc.com, Pamela Giambo, Abt Associates; Pamela_Giambo@abtassoc.com, Martin Frankel, Abt Associates and Baruch College, CUNY; Martin_Frankel@abtassoc.com, and Ali Mokdad, Centers for Disease Control and Prevention; AMokdad@cdc.gov**
It is becoming more difficult to obtain valid and reliable public health data through the use of random digit dial (RDD) telephone surveys. As a result, researchers are considering alternative modes such as mail surveys as complements or alternatives to telephone survey approaches. Traditionally, however, lack of a representative sampling frame has been a primary barrier to conducting mail surveys of the general public. Is this still the case? Advances in electronic record keeping combined with database mining may now allow researchers to develop and sample from a frame of addresses with coverage that rivals most current RDD sampling methods. As part of an on-going effort to improve coverage and response to the Behavioral Risk Factor Surveillance System (BRFSS), a quantitative assessment of several potential address-based sampling frames for use with a mail survey was conducted. This paper provides an overview of the issues related to the development of a valid and reliable address-based sampling frame; comparison of estimated coverage rates between the address frame alternatives and data from the 2000 Census and post-censual estimates; and details of the sampling strategy used in selecting addresses for a six state pilot test comparing the traditional, telephone-based BRFSS with a mail version completed by a random sample of respondents drawn from an address-only frame.

Measuring the efficacy of a clustered design in producing small area estimates in a survey of health insurance status, **Tom Duffy, ORC Macro; thomas.p.duffy.jr@orcmacro.com, and Ronaldo Iachan, ORC Macro; ronaldo.iachan@orcmacro.com**
The Ohio Family Health Survey (FHS) is a telephone survey of the health and health insurance status of adults and children in Ohio. The FHS prescribed confidence intervals for estimates of insurance status for several population subgroups: rural regions, ethnic minorities, families in poverty, families with children, etc. These constraints required a complex sample design that over-sampled on many levels. One major consideration in the allocation of sample across the 88 counties in Ohio, was obtaining reliable estimates of the health insurance status of children under the age of 18; precision constraints called for reliable estimates either for individual counties (in the case of the largest counties), or for clusters of similar counties. Sampling clusters were created using a k-means iterative process, using variables shown to correlate with insurance status in other research (e.g. household income, employment status, race). Geographic contiguity was one of several variables in the model. The objective was to obtain reliable estimates for clusters of similar, small counties without having to be completely dependent on other small area estimation techniques. As a result, clustering was used to produce design strata that would most effectively meet this analytic objective. This paper will assess how well the clustered design did in approximating county-level estimates of health insurance coverage.

The Impact of Cluster (Segment) Size on Effective Sample Size, **Steven Pedlow, NORC/University of Chicago; pedlow-steven@norc.uchicago.edu, Yongyi Wang, NORC at the University of Chicago; wang-yongyi@norc.org, and Colm OMuircheartaigh, NORC, Univeristy of Chicago; colm@uchicago.edu**
Area probability surveys with face-to-face interviewing usually involve a cluster sampling approach. Cluster sampling entails selecting small areas (we call these "segments") in which every housing unit is enumerated. The lists of enumerated units are then subsampled for inclusion into the sample. People who live in these clusters of sampled housing units often are similar on some of the outcome variables, especially financial variables. The similarity of the sample clusters is measured by the intraclass correlation, which is usually positive. Positive intraclass correlations indicate that the sample is not quite as good as a simple random sample of the same sample size. This inefficiency is measured by the design effect (DEFF), and results in a smaller effective sample size (n/DEFF). Larger segments would tend to result in smaller intraclass correlations (e.g., see Hansen, Hurwitz, and Madow, 1953), but would also result in higher costs (to enumerate every housing unit). Most area probability studies make a choice on segment size, but then cannot examine the effect of different segment sizes on the intraclass correlations. Our talk will focus on data from the Making Connections project, sponsored by the Casey Foundation, which collects data from entire neighborhoods. With this data, we compiled a frame of all addresses in the neighborhood and drew an element sample. We are now able to divide neighborhoods into segments of varying sizes to study the impact on the intraclass correlations for various response variables. This research can help choose appropriate segment sizes in future surveys.

Matched Sampling, **Douglas Rivers, Stanford University and Polimetrix, Inc.; doug@polimetrix.com**
This paper describes a method for selecting a sample designed to represent a listed population when the sample source is a large panel with an unknown probability selection mechanism. The method produces samples that can be balanced on a large number of variables. The method has several advantages relative to conventional post-stratification or raking procedures. The method is illustrated using over two million interviews conducted by Polimetrix during the 2004 U.S. Presidential election campaign.

YOUNG AND RESTLESS VOTERS

First Time Voters in the 2004 Election: Who Are They and What Was On Their Minds?, **Jennifer De Pinto, CBS News; jdp@cbsnews.com, and Megan Thee, The New York Times; thee@nytimes.com**
Thirteen million more Americans cast a vote for president in 2004 than did so in 2000 and many of them were voting for the first time. The number of groups mobilizing and registering new voters in 2004 was unprecedented, and in turn, several states reported record voter registration numbers. This paper will set out to determine: 1.) What impact first time voters had on the increased total voter turnout in 2004, 2.) Who these first time voters are, 3.) How they compare to first time voters in previous presidential elections, 4.) What issues drove these voters to the polls, 5.) Which presidential candidate they supported. The analysis will draw on data from the 2004 NEP national exit poll, as well as trend data from past national exit polls. In addition, pre- and post-election polls conducted by CBS News/The New York Times will be used to examine the attitudes of potential and actual first time voters and the likelihood that they will continue to participate in the political process.

Young Cell Phone Users and Voting Behavior in Georgia in 2004, **James Bason, University of Georgia; jbason@uga.edu**
Prior to the 2004 general election, some political observers suggested that public opinion pollsters may have underestimated support for Democrat John Kerry due to cell phone only users, primarily young, urban first time voters. Zukin (2004), based on Edison/Mitosfsky exit poll data, has suggested that this supposition was largely unfounded, as the 30 and under age group split 54 percent for Kerry and 45

percent for Bush, numbers very close to the results obtained by Zogby in a text message study among 120,000 cell phone users subscribed to a MTV Rock the Vote text message mailing group. The current research examines voting behavior of a random sample of 2,000 University of Georgia students during November 2004 to determine the voting behavior of cell phone only users during the 2004 general election. Using a web based survey instrument, respondents were asked a variety of items related to political attitudes, including voting behavior during the 2004 general election. Respondents were also asked to indicate cell phone usage patterns. The unique nature of the study will also allow comparison of University of Georgia students' voting behavior following the general election to a voting behavior of a similar age group of prior to the general election in a statewide RDD telephone survey conducted just three weeks before the election. It is generally assumed that cell phone only users are primarily young people, so using a sample of undergraduate students at a major state university will allow us to directly test this assumption, as well as determine if the group of cell phone only users differed from non-cell phone only users during the 2004 general election.

***Campaign Involvement and the Internet: A Panel Study of Howard Dean Supporters*, Courtney Kennedy, Pew Research Center; ckennedy@survey.umd.edu, Scott Keeter, Pew Research Center; keeters@people-press.org, and Cary Funk, Pew Research Center; cfunk@pewresearch.org**

In his quest for the 2004 Democratic presidential nomination, Howard Dean energized hundreds of thousands of supporters nationwide, many of whom were engaging in the political process for the first time. In addition to its ideological appeal, the Dean campaign was considered to be revolutionary in its use of the Internet for facilitating donations and coordinating volunteers. This paper presents findings from a unique, web-based panel survey of Dean campaign supporters conducted by the Pew Research Center in September and November of 2004. The study examines the political orientation and demographic composition of the campaign volunteers as well as the social dynamics of their digital interaction. The study also explores the supporters' evaluations of the presidential election and the future of the Democratic Party. Among the many intriguing results from the study: 44% of Dean supporters were participating in their first political campaign; the war in Iraq was the most important issue for 69% of supporters (99% oppose the war, but 45% oppose a quick withdrawal). Dean for America drew random samples of its supporter database and contacted these supporters by e-mail on behalf of Pew. With one reminder e-mail, the first survey drew 3,925 respondents for a response rate of 13%. These respondents were subsequently invited to complete a post-election follow-up survey, which is underway as of the end of November. A second post-election survey is also underway with a new sample of supporters. Non-response bias in the survey will be estimated using participation data contained in the original database of supporters.

CONTEXT AND FRAMING

***The Impact of Question Order on the Measurement of Party Identification*, Trevor Tompson, The Associated Press; ttompson@ap.org, and Mike Mokrzycki, The Associated Press; mmokrzycki@ap.org**

The measurement of party identification in public opinion polls became a subject of controversy in the 2004 presidential election campaign. Most public opinion polls ask party identification as part of a battery of demographic questions at the end of a survey. We have conducted a year-long experiment involving more than 20 national RDD telephone surveys and 25,000 interviews. In each survey, half of the respondents were randomly assigned to an experimental condition where a party identification question was asked at the start of the interview, and the other random half were asked the same question at the end of the interview along with the demographic questions. About half of the questionnaires in the experiment were standard tracking surveys, where we can observe the impact of the experiment over time with a similar set of questions. The content of the other questionnaires varied widely, including political, social, economic, and health-related subjects. Our research shows that in some situations, responses to the party identification question can be subject to significant question order effects. The experimental design also allows us to show that asking party identification early in the questionnaire can impact the responses to subsequent questions. We also present a rare party identification trend covering the 2004 election campaign that is unaffected by potential question order effects, and compare it to the party identification trend obtained when asking the question at the end of the interview. We use our findings to inform a broader discussion of party identification and polling in the 2004 election.

***Order Effects and Vote Preference in the 2004 Presidential Election*, Janice Bell, Ipsos-Public Affairs; janice.bell@ipsos-na.com**

Standard practice in public opinion surveys leading up to the presidential election this past November was to place horserace questions before questions on approval of incumbent performance and other policy-related issues. One could easily argue that asking vote intention before approval would bias the results (respondents with strong intention to vote for the incumbent would respond more positively about job performance, and vice versa). Over the course of 2004, however, polls conducted for The Associated Press by Ipsos-Public Affairs typically placed questions relating to presidential job approval at the front of the survey, before horserace questions. In a nationwide survey conducted the last weekend before the 2004 presidential election, the order of these two sections was randomized. This paper finds that the order of these sections produces interesting effects not only for overall support for Bush, but also for the share of "definite" Bush and Kerry supporters. In addition to findings regarding order effects, this paper also provides a broad perspective on the context of the election, as evidenced through an examination of the relationship between the most important issues of the election and overall candidate support.

***"Did Ballot Order Matter at 'The Epicenter'?" : An Evaluation of Candidate Ballot Order Effects in the 2004 Ohio Elections*, Eric Rademacher, University of Cincinnati, Institute for Policy Research; eric.rademacher@uc.edu, Jason Minser, UC, Institute for Policy Research; jason.minser@uc.edu, and Kim Downing, University of Cincinnati; kim.downing@uc.edu**

In Ohio, the so-called 'Epicenter' of the 2004 national elections, did the order in which presidential candidate names were read matter in the levels of support they received in telephone surveys? Researchers conducting pre-election surveys by telephone often administer multiple forms of trial heat questions that rotate the order in which candidate names are presented. In Ohio, the University of Cincinnati's Ohio Poll uses this practice because 1) codified directives mandate rotation of candidate order on Election Day ballots and 2) research has shown that formulating research designs sensitive to the potential for response order effects improves pre-election measures of voter preferences (see, e.g. Rademacher and Smith, 2001; Visser et al. 2000; Miller and Krosnick 1998). In this research, we examine whether ballot order impacted voter preference distributions in telephone surveys conducted in a state widely portrayed as one of the key battlegrounds in the race for the presidency. During the 2004 election campaign the University of Cincinnati's Institute for Policy Research conducted telephone surveys designed to measure voter preferences in statewide races for president and U. S. Senate. These surveys included numerous ballot order experiments. While the 2004 race for the presidency in Ohio was highly competitive, and received a great deal of campaign and

media attention, the race for U.S. Senate was not competitive, and was relegated to the equivalent of a "down-ticket" race as a result. Analyses will examine the impact ballot order had on reported vote preferences in these two contests for president and U.S. Senate. In addition, we will also report the results of ballot order experiments in various types of races, including for president and U.S. Senate, conducted using the Ohio Poll in 2000 and 2002.

***Improving Data Quality: Do Changes in Item Formatting and Placement Help Older Respondents Provide Better Data?*, Judith Lynch, RTI International; jtl@rti.org, Anne Kenyon, RTI International; aek@rti.org, Scott Scheffler, RTI International; sscheffler@rti.org, Stephanie Rizk, RTI International; srizk@rti.org, Katherine Jackson, RTI International; kathjackson@rti.org, Jiantong Wang, RTI International; wang@rti.org, and Mildred Duke, RTI International; mcsov@rti.org**

The Medicare CAHPS Disenrollment Reasons Survey collects data from Medicare beneficiaries about the reasons they disenroll from their Medicare managed care health plan. The survey is administered quarterly to samples of approximately 12,000 to 24,000 cases each quarter. In the survey, beneficiaries are first asked a series of 34 closed-ended questions about reasons they disenrolled from their plan, after which they are asked to write in their single "most important" reason for leaving their plan. Text from the open-ended "most important reason" item is coded to general categories related to quality of care and costs and benefits. The "most important reason" item is one of the most critical analytic items in the questionnaire. The Centers for Medicare & Medicaid Services uses the survey results to help explicate disenrollment rates which it is required to provide to Medicare beneficiaries. Survey results are also provided to Medicare managed care organizations for use in their quality improvement initiatives. CMS also uses the results to monitor the performance of Medicare managed care organizations. In this paper, we examine whether formatting and question placement changes to the "most important reason" item improve the quality of the data collected. We compare data collected in survey quarters before and after formatting and placement changes were made. We examine rates of missing data, interactions between the "most important" reason item and the "other reasons" for leaving item, ability of older respondents to provide a single reason vs. multiple reasons, and the percent of open-ended entries that are uncodable. Our hypothesis is that making some targeted changes to the item will improve the quality of the data collected, leading to better information being available for beneficiaries making health plan choices.

***What is sexual harassment? It depends on who asks! Framing effects of sponsorship on survey responses*, Roger Tourangeau, Joint Program in Survey Methodology; rtourangeau@survey.umd.edu, Frederick Conrad, University of Michigan; fconrad@isr.umich.edu, Mick Couper, University of Michigan; mcouper@umich.edu, and Mirta Galesic, Joint Program in Survey Methodology; mgalesic@survey.umd.edu**

Numerous studies have demonstrated that the context of an item can influence the answers the respondents give. The context is typically defined quite narrowly, usually as the preceding few questions in the questionnaire. In this paper we are testing the hypothesis that other elements of the survey presentation, such as its sponsorship, can also define the context of the survey items and affect the answers obtained. The underlying mechanisms are essentially the same as those for the other context effects: the sponsorship can provide an interpretive framework for the questions; or it can trigger the recall of specific information useful in answering the questions. This means that asking the "same" question in two different questionnaires won't necessarily yield the same answers. We present the results of an experiment in which 2617 web respondents were randomly assigned to one of the two sponsorship conditions, and then asked a series of questions about sexual harassment at the workplace. One group was told that the sponsor of the survey was a feminist organization, Women Against Sexual Harassment; another that the sponsor was the neutral Work Environment Institute. The results show that the same workplace situations are interpreted as instances of sexual harassment more often when the sponsor is the feminist organization then when the sponsor is the neutral research institute.

ISSUES, CANDIDATES AND THE 2004 ELECTIONS

***The Separate Realities of Bush and Kerry Supporters*, Steven Kull, Program on International Policy Attitudes; skull@pipa.org**

The 2004 US presidential elections was striking in that supporters of President Bush had sharply different views of reality than supporters of Senator Kerry on a variety of issues. While large majorities of Bush supporters believed that before the war Iraq had weapons of mass destruction or a major program to develop them a large majority of Kerry supporters believed this was not the case. Large majorities of Bush supporters assumed that Iraq was providing substantial support to al Qaeda and even that evidence of such support had been found, while Kerry supporters believed the opposite. Bush supporters even assumed, incorrectly, that the Duelfer report concluded Iraq had WMD and that the 9/11 Commission concluded that Iraq had provided substantial support to al Qaeda. Bush and Kerry supporters differed sharply in their assumptions about world public opinion on the US going to war with Iraq. Bush supporters had incorrect assumptions about Bush's positions on a variety of issues including the Kyoto treaty, the comprehensive test ban treaty, and the land mines treaty. Even after Bush stated his opposition to the International Criminal Court in the debates a majority continued to assume that he favored it. The roots of these varying perceptions will be discussed drawing on the concepts of cognitive dissonance and the mirror image effect. The role of the media will be discussed drawing on a comparative analysis of the rate of misperceptions depending on respondents' primary source of news. It was found that respondents' perceptions varied sharply depending on their primary source of news: those who primarily received their news from Fox News were found to have the highest rate of misperceptions.

***Economic Voting and Political Sophistication in the U.S.: A Reassessment*, Jean-Francois Godbout, Northwestern University; godbout@northwestern.edu, and Eric Belanger, McGill University; eric.belanger@mcgill.ca**

Better informed citizens are much more likely to perceive the broader social implications of economic inequality (Bartels 2003). Since a certain amount of information is necessary for voters to form economic judgments, political sophistication appears to be a promising avenue of research for economic voting. Studies have already shown that individuals partially base their evaluations of incumbent governments on personal (the pocketbook hypothesis) and national economic conditions (the sociotropic hypothesis). It has also been demonstrated that these judgments are affected by past (retrospective hypothesis) and future (prospective hypothesis) economic evaluations. What we do not know is why some voters rely on one kind of economic perceptions instead of another. It seems intuitive to think that better informed citizens are more likely to recognize the implications of national economic policies on their own economic well-being and on the broader political context. However, scholars have yet to reach a consensus over the relationship between political sophistication and economic voting. We believe that this current debate is misleading since nearly all the differences in results are attributable to model misspecifications. In effect, part of the problem lies in the fact that studies on sophistication and economic voting

compare macro- and micro-level results and fail to include all relevant dimensions of economic evaluations in their models. Furthermore, the external validity of current studies is greatly limited since the typical analysis generally focuses on one or two elections. In the following, we propose to reassess the potential role of political sophistication on economic voting by pooling data from four National Election Studies (88-00), and by looking at all measures of economic perceptions. This should broaden our understanding of the relationship between political sophistication and economic voting and correct for some of the internal and external validity problems found in the literature.

***The Role of Issues in the 2004 Election*, Merrill Shanks, University of California, Berkeley; jms@csm.berkeley.edu, and Douglas Strand, University of California; dstrand@csm.berkeley.edu**
This paper will address several aspects of one general question: how should we explain the decisions of American voters in the 2004 Presidential election? Analyses will be based primarily on the Public Agendas and Civic Engagement Survey (PACES), but will include some comparisons with evidence from national exit polls. The 2004 PACES project is a continuing survey of the national electorate based on a comprehensive set of questions concerning policy-related controversies, and includes post-election re-interviews as well as initial interviews between February and election day. The paper begins by documenting the dominant role of voters' own views concerning the combination of many familiar policy-related controversies, and adresses the following questions: 1. In what way did the war in Iraq appear to dominate other issues in shaping voters choices between Bush and Kerry? Which specific issues concerning terrorism and the continuing conflict in Iraq appeared to play the most decisive roles in shaping those choices? 2. Which specific issues concerning traditional family values appeared to play the most important roles? Should the combined impact of those issues be primarily described in terms of controversies concerning gay marriage or civil unions, or was that impact spread over other "cultural" controversies - including abortion and the role of religion? 3. Which issues concerning economic policy issues played some visible role in shaping voters' choices - and which were eclipsed by other types of issues, such as Iraq or family values? 4. How divided is the US public? How large is that portion of the electorate that holds consistently conservative, or consistently liberal, views about most current policy-related conflicts -- and is therefore "unavailable" to one side in the election? 5. Finally, based on the above analyses, which policy-related controversies played the most important role in shaping George Bush's victory?

***What motivated Americans' views of the candidates and vote preferences across the 2004 presidential campaign?*, Neil Malhotra, Stanford University; neilm@stanford.edu, Jon Krosnick, Stanford University; krosnick@stanford.edu, Gary Langer, ABC News; gary.e.langer@abc.com, and Daniel Merkle, ABC News; dmmerkle@aol.com**
What motivated Americans' views of the candidates and vote preferences across the 2004 presidential campaign? Our paper will draw on a robust dataset - a dozen ABC News telephone polls conducted from January through October 2004, comprising more than 30,000 RDD interviews - to explore possible explanations for the public's candidate preferences. These surveys include measurements of George W. Bush's ratings for handling his job overall and specific issues, political party identification, ideology, prospective judgments about the candidates' performance on issues, perceptions of the candidates' personality characteristics, vote preferences and more. Also included are open- and closed-ended measures of the perceived weight the respondent would place on various issues when formulating vote choices. Our aim will be to discern what criteria Americans used to evaluate the candidates at various points in the campaign. We will build and estimate the parameters of logistic regression equations predicting candidate choice and compare the coefficients across these surveys. To test the validity of introspective measures of decision-making strategy, we will estimate regression equations to gauge the impact of various issues among people who indicate they will or will not place substantial weight on those issues. We will estimate the parameters of OLS regression equations predicting overall presidential job performance evaluations with domain-specific performance evaluations. And we will note whether the resulting coefficients vary across time in concert with the volume of national news media coverage of each issue, as gauged by story counts generated using Lexis-Nexis.

RACIAL ATITUDES

***The Interdependence of Determinants for the Strength and Direction of Social Desirability Bias in Racial Attitude Surveys*, Volker Stocké, University of Mannheim, SFB 504; vstocke@rumms.uni-mannheim.de**
Empirical evidence suggests that the respondents' approval motive, their desirability beliefs and the privacy of the response situation determine how prone survey answers are to social desirability bias. Previous research analyzed these factors separately and has not taken their possible interdependence into account. This paper examines the predictions from rational-choice theory that a strong approval motive, clear differences in the perceived desirability of response options and a lack of privacy are all necessary but not sufficient conditions for social desirability bias. This prediction was tested in our first study. With data from a local random sample, we found the beliefs about whether positive or negative racial attitudes are more desirable, and to what extent this is the case, to differ considerably between individual respondents and demographic groups. Respondents' racial attitude answers were in agreement with these beliefs and this significantly stronger, when the attitude responses were recorded interviewer- rather than self-administered. Furthermore, this association substantially increased when subjects had a stronger approval motive. This threeway interaction between the desirability beliefs, response privacy, and the strength of the approval motive was theoretically predicted. In a second study, we addressed the issue of whether the results found in the first study are externally valid. Since attitude answers and desirability beliefs were collected in the same interview, the observed associations may be an artifact due to the subjects' sensitization towards social desirability concerns. We thus collected in a separate study only racial attitude answers under conditions of varying response privacy. Aggregated differences in the desirability beliefs according to the attitude items, the respondents' social status and their education observed in the first study were matched with the response behavior under the same conditions in the second study. The results from the first study were replicated with this method.

***Blacks' Perceptions of their Racial Context and Prejudice Towards Whites*, David Wilson, Gallup/Michigan State University; david_wilson@gallup.com**
Much has been written of the need to study Blacks' racial attitudes, especially Blacks' prejudice. The impact of racial context, operationalized in terms of racial composition (i.e., racial group size) and racial threat, on racial attitudes is well established in studies of White Americans, yet the relationship is rarely studied for Black Americans. The present analysis centers on the effects of perceived racial context and threat, on self-reported prejudice. Using a nationally representative over-sample of Black Americans, collected by the Gallup Organization, I show how perceptions of racial group size, and perceived racial threat influence Blacks prejudice towards Whites. This

relationship is best explained by [sense of] group position theory, and related power-threat models. Generally, Blacks' prejudice towards whites is most influenced by perceptions that community Whites are prejudice. This threat is conditioned by Blacks' perceptions of a racially negative context, including perceptions of racial unfairness and discrimination, and limits on socio-economic opportunities. The implications focus on the larger context of improving race relations across growing minority populations, as well as gaining a general understanding of racial prejudice.

***When Crime Spending Views Reflect Punitiveness: The Context of Racial Prejudice*, Steven Barkan, University of Maine; steven_barkan@umit.maine.edu, and Steven F. Cohn, University of Maine; steve_cohn@umit.maine.edu**
The criminology literature includes many studies of the extent and correlates of punitiveness, or the belief that criminals should be punished more harshly. Two common measures of punitiveness include the General Social Survey's standard items on views about the death penalty and about the harshness of local courts. A growing number of studies link views on these two issues to racial prejudice among whites. Another standard GSS item, on "spending on halting the rising crime rate," has received virtually no attention in the criminology literature despite the great increase in criminal justice expenditures since the 1970s that is thought to reflect public preferences for such expenditures. On its face, this variable does not measure punitiveness as clearly as the other GSS items on the death penalty and courts' harshness. Moreover, this variable is only weakly correlated to the other two items at the bivariate level. However, punitive aspects of crime spending views emerge when they are studied among whites. Within this group, beliefs about blacks' proneness to violence predict support for greater crime spending, net of other variables. This association is specified when racial prejudice is taken into account, as it becomes especially strong for whites who score above the median on a racial prejudice scale reflecting aversion to contact and stereotyping, and it disappears for whites at or below the median. These results indicate that crime spending views among racially prejudiced whites reflect a desire that African Americans be punished by the criminal justice system.

Friday 10:00 am -11:30 am - AAPOR ABSTRACTS

DATA QUALITY I

***The Effects of Digital Recording of Telephone Interviews on Survey Data Quality*, Danna Basson, UW Survey Center; dbasson@ssc.wisc.edu**
Because digital recording is relatively unobtrusive, we assume it does not affect interviewer-respondent interaction and that recordings capture exactly how the interview would have proceeded otherwise. However interviewers may behave differently knowing they are recorded. Also, the technology is not completely unobtrusive and interviewers notice differences in the flow of the interview as recordings are saved on the computer. Finally, we also have to ask respondents for permission to record, which may affect their willingness to participate or their responses. The aim of this research is to examine how recording affects survey data quality. This paper analyzes Badger Poll data, a RDD poll of 500 Wisconsin residents conducted by the UW Survey Center in June 2004. Cases were randomly assigned to one of three recording conditions that were intended to isolate potential differences that are due to the recording equipment itself and to changes in interviewer behavior as a result of being recorded. The primary outcome measures examined here include response rates, break-offs during the interview, unit non-response, the time it takes an interviewer to convince a respondent to participate (after dialing the phone number), and the time to completion. Recording might affect interviewer behavior in conscious or unconscious ways, such as: reading questions more slowly, repeating answer categories in full rather than just in part, and tensing up or preparing before the interview begins. Thus, a subsample of interview recordings will be coded and analyzed for behavioral clues, such as interviewer speech rates, interviewer clarifications, and respondent hesitations before participation. Since recording can be used to improve the quality of surveys by monitoring interviewer performance and analyzing pretest interviews, the ease of digital recording makes the issue of how it affects the phone interview potentially relevant to more researchers.

***Use of Dependent Interviewing Procedures to Improve Data Quality in the Measurement of Change*, Jeffrey Moore, U.S. Census Bureau; jeffrey.c.moore@census.gov, Nancy Bates, U.S. Census Bureau; nancy.a.bates@census.gov, Joanne Pascale, U.S. Census Bureau; joanne.pascale@census.gov, and Julia Griffiths, U.S. Census Bureau; julia.e.klein.griffiths@census.gov**
Panel surveys often (always?) suffer from seam bias, the tendency for estimates of change measured across the "seam" between two successive survey administrations to far exceed change estimates measured within a single interview. Much research has documented the existence of seam bias (e.g., Moore and Kasprzyk, 1984; Burkhead and Coder, 1985; Hill, 1987; Martini, 1989; Young, 1989; Kalton and Miller, 1991; Brown, Hale, and Michaud, 1998); attempts to reduce it, however, have met with only limited practical success (Moore, Marquis, and Bogen, 1996; Brown, Hale, and Michaud, 1998; Mathiowetz and McGonagle, 2000; Rips, Conrad, and Fricker, 2003). The U.S. Census Bureau recently completed a multi-year research program to improve the questionnaire for the Survey of Income and Program Participation (SIPP). One important goal of this program was a significant reduction in seam bias; the questionnaire revision which was intended to accomplish this was a more extensive and more focused use of dependent interviewing procedures. Evaluations of these procedures in two field experiments found some positive results (Moore and Griffiths, 2003). The most recent SIPP panel, launched in February 2004, incorporates the new dependent interviewing procedures in the production SIPP questionnaire. This paper describes those procedures, and examines their impact on estimates of month-to-month change across waves 1 and 2 of the new panel for a number of characteristics (e.g., participation in government transfer programs, school enrollment, employment, and health insurance coverage) through a comparison with similar estimates derived from 2001 SIPP panel data. We find some evidence of a positive change with the new procedures; compared with 2001 panel estimates, estimates of month-to-month change at the interview seam from the first two interview waves of the 2004 panel are generally less out of line with off-seam estimates. Even with the improvement, however, much seam bias still remains.

***Reporting the frequency of household tasks by elderly respondents: the effect of different interview strategies on data quality*, Johannes van der Zouwen, Vrije Universiteit at Amsterdam; j.van.der.zouwen@fsw.vu.nl, Johannes Smit, Vrije Universiteit Medical Centre Amsterdam; jh.smit@mdw.vu.nl, and Marleen Van der Horst, Vrije Universiteit Medical Centre; MHL.vanderhorst@vumc.nl**
Detailed self-report questions are hard to answer, especially for elderly respondents (Schwarz et al., 1998). So one would expect that questions, posed to elderly respondents (65+) about the frequency and duration of their performance of household tasks during the last two weeks, lead to question-answer sequences resulting in much item non-response and many inadequate, irrelevant and incomplete answers.

Contrary to that expectation, in a survey after physical activities and fall incidents among elderly people, hardly any item non-response was reported. Moreover, the answers to these self-report questions corresponded quite well with independently collected information, by means of a 7-day diary, about the performance of various household tasks (Stel et al., 2004). In order to explain the difference between the expected and the actual interview process, we focused the analysis on the behavior of the interviewers. It turned out that they used different strategies to help the respondents to come up with relevant and complete answers. In the 'integral' strategy, the interviewer -in accordance with the questionnaire- poses questions about the frequency and duration of entire groups of household tasks. In the 'partial' strategy, the interviewer -deviating from the questionnaire- rephrases the questionnaire and poses the questions about frequency and duration for each individual task separately, therewith making the questions easier to answer. The 'mixed' strategy is a combination of these two interview strategies. The interview strategy employed has a strong effect on the duration of the question-answers sequences and on the information eventually provided by the respondent. References Schwarz, N., Park, D., Knaüper, B. and Sudman, S. (Eds.)(1998), Cognition, Aging, and Self-Reports. Psychology Press, Philadelphia, PA Stel, V., Smit, J.H., Pluijm, S.M.F., Visser, M., Deeg, D.J.H., and Lips, P. (2004), Comparison of the LASA Physical Activity Questionnaire with a 7-day diary and pedometer. Journal of Clinical Epidemiology 57, 252-258

***Managing Data Quality on the Survey of Consumer Finances 2004*, Leslie Athey, NORC; athey-leslie@norc.org, and Arthur Kennickell, Federal Reserve Board; Arthur.Kennickell@frb.gov**

There are many reasons why researchers would elect to gather factual information through personal interviews rather than by consulting records that can be difficult, or even impossible, to assemble and abstract. In choosing to gather facts directly from respondents, however, researchers must address potential sources of error in respondent reporting and recording of that information by interviewers. The Survey of Consumer Finances gathers detailed household financial information from a sample of 4500 Americans. This survey faces the challenge of collecting high-quality numerical data from a complex interview on a subject that is often poorly understood by respondents and untrained interviewers alike. For the 2004 round of this triennial survey, the Federal Reserve Board and NORC teamed to improve the quality of the financial information collected by interviewers. Attempted improvements included changes to interviewer recruitment, training, testing and retention strategies; and provision of both rapid turnaround and more in-depth feedback to interviewers about their data quality during data collection. We present both a description of the attempted improvements and an estimate of their impact on the quality of the financial information collected.

THEORIES AND MODELS OF NONRESPONSE

***How could they ever, ever persuade you? Are some refusals easier to convert than others?*, Nick Moon, NOP World; nmoon@nopworld.com, Nickie Rose, NOP World; nrose@nopworld.com, and Nikki Steel, NOP World; nsteel@nopworld.com**

With the increasing pressure on response rates, the reissuing of non-responding sample members - typically refusals and non-contacts - becomes more and more important if target response rates are to be achieved. Despite this, there is relatively little literature on the subject: a search of the POQ archive for "refuse convert" yields only the Singer, van Hoewyk and Maher 2000 article on incentives in telephone surveys. NOP is currently approaching the end of fieldwork on the second wave of a major face to face longitudinal survey, and we are engaging in a process of reissuing refusals and non-contacts. This process is normally managed by the Field department, but on this occasion the decision as to whether refusals should be reissued or not was made by researchers. The original plan was to send targeted letters, aiming to address the refusers' particular reason for refusing, and although in the end a standard letter was used for all, the refusals were all sorted by type - "too busy", "one wave was enough", "too personal" and so on. The reissue program is still underway, but when the program is finished we will be able to look at whether certain types of refusal were more likely to respond than others. The paper will thus compare conversion rates among different refusal types, and make suggestions as to strategies for reissuing in other surveys. Also, because the field period is long, some of the noncontacts will have been reissued some time after they were initially returned. One can theorise that the longer the gap before reissue the greater the chance of conversion, as it means that people who were away long term may have returned home. The paper will also address the extent to which this theory was borne out, and again suggest strategies for other surveys.

***A Theory-driven Approach to Reducing Nonparticipation of Twelfth Graders in National Surveys*, Young Chun, American Institutes for Research; ychun@air.org**

Understanding nonresponse bias has become more important in the National Assessment of Educational Progress (NAEP) of twelfth graders where participation rates have declined in recent years. NAEP, a national survey providing the so called "the Nation's Report Card," is the only nationally representative and continuing assessment survey of what America's students know and can do in various subject areas. The purpose of this paper is to explain nonparticipation of twelfth graders in NAEP by applying "social isolation theory" and using measurement and analysis model of nonresponse developed by Groves and Couper (1998) and further applied to education sruveys by Chun and Scott (2003). We have used data on 12,000 twelfth graders in the 2000 High School Transcript Study (HSTS), linked with NAEP surveys of students, teachers and principals all administered by the National Center for Education Statistics. The NAEP HSTS study is unique in providing us with measures of individual characteristics including achievement factors and absenteeism, and school-and-district-level indicators of social isolation for nonparticipants as well as participants in assessments. We dissect the effects of key variables on nonparticipation to explore level-specific impacts (i.e., student, school, and district), and then combine them simultaneously across levels to understand the effect of a full set of factors on participation. The findings address both theoretical implications for the external validity of a social isolation theory of nonparticipation and practical implications about measures of interventions to reduce nonparticipation in the national assessments of students. The knowledge we glean from this research would help develop effective interventions that control for student nonparticipation in NAEP or national assessment surveys.

***Participation Decisions Throughout and Across Surveys*, Andy Peytchev, University of Michigan; andrey@isr.umich.edu, Scott D. Crawford, Survey Sciences Group, LLC; scott.crawford@msiresearch.com, Karen Kurotsuchi Inkelas, University of Maryland; ki21@umail.umd.edu, and Sean McCabe, University of Michigan; plius@umich.edu**

Studies have found that survey break-off rates (defined as proportion of respondents who fail to complete a survey once it is started) have been as high as 50% in self-administered surveys. To the extent that those who break-off are systematically different from those who complete, bias may be introduced. In the spring of 2003, approximately 5,000 undergraduate students were invited to a university residential program survey, resulting in a 26% break-off rate (10% among invited). Later in the same semester a survey on student life issues included the same undergraduate students, resulting in a 7% break-off rate (4% among invited). Finally, a year later the first survey

was repeated on a mix of the same and new respondents, resulting in a 45% break-off rate (7% among invited). Respondents in the studies were linked such that we could track a single respondent's behavior across all three surveys, providing a rare look at within respondent patterns of participation. Basic demographic variables (e.g., race, gender) were made available for the full sample. The surveys were conducted by the same vendor and software but varied in the incentive, sponsor, topic, and format of questions. Half of the initial sample had participated in a program related to the residential topic. These differences in design allow comparison of reasons for break-off over the course of the surveys and of the subsequent effect on nonresponse bias. We will discuss how respondent, survey, and question characteristics jointly lead to break-offs and what conditions produce more potential for bias. The nature of the link between break-offs and unit nonresponse in the same survey will be examined. We will look at how break-off location within the questionnaire could be predictive of different response propensities to subsequent survey requests. Implications, such as using reasons for break-off in weighting will be considered.

***Exchange and Compliance: Integrating Theories of Survey Nonresponse*, John Goyder, University of Waterloo, Dept. of Sociology; jgoyder@uwaterloo.ca, Luc Boyer, University of Waterloo; lucboyer@sympatico.ca, and Guil Martinelli, University of Waterloo; guil_martinelli@yahoo.com**

Interest in theorizing about survey nonresponse has accelerated over the past decade. Earlier in the history of survey research, response decisions often seemed to be treated as unfathomable black box phenomena. Scholars such as Dillman, Groves and Couper, however, have drawn from across the social sciences to extract conceptual schemes for why people decide to accept or decline requests for surveys. With the conceptual enrichment, however, has come complexity and some confusion. Beginning with analysis of current usage in articles on survey nonresponse in leading journals, the argument behind this paper is that the relationship between social exchange theory and the theory of social psychological heuristics has not heretofore been systematically addressed. This we endeavour to do, using a two-dimensional scheme. Building on work by Groves and Couper, we define one dimension as "amount of decision-making." This helps locate the heuristic interpretation of response decisions. The second dimension is termed "strength of social factors " and helps to locate the different varieties of social exchange theory that pertain to response behavior. Viewed thus, it becomes clear that one source of confusion in theorizing about survey nonresponse is the heterogeneity of the various approaches which gather under the umbrella term "social exchange." At the most general level, social exchange is the background landscape behind most theorizing about survey nonresponse, but once defined more specifically a fit into one or another quadrant within our scheme becomes possible. The scheme proposed is sufficiently abstract that alternative conceptual approaches, differentiated both from social exchange and from heuristics, can be included. It is proposed, for example, that communications theory around how people handle information overload is relevant to response decisions and also that ideas from psychotherapy theory help define the successful survey interviewer.

NONRESPONSE BIAS

***Potential Non-Response Bias from Hard to Reach Respondents*, Peyton Craighill, Pew Research Center for the People and the Press; craighillp@people-press.org, Cary Funk, Pew Research Center; cfunk@pewresearch.org, and Michael Dimock, Pew Research Center for the People and the Press; dimockm@people-press.org**

Throughout 2004, the Pew Research Center has compiled data on how difficult or easy it was to complete each interview in our surveys toward the purpose of better understanding potentially subtle differences in the political attitudes, behaviors, and characteristics of respondents who are relatively easy to reach in a standard national RDD survey and those who are more difficult to reach. This data collection effort is a complement to Pew's 2003 Methodology Study presented at AAPOR 2004, in which a standard 5-night survey was compared with an extended "rigorous" study in which every effort was made (including mailings and incentives) to maximize response rates over a three month period. While the comparison of these "hardest to reach" respondents with the standard sample found few substantive differences in political attitudes or other personal information, the limitations on sample sizes left some ambiguities. By combining data from roughly 20,000 interviews, we increase our sample size and have the ability to pinpoint small but relevant differences between easy and hard to reach respondents (though we still must rely on the 2003 Methodology Study to understand those who are never captured in a standard survey protocol). The research design looks for disparities between interviews completed with little effort (few calls, no refusal conversions) and those which required extensive callbacks and/or refusal conversion efforts in the five day period. The subjects for comparison include a broad gamut of political and social attitudes, including electoral preferences, intention to vote, support and opposition to Iraq, gay marriage, foreign policy attitudes, media sources and news interests, as well as a full battery of demographic characteristics. The implications of the results are particularly important for researchers conducting one- and two-night public opinion surveys, in which only a limited effort is possible to reach all potentially accessible respondents.

***Are We Under-Counting Blue Counties?: Correcting for Disproportional Response Rates at the County Level in National RDD Samples*, Michael Dimock, Pew Research Center for the People and the Press; dimockm@people-press.org, and Jonathan Best, Princeton Survey Research Associates International; Jonathan.Best@psra.com**

Analysis of national RDD samples conducted by the Pew Research Center and Princeton Survey Research Associates International finds a systematic and consistent underrepresentation of respondents in counties with the highest population density. The underrepresentation results from significantly lower cooperation rates in these counties, and leads to samples that, even when corrected for age, gender, race, ethnicity, education and census region, fail to reflect the true distribution of the public across all parts of the country. Because high population-density counties voted heavily for John Kerry, this geographic disproportionality, if uncorrected, would result in an overestimation of Bush's margin over Kerry by roughly two percentage points in pre-election polls conducted in the 2004 election cycle. Moreover, evidence from the 2003 Pew Research Center Methodology Study, which was in the field for over three months and made every effort to complete interviews in even the hardest to reach and least cooperative households found that the disparity in response rates cannot be overcome by greater effort in the phoneroom.

***Unit Nonresponse and Error in a National Public Opinion Survey: A Census Matching Approach*, Allyson Holbrook, University of Illinois at Chicago; allyson@uic.edu, Timothy Johnson, Survey Research Laboratory; timj@uic.edu, Young Cho, University of Illinois at Chicago; youngc@srl.uic.edu, and Jaime Brugueras, University of Illinois at Chicago; jaimeb@srl.uic.edu**

Response rates to social and behavioral surveys have been declining for several decades, increasing the likelihood that differences between respondents and nonrespondents may bias survey estimates. Response rates are impacted by interviewers' ability to contact respondents and to gain their cooperation, and the association of contact and cooperation with important respondent characteristics may vary for face-to-

face and telephone surveys. Records match analyses that compare the personal and/or community characteristics of households that do and do not respond to survey requests have been employed to investigate the potential effects of unit nonresponse on survey estimates. One key advantage of this approach is that information regarding all sampled households is available for analysis. We propose a unique application of this methodology to compare nonresponse in face-to-face and telephone surveys. Specifically, we will compare nonresponse bias in the RDD telephone and area probability face-to-face surveys conducted as part of the 2000 National Election Studies Survey. This approach has not been previously employed to compare nonresponse in telephone and face-to-face cross-sectional national public opinion surveys. In the current research, indicators of survey response, refusal, and noncontact will be constructed using final sample dispositions. Data from completed surveys and census data on demographic variables related to socioeconomic status and racial composition at the zip code level will then be merged with the sample frame using the telephone exchange/area code combination for the telephone survey sample and the zip-code for the face-to-face survey sample. These data will allow us to examine the associations between each demographic variable and household nonresponse, survey refusal, and survey non-contact. Because the survey being analyzed is nationwide in scope, and provides random samples of both telephone and face-to-face interviews, these data will also allow us to develop and contrast models that describe the nonresponse mechanisms associated with telephone vs. face-to-face surveys.

***Non-response Bias in the Measurement of Radio Listening: Is it Necessary to Make 13 Contact Attempts?*, Robin Gentry, Arbitron Inc.; robin.gentry@arbitron.com**
Response rates have become an increased focus of survey researchers in the past decade due to the escalating effort and resources required to maintain traditional levels of respondent cooperation. In addition, potential non-response bias has led to a marked rise in the number of call attempts researchers are making in hopes of contacting respondents. The goal of additional attempts is to bring hard-to-reach respondents into the sample -under the assumption that they are significantly different from easy-to-reach respondents-thereby reducing non-response bias. A number of recent studies (e.g. Curtin et. al, 2000; Keeter et. al., 2000) suggest that simply adding contact attempts may have little effect on decreasing non-response bias in the estimates. What is unknown is whether this trend holds across variables and across surveys. For example, a number of respondent characteristics correlated with radio listening behavior (i.e., age, race/ethnicity, employment status, and income) have traditionally been associated with differential non-response-possibly leading to a higher risk of non-response bias than in other types of surveys. The current analysis of over 100,000 households who participated in a large scale syndicated radio ratings survey explores potential non-response bias if fewer call attempts had been made to the household. The analysis consists of a simulation, similar to that used by Curtin et al. (2000), whereby the database is randomly split and the typical 13-call design is compared to what would have occurred if fewer calls had been placed. A measure of bias was calculated for each pair of simulated call-designs (i.e., 1 call vs. 13 calls, 2 calls vs. 13 calls, etc.) Results suggest that fewer call attempts could be utilized without adding substantial bias to the key estimates (e.g., total quarter hours of listening, number of stations, and favorite station format).

***Nonresponse Bias in Reliability and Validity Estimares for Attitude Items*, Zachary Arens, Gallup; zac_arens@gallup.com**

RESPONSE SCALES

***Rating versus Comparative Trade-off Measures: Effects of Task, Topic, Element Differentiation, and Number of Elements on Validity*, Randall Thomas, Harris Interactive; rthomas@harrisinteractive.com, Kerri Miller, Harris Interactive; kmiller@harrisinteractive.com, and Alyssa Johnson, Harris Interactive; amjohnson@harrisinteractive.com**
While we have found that comparative trade-off tasks had inferior validity compared to rating tasks, our prior studies had respondents evaluate elements that were highly differentiated in terms of importance or favorability. However, comparative tasks have been proposed to function better when the elements to be evaluated are harder to differentiate in terms of importance or favorability. Based on the results from a pilot study, we created element sets that were either clearly differentiated from high to low in importance and positivity or were less differentiated (all more desirable/important). Respondents were then randomly assigned to 1 of 6 evaluative tasks (rating, 5 comparative tasks - full ranking of elements, top 3 ranking, top-bottom ranking, constant sum, and paired comparison). We replicated and extended many of our earlier results. Some notable differences between the tasks emerged when the elements were less clearly differentiated.

***How much are you willing to give? The response scales can influence the size of donations*, Mirta Galesic, Joint Program in Survey Methodology; mgalesic@survey.umd.edu, Frederick Conrad, University of Michigan; fconrad@isr.umich.edu, Mick Couper, University of Michigan; mcouper@umich.edu, and Roger Tourangeau, Joint Program in Survey Methodology; rtourangeau@survey.umd.edu**
The candidates in the last presidential race relied substantially on donations collected through their web sites. Many other organizations use the web for collecting contributions from their supporters. These web sites differ in the way they formulate the donation request. Most of them offer some kind of response scale with various amounts of donations as the possible answers. The number and range of the scale points, i.e. possible dollar amounts, as well as the intervals between them, vary widely from site to site. It is well known that response scales can influence respondent's answers. For example, the respondents can interpret the scale midpoint as the answer of an average person. Is it possible that different response scales can also change the amount people donate to an organization? In this paper we test the hypothesis that the midpoint of the scale of proposed dollar amounts can influence the amount of money that people are willing to give. We conducted an experiment in which 2144 web respondents were randomly divided in four groups. Each group was given the same description of a fictitious medical fund that would be used for financing important medical research, and asked how much money they would be willing to donate. Three of the four groups were given a response scale with a fixed number (9) and range (from $5 to $1000) of scale points, but with different amounts at the visual midpoints: $50, $250, and $500. The fourth group was given a numerical field and asked to input an amount themselves. Of the three groups that received the response scales, the one with the highest midpoint was willing to donate the most money. This amount was similar to the average amount given by the fourth group.

***Response Option Ordering: Reconciling meanings conveyed by rating scale label and position*, Philip Garland, Stanford University; garland@stanford.edu, and Jon Krosnick, Stanford University; krosnick@stanford.edu**
In questionnaires that present fully-verbally-labeled rating scales visually to measure attitudes, response choices are usually shown in ascending or descending order from most positive or negative to least positive or negative. This research explores whether and how departures from this convention might affect respondent behavior. We manipulated the wording of answer choices and the order in which they were presented, and measured answer choices, time to complete the questionnaire, and respondent satisfaction with the survey

experience. Respondents in a large national survey were randomly assigned to one of three experimental conditions: the "standard order" (SO), "moderately unsystematic order" (MUO), and "random order" (RO) conditions. In the SO condition, rating scale response options were presented in a conventional order from most positive to most negative. In the MUO condition, the response options appeared in that same order except for the reversal of the ordering of the two response options at either the beginning or the end of each scale. In the RO condition, the same response options appeared, but in random orders. If respondents rely only on the denotative meanings of words in the response options, then the frequency with which people chose each option should not vary depending on presentation order. But if people quickly perceive each rating scale as representing a continuum, respondents in the MUO condition may ignore the small denotative meaning differences between the reversed adjacent scale point labels and place themselves along the continuum at the same places as they would have in the SO condition. And if people pay at least some attention to the denotative meanings of the verbal labels, then they should quickly detect the random order and choose the labels in the RO condition with the same frequency as they do in the SO condition, ignoring the order of presentation.

***Comparing Check-All and Forced-Choice Question Formats in Web Surveys: The Role of Satisficing, Depth of Processing, and Acquiescence in Explaining Differences*, Jolene Smyth, Washington State University; jsmyth@wsu.edu, Don Dillman, Washington State University; dillman@wsu.edu, Leah Christian, Washington State University; lmchristian@wsu.edu, and Michael Stern, Washington State University; mstern@wsu.edu** It is common survey practice to convert a series of yes/no (forced-choice) items in telephone surveys to check-all-that-apply items in web and mail surveys. However, relatively little is known about how these different question formats may influence answers. This paper reports results from two web experiments and a comparison paper experiment in which nine different questions, varying in substantive topic and type (opinion-based and fact/behavior-based) were asked in 16 experimental comparisons of the check-all and forced-choice formats. Our purpose was to determine whether this change in question format influenced the number of response options marked affirmatively within each question and why any differences might occur. Results revealed that in every instance respondents marked significantly more items in the forced-choice format than in the check-all format. Given these results, detailed analyses of response patterns within questions, answering time, and alternative wording structures of questions were undertaken to examine which of three theories (satisficing, depth of processing, and acquiescence) best accounted for the response differences across question formats. These analyses indicated that the forced-choice format appears to invoke deeper processing and to eliminate satisficing behavior that occurs among some respondents to the check-all format, but that acquiescence does not seem to be an issue in the forced-choice format. Thus, it appears that the use of the forced-choice question format is a desirable alternative to the use of the check-all question format for multiple answer questions. In addition, the findings reported here give ample reason to be concerned about the current practice of automatically converting items from the forced-choice format to the check-all format or vise versa when switching between telephone and paper or web surveys.

HEALTH SURVEYS I

***NHANES Converted Refusals: Are they different from willing respondents in socio-demographic composition?*, Yinong Chong, Centers for Disease Control and Prevention; ychong@cdc.gov, Margaret Carroll, CDC; mcarroll@cdc.gov, and Pat Montalvan, Westat; Pat Montalvan MONTALP1@WESTAT.com** The National Health and Nutrition Examination Survey (NHANES) is the only national survey collecting data through both in-person interviews and medical examinations. Over the years NHANES has maintained interview response rates over 80 percent and medical examination response rates of 75 percent or better, due to extensive interviewer efforts to obtain cooperation. This paper examines whether converted refusals differ from willing respondents in socio-demographic characteristics -- important correlates of many health conditions in the U.S. population. The study population included adults 20 years and older who had completed both interviews and medical examinations (n=10267). Administrative data from 1999-2002 were used to identify respondents requiring refusal conversion for the household screening/health interview or for the medical exam. Demographic attributes for these two groups were compared with those of willing respondents. The data indicates that respondents who initially refused to participate in the medical exam tended to be younger than willing respondents (40 versus 46 years old), and were more likely to be male (53 versus 48%) and African American (18 versus 11%). However, those who were converted at the time of the screening or health interview and consequently participated in the medical exam showed no significant difference in mean age, gender or racial/ethnic composition from willing respondents. Furthermore, respondents initially refusing the medical exam also had lower education and income levels, whereas those initially refusing the screening or health interview did not show such disadvantages. These findings indicate that the group initially refusing the medical examination was somewhat different from willing respondents in socio-demographic attributes. Using the NHANES examination data, future studies should investigate whether the socio-demographic differences in these two groups translate into differences in health profiles and conditions.

***Measuring Knowledge of General Cancer Risks - Open Listing versus Targeted Questioning in the 2002 HINTS*, Michael Massagli, Dana-Farber Cancer Institute; mikemassagli@comcast.net, Vish Viswanath, Harvard School of Public Health/Dana Farber Cancer Institute; Vish_Viswanath@dfci.harvard.edu, Bradford W. Hesse, National Center Institute and Richard P. Moser, National Cancer Institute** The National Cancer Institute's (NCI) Health Information National Trends Survey (HINTS) gathers information from a nationally representative sample (with over-samples of ethnic-minority populations) about cancer communication practices, information preferences, risk behaviors, attitudes, and cancer knowledge. HINTS data may guide NCI's information planning efforts and development of effective health communication strategies. It is paramount that valid and reliable measurement of cancer knowledge be demonstrated. Our analysis points to difficulties in current measurement of knowledge of general cancer risk in the HINTS. We contrast estimates of knowledge of specific risk factors obtained using open-listing and targeted questioning for population subgroups defined by race, language of interview, sex, age, education, and cancer experience. For example, all respondents were asked: "Can you think of anything people can do to reduce their chances of getting cancer? Anything else?" 64% of non-Hispanic Whites (NHW) mentioned not smoking or quitting smoking, compared with 50% of non-Hispanic Blacks (NHB) and 49% of Hispanics. A random-half of respondents were also asked: "I'm going to read you some things that may affect a person's chances of getting cancer. Do you think that smoking increases a person's chances of getting cancer a lot, a little or not at all or do you have no opinion? 94% of NHW responded a lot or a little, compared with 91% of NHB and 94% of Hispanics. Additionally, estimates deriving from targeted questioning may be affected by priming. Among NHW who did not mention smoking in response to the open-listing question, 86% responded a lot or a little to the targeted question, compared with 98% of those who did mention smoking. For NHB, the contrast is 86% versus 96%; for Hispanics the contrast is 92% versus 95%. We present additional examples and discuss implications for development of items to measure knowledge of cancer risks.

***Impact of Reports of Hepatitis B Vaccination on Hepatitis A Vaccination Reports*, Lisa Carley-Baxter, RTI International; lcbaxter@rti.org, Doug Passaro, University of Illinois at Chicago; doug@uic.edu, Paul Levy, RTI International; levy@rti.org, Susan Twiddy, RTI International; twiddy@rti.org, and Ron Hershow, University of Illnois at Chicago; RCHersho@uic.edu**

Hepatitis A and Hepatitis B are both viral diseases affecting the liver, however they are transmitted via different means and often result in different spectra of health effects once infected. Vaccinations are available for each of these diseases. It is hypothesized that many lay people are not familiar with differences between Hepatitis A and Hepatitis B and may confuse them due to the similarity of their names. For this paper, we report the results of an experiment that was conducted in the fall of 2004 embedded within a retrospective study of childhood vaccination to determine whether reports of Hepatitis A vaccination were influenced by the inclusion of questions about Hepatitis B vaccination. This random-digit-dial survey was conducted with parents of children between 2.5 and 15 years of age. Parents were asked for reports of childhood vaccinations, permission to contact medical care providers to obtain data on these vaccinations, and the demographic characteristics of the household. Interviews were completed with 650 households, representing approximately 1,200 children, in Arizona and Oregon. These states were chosen because of their high incidence rates of hepatitis A. For those cases where parental permission was obtained, medical care providers were contacted to obtain vaccination records for the children. We further investigate whether these vaccination reports differ by demographic characteristics, the number and age of children, presence of shot records, and the number of vaccination providers. We extend this analysis by also investigating the accuracy of the reports of Hepatitis A vaccinations from parents who were also asked the Hepatitis B questions. Finally, we compare parent and provider reports of Hepatitis A vaccinations.

***A balance between quality and cost- conducting an Adult Tobacco Survey with multiple stakeholder interests.*, Barbara Fernandez, ORC Macro; barbara.m.fernandez@orcmacro.com, Randal ZuWallack, ORC Macro; randal.s.zuwallack@orcmacro.com, Kristie Hannah, ORC Macro; kristie.m.hannah@orcmacro.com, Anne Gorrigan, ORC Macro; Anne.M.Gorrigan@orcmacro.com, Kisha Bailly, ORC Macro; Kisha.M.Bailly@orcmacro.com, and Peter Mariolis, CDC; Pmariolis@cdc.gov**

Abstract Problem/Objective: Federal, state, and local organizations have gradually joined forces to cultivate a standardized survey and study protocol to measure adult tobacco use. Conducting a state-based Adult Tobacco Survey (ATS) involves multiple stakeholders with competing interests -- in addition to the common goal of collecting high-quality data. The state or organization wants the study conducted in a timely and cost-effective manner; the CDC wants comparability between state surveys; multiple Independent Review Boards (IRBs) must assure that subjects are not harmed; and the vendor wants the study to run smoothly and efficiently. Ensuring that all demands are met, while allowing for the highest data quality so that the findings are disseminated with confidence, is often challenging. Standardized calling protocols and survey design minimize sampling and non-sampling survey error within reasonable cost limits. Methods: Data from states that have and have not used the CDC's ATS Protocol Guidelines will be analyzed. Key smoking indicators will be examined in light of guideline adherence and transgressions. Research will include comparisons of smoking estimates and the demographics of participants based on: the number of call attempts; refusal conversion procedures; the wording of the introduction; and whether or not introductory screening procedures are in place to identify subgroups such as smokers. Results: (Results will be available in January) Conclusions: Adhering to the Guidelines facilitates the process and ensures that the data is collected via standardized procedures proven to result in high data quality in that they allow for: 1) confidence in results as preliminary research has shown that the recommended protocols reduce the risk of potential bias; 2) comparability between surveys; and 3) the survey to be fielded in a timely and cost-effective manner.

SAMPLING II

***Improving web based intercept surveys. A framework for the Active User Sampling.*, Lars Kaczmirek, ZUMA, Centre for Survey Research and Methodology; kaczmirek@zuma-mannheim.de, and Wolfgang Neubarth, ZUMA, Center for Survey Research and Methodology; neubarth@zuma-mannheim.de**

The implementation of popups in intercept surveys is a commonly used sampling method for data gathering purposes. As advertisements are using the same technique for distribution, popups are becoming increasingly unpopular. Due to the development and increasing use of popup blockers in the most common browsers and in security software like firewalls, the possibility of reaching participants is decreasing. Similarly, there is a tendency for internet users to learn to ignore unexpected popups. To overcome these problems and to call the user's attention to possible participation we propose to intercept the user in his or her navigation behaviour. Clicking on a link leads to an interception page where the user is informed about the survey. Either an action of the user or a timeout leads to the previous destination page or the survey. After completing or abandoning the survey, the user is then directed to the previous target page. Avoiding popups and implementing this active user sampling method leads to several advantages: (1) Only users who actively navigate through the promoting website are invited, (2) the invitation may be build in accordance with the links and topics users choose, (3) users are more prone to actively react to the interception page. The paper integrates and explores the use of dynamic pages, javascript, cookies and other session management and user identification tools into a framework for an active user sampling method. A case study conducted to evaluate an information portal demonstrates the information flow needed for implementation. Fallback methods will be discussed to allow for the variety of browsers and system configurations to fit into the framework as well as to maintain the user experience of the website.

***Simultaneous Sampling*, Karol Krotki, RTI International; kkrotki@rti.org, and Cynthia Bland, RTI; cbland@rti.org**

Simultaneous Sampling This research addresses situations in which it is possible and desirable to draw two or more independent samples from the same frame, or from two or more similar frames. The paper begins with a discussion and comparison of methods that can be used to select one sample followed by a second sample conditional on the first. Several issues need to be addressed, the primary one being the correct calculation of sampling probabilities. The main part of the paper addresses the challenge of selecting two simultaneous samples in such a way as to reduce overall field costs compared to the cost of implementing the samples independently. The goal is to select the two samples in pairs or cluster such that, one, each cluster consists of units from the two samples and, two, the units are geographically compact. Several methods are described. One group of methods involves a multistage selection of units starting with appropriately-sized primary sampling units (PSUs). Another general approach is to resort to list sampling and arranging dynamically for the geographic linkage of units from the two frames. Specific application of the methods will be discussed in the context of the 2006 PISA/PIRLS surveys.

***Comparison of Traditional Listings and USPS Address Database as a Frame for National Area Probability Samples*, Colm OMuircheartaigh, NORC, Univeristy of Chicago; colm@uchicago.edu, Stephanie Eckman, NORC, University of Chicago; eckman-stephanie@norc.uchicago.edu, Ned English, NORC, University of Chicago; english-ned@norc.net, James Lepkowski, ISR, University of Michagan; lepkowskij@norc.uchicago.edu], and Steven Heeringa, ISR, University of Michagan; herringas@norc.uchicago.edu]**

The paper will present empirical findings from a comparison of two major frames for national area probability sampling. Since 2000, survey samplers have been exploring the potential of the United States Postal Service (USPS) address lists to serve as a sampling frame for probability samples from the general population. Though the early work has demonstrated the strengths of the USPS address lists, there has not been a comparison of the two methods on a national scale. NORC and ISR are now working together to compare two national area-probability sampling frames for household surveys: (i) the frame produced by traditional listing, using survey field staff, and (ii) the list of postal addresses compiled by the USPS. We are conducting this comparison in an ongoing survey operation which combines the next wave of the Health and Retirement Survey (HRS) with the first wave of the National Social Life, Health, and Aging Project (NSHAP). We will present preliminary results on the relative coverage properties of the two frames, and describe challenges in carrying out the comparison. The research provides insight into the coverage and cost/benefit trade-offs that researchers can expect from traditionally listed frames and USPS address databases.

***A new and better method for sample designs with disproprtionate sampling*, Iain Noble, Department for Education and Skills; iain.noble@dfes.gsi.gov.uk, Peter Lynn, ISER, Essex University; p.lynn@essex.ac.uk, and Patten Smith, BMRB International; patten.smith@bmrb.co.uk**

Survey designers often want to oversample specific sub-populations to improve precision and robustness. Nearly all existing methods in two stage sampling produce explicit problems, with specific weaknesses, because domains (sub-populations) targeted for over sampling usually cross-cut the clusters that form the PSUs. This paper sets out a new method of sampling giving major gains over previous methods, in both effectiveness and efficiency, by simultaneously constraining three factors: - sample sizes within domains that cross-cut clusters; - sample sizes within clusters; - variation in selection probabilities, both within domains and overall. This is achieved through use of a size measure which, instead of being a simple count of the second stage units, is a weighted sum of the counts of units in each domain within the cluster. This requires knowledge of the distribution of units over domains within each cluster, which is often possible when administrative data are available. These data can be exploited to produce samples with closely controlled composition of PSUs, that are more effective in boosting sub-samples than standard methods and have, through reduced design effects and increased effective sample size, greater cost efficiency than otherwise possible. The paper comprises a description of the method, its underlying theoretical framework and a case study of its application in the design and implementation of the sample for a major longitudinal study of young people. The overall design was a conventional two-stage one with PPS, with schools being PSUs and individuals sampled from their registers. A critical objective was, however, boosting sample numbers for six ethnic minority groups. The paper sets out the advantages of the new method in this case, problems encountered and their solutions. The conclusion is that this method has substantial advantages over others and should be the default method for use in similar circumstances.

***Assessing the Effectiveness of an Optimal-Allocation Sample Design to Obtain Reliable Estimates of Health Insurance among Minority Populations*, Tom Duffy, ORC Macro; thomas.p.duffy.jr@orcmacro.com, and Ronaldo Iachan, ORC Macro; ronaldo.iachan@orcmacro.com**

The Ohio Family Health Survey (FHS) is a telephone survey of the health and health insurance status of adults and children in Ohio. The FHS prescribed confidence intervals for estimates of insurance status for several population subgroups: rural regions, ethnic minorities, families in poverty, families with children, etc. These constraints required a complex sample design that over-sampled on many levels. To obtain reliable estimates of the health insurance status of African Americans, especially those with lower incomes, an optimal allocation design was employed. This allocation involved stratifying Ohio's largest metropolitan counties into High, Medium, and Low density strata based on Census estimates of the proportion of households with at least one African American adult. Preliminary estimates of the costs of screening households, and the costs of completing interviews, were entered into the allocation model. The resulting disproportionately-stratified design attempted to optimally allocate interviews across strata with minimal impact on the variance of survey estimates. This paper will review the optimal allocation design, compare preliminary estimates of demographics to what was obtained during fielding, and examine the efficiency of the design with respect to the precision of key survey estimates as well as the costs of screening households.

RECENT METHODOLOGICAL RESEARCH ON THE KNOWLEDGE NETWORKS APPROACH

***Comparing the Knowledge Networks Web-Enabled Panel and the In-Person 2002 General Social Survey: Experiments with Mode, Format, and Question Wordings*, Tom W. Smith, NORC, University of Chicago; smitht@norcmail.uchicago.edu, Li Rick, Knowledge Networks; rli@knowledgenetworks.com, and Mike Dennis, Knowledge Networks; mdennis@knowledgenetworks.com**

This paper is an examination of the interplay of mode of data collection and item nonresponse caused by 'Don't Know' responses. A series of experiments were carried out to replicate and extend earlier comparisons between an in-person survey (the General Social Survey) and Web-enabled Knowledge Networks surveys. Results indicate that levels of Don't Know on e-survey are highly contingent on format and conditioning effects and that a Web-based format for Don't Knows can be designed to match DK levels found in in-person surveys. Moreover, differences in distributions across modes are notably reduced with DK responses are excluded, but do not disappear. Support for public spending was generally greater on the in-person GSS than on the KN e-surveys and the differences were notable on items concerning the underclass. Finally, wording experiments produced results that were consistent in direction across modes, but not always comparable in magnitude. Thus, while e-surveys produce similar findings to in-person surveys in many circumstances, notable differences occur in other situations.

***Data Collection Mode Effects Controlling for Sample Origins in a Panel Survey: Telephone versus Internet*, Cindy Chatt, Gallup; cchatt27@msn.com, Mike Dennis, Knowledge Networks; mdennis@knowledgenetworks.com, Li Rick, Knowledge Networks; rli@knowledgenetworks.com, and Paul Pulliam, RTI; pulliam@rti.org**

We evaluate telephone and Internet-based modes of survey data collection by controlling for sample origin. Previous research has focused on sample effects only. Our main result is that substantive response differences are primarily associated with mode of data collection and not with sample origin. Sample origin is controlled by conducting both Internet and telephone interviews with members of the Knowledge

Networks (KN) web-enabled panel. The survey, which was sponsored by RTI International, measures policy and civic attitudes regarding 9/11 in early 2002, and was designed by RTI International and the Odum Institute at the University of North Carolina. The survey analysis is based on 2,979 web interviews with KN panelists, 300 telephone interviews with KN panelists, and 600 telephone interviews with persons that refused to join the KN panel or else take the web panel survey. The differences caused by mode in this Internet versus telephone study were strikingly similar to the telephone versus mail mode effects found in civic attitude studies by Tarnai and Dillman and in telephone versus face-to-face mode effects by Krysan. These studies found a tendency (which we confirm) for telephone respondents to answer on the extreme positive end of the scale. The Internet respondents are more likely than both telephone sample groups to use the full range of scales.

***Statistical Tests of Data Quality in a Contingent Valuation Survey Using Knowledge Networks Data*, Joel Huber, Duke University; joel.huber@duke.edu, Jason Bell, Duke University; jason.bell@duke.edu, and W. Kip Viscusi, Harvard University Law School; kip@law.harvard.edu**

This paper uses statistical tests on Knowledge Networks data to determine the extent of survey bias in a contingent valuation study about water quality with randomization of survey versions. The validity tests include examining the data against theoretical predictions and an across-person test requiring respondents to be sensitive to the scope of differences in cost of living and water quality. There are also validity tests determining whether panel membership influenced the valuation results. There are four variables in the regression of the determinants of the value of water quality benefits. The first variable is whether the respondent stopped and then continued the survey at a later time. The second variable is the time the respondent has been a member of the Knowledge Network panel. The third variable is the number of days the respondent took to complete the survey. The final survey methodology variable tested is whether the respondent subsequently quit the panel at any time until May 2004. Although the sample is nationally representative it is useful to test for possible selection biases arising from panel members who were invited to participate but did not successfully complete the survey, and these results are also included. Overall, there is no indication that any of these key aspects of the panel methodology bias the survey responses. There were no significant effects of any of the Knowledge Networks panel variables so that there is no evidence that national performance of the survey task is importantly influenced by any of these variables.

***: Correcting for Self-Selection Bias Using the Heckman Selection Correction in a Valuation Survey Using Knowledge Networks*, Trudy Cameron, University of Oregon; cameron@darkwing.uoregon.edu, George M (J.R.) Deshazo, UCLA; deshazo@ucla.edu, and Mike Dennis, Knowledge Networks; mdennis@knowledgenetworks.com**

Researchers frequently acknowledge several reasons for possible non-representativeness in surveys using panel samples. We model the selection process for the Knowledge Networks panel, starting with a random-digit-dialed set of initial contacts and following these cases through a number of distinct attrition opportunities, ending with one sample drawn for an actual survey and the individuals who chose to respond to it. Using GIS methods, we match over 525,000 RDD addresses or telephone exchanges to the corresponding county and the most appropriate census tract. We use a set of fifteen orthogonal factors based on census tract characteristics, plus county voting percentages for candidates Gore and Nader in the 2000 Presidential election. We find many statistically significant determinants of attrition at our different attrition opportunities. To illustrate the effects of selection, we consider a second subsample where survey respondents expressed their opinions about the proper role of government in terms of environmental, health and safety regulations. In a formal maximum likelihood selection model, we find some evidence of a slight liberal (pro-regulation) bias that may stem from non-random selection, but the effect is not statistically significant and the hypothesis of "no liberal/conservative bias" cannot be summarily rejected. Less sophisticated models in the class of "propensity score" corrections show minimal significant effects on selection propensity in the regulatory preference outcome models but the distortions are quantitatively very tiny.

Friday 11:30 am-12:15 pm - AAPOR ABSTRACTS

POSTER SESSION

***Analysis of Reinterview Data from NLSY97 rounds 2 through 5*, Parvati Krishnamurty, NORC; KRISHNAMURTY-PARVATI@norc.uchicago.edu**

In the National Longitudinal Survey of Youth 1997, approximately 10% of respondents go through a telephone reinterview after each round. The main goals are to assess the quality of the interview data and to make sure the interview was conducted properly. Questions asked include selected questions from the main survey, some of which change from year to year. In this paper, we examine reinterview data from rounds 2 through 5 to identify unreliable survey questions, measure response variance and explore methodological issues that may explain inconsistencies in responses. For questions that are asked in multiple rounds, we look at whether responses have become more reliable in successive rounds as respondents have become more familiar with the questionnaire. Finally, we look at demographic characteristics and time elapsed between interview and reinterview to see if these are related to inconsistency of responses. We find that there is a higher degree of inconsistency for questions on income and highest grade attended while responses to other questions like the type of housing are quite consistent. Looking at successive rounds, we find that the reliability of responses to the income question has improved. Estimates of the simple response variance such as the gross difference rate are unbiased only if the reinterview is independent and an exact replication of the interview. In practice, there are operational constraints which do not allow us to achieve such an exact replication of the interview questions in the reinterview. Differences in wording, mode and context of the questions in the interview and reinterview could potentially explain some of the variability in the responses.

***A Transdisciplinary Approach to Protocol Development: A Case Study*, Melissa Clark, Brown University; Melissa_Clark@brown.edu, Suzanne Colby, Brown University; Suzanne_Colby@brown.edu, Julie Boergers, Rhode Island Hospital; JBoergers@lifespan.org, Christopher Kahler, Brown University; Christopher_Kahler@brown.edu, Susan Ramsey, Rhode Island Hospital; SRamsey@lifespan.org, Charles Neighbors, The Miriam Hospital; Charles_Neighbors@brown.edu, and Johanna Lewis-Esquerre, Brown University; Johanna_Lewis-Esquerre@brown.edu**

Transdisciplinary science involves the integration of theoretical and methodological perspectives from different disciplines. In 1999, Brown University, in collaboration with Harvard, Yale, and Brandeis Universities, was one of seven institutions to be awarded a 5-year Transdisciplinary Tobacco Use Research Center grant. This collaboration, referred to as the New England Family Study (NEFS), included

three major projects organized around a unifying theme. In addition, six cores served as resources for the project. Participants in the NEFS included three generation families. The Measures and Methods Core used a two-phase process to develop and insure the fidelity and integrity of face-to-face assessments of constructs related to tobacco use and dependence as well as other co-morbid conditions. The assessments were categorized by three types of instrument development: Standard, Significantly Adapted, and Newly Developed. Standard Assessments consisted of instruments from previous studies that only required formatting for consistency, minor word editing, instructions, and other practical applications for differences in cognitive abilities and experiences of each participant age group. Significantly Adapted Assessments consisted of instruments from previous studies that we extensively modified for use in the NEFS. Types of modifications made to Significantly Adapted Assessments included: the addition and/or revision of items; changes in inclusion or exclusion criteria to the instrument; and refinement of particular constructs measured by the assessment. Newly Developed Assessments consisted of instruments developed specifically for NEFS. Overall, there were 47 Standard, 59 Significantly Adapted and 32 Newly Developed assessments. Our experiences indicate that the major strength of our transdisiplinary approach was a set of measures and assessment methods that utilized the expertise of researchers from multiple disciplines. Weaknesses included the increased length of time to complete instrument development and the concern that some constructs were not measured with the ideal amount of detail that would occur in a single discipline approach.

***Too much of a good thing? Working through establishment gatekeepers*, Andrew Zukerberg, U.S. Census Bureau; andrew.l.zukerberg@census.gov, Andrew Soderborg, Census; Andrew.J.Soderborg@census.gov, Randy Parmer, Census; Randall.j.Parmer@census.gov, and Steve Tourkin, Census; steven.c.tourkin@census.gov**

In establishment surveys, gatekeepers often prevent interviewers from reaching the sampled person. Many surveys have developed methods to get around gatekeepers or enlist them as agents in the survey process. Often these efforts target an individual. For the Schools and Staffing Survey (SASS), school districts function as gatekeepers for the schools under them. Three scenarios were anticipated for the 2003-04 SASS. 1) If a district is contacted before the school and gives permission to conduct SASS, it could increase overall response rates. 2) If a district is contacted before the school and refuses to participate, it could lower overall response rates. 3) If districts are not contacted before the school, schools could request district permission to participate, delaying completion of the survey and increasing costs. In order to determine the best way to handle district contacts, an experiment was conducted in three Census Bureau regional areas. Approximately half of the school districts in each office were contacted by phone several months before the survey was conducted to discuss the survey and any information they would need before approving the survey. If information or formal application was required, it was prepared and sent to the district shortly after the call. In the other half of districts, a standard prenotice letter was sent to the district at the start of data collection. This paper reports on the impact on school response under those scenarios and makes recommendations for handling establishment gatekeepers.

***Dual-Incentive Experiment with New Business Owners*, Janice Ballou, Mathematica Policy Reseach Inc; jballou@mathematica-mpr.com, Dave DesRoches, Mathematica Policy Research, Inc; ddesroches@mathematica-mpr.com, and Frank Potter, Mathematica Policy Research; fpotter@mathematica-mpr.com**

An important dimension of survey quality is the participation of selected sample members in a study. The use of incentives to encourage participation has been one of the factors survey researchers have studied. Overall, research on incentives suggests that some type of incentive will improve participation. However, research on incentives needs to be conducted with various types of populations to identify when an incentive is more or less effective; and, in addition, to identify the optimal timing of giving the incentive--before or after the sampled study member has completed the questionnaire. New business owners are a particularly challenging population to interview. A pilot test for a survey of new business owners sponsored by the Ewing Marion Kauffman Foundation will include two incentive experiments. Incentives will be offered at two time points in time to randomly selected new business owners--(1) pre: included in the introductory letter and (2) post: when eligibility has been confirmed. For Experiment 1, half of the sample will receive $1.00 in the introductory letter and half will not receive any pre-incentive. Experiment 2 will offer a $50.00 post-incentive to half of the new business owners who qualify after a series of eligibility questions. At each stage we will analyze the impact of the incentive on survey participation. In addition, among the 1,000 completed interviews there will be an analysis of the four resulting experimental groups: (1) dual incentives-received both pre and post, (2) $1.00 pre-only, (3) $50.00 post only, and (4) no incentive.

***The RAS Model: A Simple Test*, Agnieszka Dobrzynska, Universite de Montreal; agnieszka.dobrzynska@umontreal.ca, and André Blais, Universite de Montreal; andre.blais@umontreal.ca**

In 1992 John Zaller formulated the most influential theory of opinion formation, the Receive-Accept-Sample, or RAS, model. The theory is based on four axioms about the conditions under which a message is received, and, if received, accepted or rejected. According to Zaller the reception of a message depends on the intensity of the message and on individuals' general level of political awareness. And the most sophisticated should be more able to resist a message when the latter does not accord with their predispositions. The RAS model, though highly influential, remains largely untested. The author himself provides only indirect empirical evidence that is consistent with the reception and resistance axioms. The aim of this paper is to propose a rigorous and direct test of the RAS model. The study deals with the 1988 Canadian election, an election that was fiercely fought over one central issue, the Free Trade Accord with the United States. We use the 1988 Canadian Election Study campaign rolling cross-section survey, and we test Zaller's propositions about who is most likely to receive, and then accept or reject the parties' messages about the central issue of the election.

***Knowledge, Question Format and Propensity to Guess*, Nick Allum, University of Surrey; n.allum@surrey.ac.uk, Nick Allum, University of Surrey; n.allum@surrey.ac.uk, and Patten Smith, BMRB International; patten.smith@bmrb.co.uk**

It is increasingly apparent that factual knowledge is an important determinant of heterogeneity in public opinion in a broad range of areas (Delli Carpini and Keeter 1996). Presidential voting, issue preference and electoral participation are only some of the more prominent areas of political decision making that have been shown to depend crucially on an individual's level of political knowledge. The measurement of this key construct in surveys is, however, far from straightforward. In addition to the potentially adverse effects on response of administering test batteries, contextual cues and personality related 'propensity to guess' have both been shown to bias estimates of between group differences in political knowledge (Mondak 2000; Nadeau and Niemi 1995). This has led Mondak to recommend "closed-ended items on which DKs are discouraged" (2004, p509) as the standard format for political knowledge items in surveys. This strategy is, however, potentially problematic in two main respects: (1) efforts to eliminate don't know responses may only be partially successful and (2) encouraging guessing amongst the ignorant will bias estimates of knowledge for the sample as a whole and reduce the reliability of

summated scales. In this study we randomly allocate respondents to one of three experimental conditions in order compare Mondak's proposed format with the more conventional approach set out by Delli Carpini and Keeter. Additionally, we introduce a third response format in which respondents are not asked whether a statement is true or false but whether it is definitely true, probably true, probably false, definitely false (or don't know). In the latter two conditions, initial don't know responses are subsequently probed for a 'best guess'. We compare response distributions across the three conditions relative to expectations under a guessing strategy to draw conclusions about the relative merits, both statistical and practical, of the three different approaches.

***The Use of the CSAP's Core Measures Initiative: Cross-Sectional Development but Repeated Measures Application*, Mariah Storey, University of Wyoming; riah@uwyo.edu, and Eric L. Canen, University of Wyoming; ecanen@uwyo.edu**
In an effort to identify common evaluation measures and to make comparative assessments among the numerous substance abuse prevention programs it funds, the Center for Substance Abuse Prevention (CSAP) requires all programs to be evaluated according to the Core Measures Initiative. Survey items from the Core Measures Initiative assess beliefs, attitudes and opinions regarding drug and alcohol use among youth, 12-17 years old. Most of these survey items were originally developed in a cross-sectional setting of 6th, 8th, 10th and 12th grade students, and the survey was meant to be re-administered annually or biennially to all 6th, 8th, 10th and 12th grade students under another cross-sectional design. Using this cross-sectional design, the survey items have been shown to be valid and reliable. Unfortunately, most of the prevention programming funded through CSAP uses pre and post testing and a matched repeated measures design to evaluate their effectiveness. The impact of a matched repeated measures design was not taken into consideration when the survey items were being developed and were also not addressed during the assessment of scales' reliability and validity. Data from a core measures based statewide cross-sectional survey and data from a set of repeated measures surveys with the same items will be used to explore this issue. In particular, we will examine the ways in which adaptation of a large scale cross-sectional surveys to a repeated measure designs can be both helpful and problematic. The wording of the items for the current cross-sectional design will be compared to the wording needed in a repeated measure design. Additionally, how the type of design (cross-sectional or repeated measure) impacts data results will be discussed. Finally, comparisons of reliability and validity estimates within these two designs will be analyzed and discussed.

***Using survey data to provide small-area resource allocation models*, Bill Blyth, TNS; bill.blyth@tns-global.com, Graham Moon, University of Portsmouth; gmoon@uop.ac.uk, and Gemma Holt, University of Portsmouth, UK; gemma.holt@port.ac.uk**
The growth of evidence based policy in western democracies has resulted in an increasing demand for the provision of key performance indicators. This has occurred across a wide range of public policy areas. Implementation of national policy requires local delivery and with fixed resources local needs assessments and targets. However, often the public sector's administrative systems are incapable of delivering relevant national and local information. Instead such needs often are met by large-scale quantitative surveys. Invariably these are limited to providing national and regional rather than local delivery point data. To fill this information gap in the UK public health sector TNS and the University of Portsmouth Institute for the Geography of Health have over the last four years been carrying out a programme of data collection and model development.. Using a range of data including census, national probability surveys of self-reported health and very large non-probability surveys, small-area models have been built of long-term chronic disease prevalence, undiagnosed and risk-population prevalence, and related lifestyle behaviours. Conditions modelled include cardio-vascular disease, type II diabetes, smoking and obesity. The method employed is a form of multi-level modelling. Very high degees of validity have been obtained- in some instances over 80% match on input and hold-out data. The models are built up from the smallest areas of census goegraphy and thus can be aggregated as wished and combined with other adminstartive data- for example at screening clinic level. The paper will describe this work with illustrations. It will further discuss how it can be extended not only to public health in other countries, but also to other areas of public policy.

***Interpreting Response Latencies: A Mixed Method Approach.*, Christine Carabain, Academic; cl.carabain@fsw.vu.nl, and Wil Dijkstra, Vrije Universiteit; w.dijkstra@fsw.vu.nl**
In attitude research, response latencies are often used as an indicator of attitude strength. The attitude strength depends on the associative strength between the attitude object and the object's evaluation. The assumption is that the stronger this association, the faster the respondent should be able to express an attitude. This paper provides an alternative explanation with regard to response latencies and attitude strength. In this study, Dutch students were asked both closed and open questions about their attitudes toward Muslims in the Netherlands. Half of the interviews started with the closed questions and the other half started with the open questions. The response latencies on the closed questions were compared with the answers to the highly comparable open questions. The results of this study demonstrate a link between the response latencies on the closed questions and the answers to the open questions. Respondents who verbalized multiple considerations answering open questions needed more time to answer the corresponding closed questions then respondents who did not verbalized these considerations. This effect only appeared when the interview started with the closed questions. Apparently, when respondents first answered open questions and thus already thought and talked about the subject enabled them to choose an answer alternative faster, even if they had multiple considerations with regard to Muslims in the Netherlands. These results suggests that longer response latencies do not always have to be the result of an often assumed 'weaker' attitude or uncertainty of respondents. This study demonstrates that these longer response latencies may as well be the result of respondents facing difficulties choosing an answer category because of having multiple considerations or a more nuanced attitude toward the attitude object.

***Survey Experiences Influence Survey Attitudes and Participation Propensity*, Robert Brage, Department of Psychology, Stockholm University; rbe@psychology.su.se, and Lars R. Bergman, Dep. of Psychology, Stockholm University, Sweden; lrb@psychology.su.se**
General and specific survey attitudes and survey experiences were assessed in a CATI probability sample "survey on surveys" with a total of 989 sample persons in Sweden. The three main samples were taken from 1. the general Swedish population, 2. participants in the Labor Force Survey (LFS), and 3. participants in the Survey of Living Conditions (SLC). The response rates in the main samples were 70, 83, and 64 percent, respectively. Also two samples of nonrespondents from LFS and SLC were included, but the response rates were just 28 and 26 percent, respectively, for these samples. In this presentation, the influence of different survey experiences on survey attitudes, participation intent and respondent behavior is discussed. Attitudes toward "Market Surveys" were less positive than toward noncommercial "General Surveys" and the trust in Statistics Sweden was higher than for private survey agents. Attitudes toward General Surveys vary across the samples but attitudes toward Market Surveys do not. The personal SLC interview appear, on the average, to have affected attitudes and

behavior negatively, while the LFS, with its eight short telephone interviews, seems to have fostered future respondents in General Surveys. Feeling pressured to participate was, among SLC respondents, a frequent reason for participating and the associations between General Surveys and Market Surveys were larger in this sample. With the lounge of large scale web-surveys of all kinds, there is a risk that people's ability to pick out the legitimate, purposeful surveys of value to society will get blunted. At least if there are not clear differences in recruitment strategies, quality of interaction, and communication of the purpose. And, we will perhaps in General Surveys end up with the respondents that participate in all kinds of surveys just because they enjoy participating.

***Response Propensity Patterns among Respondents Selected in Multiple Years of a Repeated Cross-sectional Survey*, Sherman Edwards, Westat; ShermEdwards@westat.com, and John Rauch, Westat; Johnrauch@westat.com**

The Medicare Managed Care CAHPS®, a mail and telephone survey of beneficiaries enrolled in Medicare Advantage health plans about their experiences getting health care, has been conducted annually since 1998. Equal-sized cross-sectional samples are selected from each eligible health plan, or, for large plans, in geographic plan sub-units. The 2004 sample represented 286 plan reporting units from 173 Medicare Advantage health plan contracts. The primary purpose of the survey is to provide consumers with information to compare health plans available in their area. While overall MMC-CAHPS response rates have been at 80 percent or higher, the response rate by plan reporting unit has varied considerably. Previous methodological research has indicated that differential response rates across plan reporting units likely do not introduce bias into the health plan comparisons. Each year's sample includes some enrollees who had been selected for one or more prior years' surveys. About 7 percent of the 2004 sample (n=170,000) were also sampled in 2003, but the proportion by plan reporting unit ranged from 0 to more than 90 percent repeat sample members. If repeaters are different from those newly sampled, this differential could bias health plan comparisons. Previous, as yet unpublished research has suggested that CAHPS measures are stable over time among repeat respondents. However, this research did not consider the role of nonresponse patterns over time. This paper will examine the response propensity patterns of repeat sample members over multiple cross-sectional iterations of the MMC-CAHPS. It will also assess how these patterns may ultimately affect comparisons of health plans given the large differences in proportion of repeat respondents across plans. While the results will not be directly generalizeable to other surveys or populations, any findings may suggest areas of inquiry for other repeated cross-sectional surveys of relatively small populations, and for longitudinal surveys.

***First-level interactions and final outcomes: Adding value to behavior coding*, Ashley Landreth, U.S. Census Bureau; ashley.denele.landreth@census.gov, and Jennifer Hunter, US Census Bureau; jennifer.e.hunter@census.gov**

Behavior coding methodology employs a systematic analysis of verbal interactions between an interviewer and a respondent. Its purpose is to identify overt problems by quantifying the interviewer/respondent behaviors that connote difficulties in both asking and answering survey questions. Typically this is accomplished by limiting the analysis to the "first-level" of interviewer/respondent interaction, because major problems present themselves when the question is first asked and as respondents initially react or respond. While first-level coding is necessary for detecting flawed questions, it does not provide information about the final response "outcome" when respondents at first fail to produce an answer that meets the measurement objective. Flawed questions with response problems that cannot be overcome through interviewer/respondent negotiation may suffer from fundamentally different issues than flawed questions that are eventually resolved. Coding the final response outcome may provide more robust analysis of question performance, without creating undue burden for coders. In this research, in addition to analyzing the first-level interaction, we also coded and analyzed the final response outcome with regard to questions' measurement objectives. The outcome was coded using a basic framework of behavioral codes applied to audiotaped face-to-face interviews that collected demographic data as part of the nonresponse follow-up to the 2004 Test Census. This research explores the advantages of coding this aspect of the interviewer/respondent interaction; it provides additional information about the response process, data quality, and potential approaches to fixing question problems. In addition, we will discuss whether the additional data generated justify the extra effort on behalf of coders.

***Evaluating the Impact on Data Quality of Conversion from PAPI to the iPAQ Handheld Computer for In-person Data Collection*, Randall Keesling, RTI International; srk@rti.org, and Scott Bell, Nielsen Media Research; Scott_bell@tvratings.com**

The National Household Enumeration Survey, an in-person interview conducted in English or Spanish annually since 1994, has been evaluating the idea of converting its current paper-based instrument to computer driven systems. A primary concern in considering conversion to the technology is the potential effect of the collection device on data quality. This presentation will focus on procedures and software features that were developed and evaluated for their impact on data quality. There were concerns that the accuracy of the recorded responses may decrease given the fact that, with the PAPI questionnaire, data recorded on the paper form were key-entered by two professional data entry operators while running a software program that would detect any keying errors. Using the handheld computer, the function of the two keyers would be replaced with a single field interviewer entering the data once with no obvious check for keying errors. Additionally, there were questions regarding the possible impact on response rates and/or the data being provided by a respondent given the presence of the handheld device in place of the paper form. Would the responses differ or be skewed relative to the data collected previously via PAPI? The objective of the evaluations was to identify and quantify the level of impact the change in data collection media had on data quality; specifically regarding completeness and accuracy. Data will be presented comparing responses collected with the handheld computer with that of the PAPI approach. Implications and recommendations for converting from PAPI to electronic field data collection will be discussed.

***A Profile of Self-Selecting Web Respondents*, Paul Guerino, American Institutes for Research; pguerino@air.org**

Web surveys are increasingly being adopted by survey organizations as an alternative to mail and telephone surveys. While there is a growing body of literature on respondent behavior to Web-only surveys, little is known about the charactistics of Web respondents in multimode surveys. The purpose of this paper is to profile the Web respondents to the National Survey of Recent College Graduates (NSRCG), a biennial survey of recent bachelor's and master's degree recipients in science, engineering, and health. In 2003, the NSRCG was conducted as a mail survey with the option to complete the survey on the Web or by telephone. This paper describes (1) who self-selected into the Web mode and how they differ from respondents to the other modes in terms of gender, race/ethnicity, age, degree field, employment status, and occupation; (2) patterns of Web behavior including login days and times; (3) interactions with email reminders, incentive offers, and other contacts, and (4) the incidence of "lurkers," that is, people who view the Web survey but do not complete it. Depending on the development costs and number of respondents, interviews completed by Web are potentially less expensive than mail or telephone interviews. The paper concludes with recommendations for maximizing Web responses in a multimode environment.

***The Arbitron/NMR PPM Response Rate Test*, Beth Webb; beth.webb@arbitron.com, and Arianne Buckley- Media Research; arianne.buckley@arbitron.com**

Arbitron has developed an electronic personal meter (Portable People Meter, PPM) that can automatically detect audio exposure to encoded radio and TV signals. We ask panelists to wear this pager-sized device everyday from the time they rise to the time they retire in order to measure their media exposure. Beginning in May 2003 Arbitron and Nielsen Media Research (NMR) fielded a response rate test to determine whether a PPM panel could achieve sample performance comparable to that of traditional set-top meter panels when using the same in-person recruitment methods. In addition, a second test group was fielded which used a hybrid approach, that began with telephone recruitment and then used in-person recruitment only when telephone recruitment was unsuccessful. The goal was to determine whether this more efficient approach could perform as well as an all in-person approach. Each PPM test group consisted of approximately 100 households each. Both test groups originated as standard NMR area probability samples. Also, both test groups employed the "basics" and "alternates" sampling method, in which refusing basics are replaced with substitute households in the same general area which has the same presence of children and cable status. Group 1, the "In-person" group, was recruited and installed using the standard NMR in-person methods. Group 2, the "Hybrid" group, began with telephone/mail, followed by in-person contact when the telephone/mail methods did not gain a "basic" install. Monthly analyses looked for meaningful patterns of differences between the test groups in cooperation rates among the basic sample, sample performance indicator (SPI) which is the accredited response metric for set-top meter panels, monthly turnover rates, demographic characteristics or compliance with the measurement method. Both test panels were maintained through May 2004 to see if the results found during the first few months of the test would hold up over time.

***Measurement Issues in Behavioral Self-report Measures: Effects of Response Format, Respondent Load, Number of Elements, and Time Frame Reference*, Randall Thomas, Harris Interactive; rthomas@harrisinteractive.com, Jonathan Klein, University of Rochester; jklein@harrisinteractive.com, Lisa Wilding, Harris Interactive; lwilding@harrisinteractive.com, Susan Behnke, Harris Interactive; sbehnke@harrisinteractive.com, and Robin Repass, Harris Interactive; rrepass@harrisinteractive.com**

While we have discovered significant differences between Yes-No Grids and Multiple Response techniques in behavioral self-reports (Thomas, Behnke, Lafond, & Smith, 2003), we sought to expand our investigation of alternative techniques and explored how the various self-report techniques varied with regard to time interval. For each specific time frame, respondents were randomly assigned to one of the following response formats: Yes-No Grid, Multiple Response, Numeric Box, 5 Category Grid, 5 Category Drop Down Box. Since one of the explanations as to why the Multiple Response format leads to lowered behavioral self-reports is that respondents are more likely to satisfice with a multiple response format, we varied the number of elements for the respondent to consider. In addition, since we believed that a multiple response format would be more prone to satisficing based on survey length, we randomly assigned respondents to complete the self-report task either at the beginning or near the end of a 20 minute survey. We replicated the finding that the Yes-No Grid led to higher behavioral self-report than the Multiple Response format across time intervals. However, our results for the various response formats as a function of survey position did not fully support a simple satisficing explanation for the differences between formats.

***Effects of suggestive and non-suggestive probing on different types of inadequate answers*, Yfke Ongena, Vrije Universiteit Amsterdam; YP.Ongena@fsw.vu.nl, and Wil Dijkstra, Vrije Universiteit; w.dijkstra@fsw.vu.nl**

In telephone interviews, respondents may follow the peripheral or the central route when processing survey questions. When respondents follow the peripheral route, they process the question with little attention, using only the question as a cue to answer, rather than response alternatives, as is usual in ordinary conversations. This conversational style of responding is likely to result in mismatch answers (answers not matching the required answering format). In case of the central route, respondents process the question quite thoughtfully and are likely to give adequate answers. However, mismatch answers may occur when questions are ambiguous or difficult (creating task uncertainty, i.e. difficulties in understanding the meaning of questions or creating state uncertainty, i.e. difficulties in retrieving relevant information). Mismatch answers necessitate the interviewer to probe in order to get scorable answers. Probes can be non-suggestive (offering more alternatives) or suggestive (offering one alternative). A non-suggestive probe forces the respondent to process the question more thoughtfully, yielding longer reaction times irrespective of the type of mismatch answer preceding the probe. However, for suggestive probes we expect differences related to the preceding mismatch answers. A suggestive probe after a conversational mismatch, leads respondents to continue following the peripheral route, using the suggestion as a cue. When a suggestive probe occurs after mismatch answers caused by task uncertainty, respondents are assumed to switch from their initial central route to the peripheral route. In both cases the reaction times will be short and the respondent is likely to accept the suggestion. However, after a mismatch answer caused by state uncertainty, respondents are likely to continue following the central route, not necessarily accepting suggestions. An experiment was conducted varying question wordings for difficulty, ambiguity and conversation-likeness. Interviewers were assigned to either a suggestive or non-suggestive probing condition. Reaction times were measured for all utterances.

***Survey non-response in a national area probability sample as a dimension of survey quality; an analysis of community characteristics*, Jibum Kim, NORC; jbk7000@naver.com, and John Sokolowski, NORC; Sokolowski-John@norc.org**

Most area probability surveys collect little to no information about non-respondents. This paper aims to explore the community characteristics of survey non-respondents in a nation area probability survey. Data selected for use is obtained from the 2004 General Social Survey, an area probability sample of 6260 housing units. Results from the survey are nationally representative of English speaking residents. We expect that wealthy and racially segregated areas are more likely to have high non-response rates. This paper examines community characteristics in relation to survey non-response and proposes ways to improve response rates among high non-response communities in national area probability surveys.

***What's Up Doc? Mixing Web and Mail Methods in a Survey of Physicians*, Timothy Beebe, Mayo Clinic; beebe.timothy@mayo.edu, Giles R. Locke, Mayo Clinic; locke.giles@mayo.edu, and Sunni Barnes, Mayo Clinic; barnes.sunni@mayo.edu**

To increase response rates, household surveys often turn to mixed-mode designs whereby instruments are designed to be administered via mail, web, telephone, and/or in-person and respondents are allowed to respond to the form most appropriate for them. The application of mixed-mode designs to physician surveys seems natural given that low response rates to single-mode physician surveys are common. Evidence presented at last year's AAPOR conference suggested that one particular mode combination, mail and web, might prove useful in extending the coverage of the survey to a broader mix of physicians because the profile of providers responding online was somewhat different from those responding to a mail survey (Losch, Thompson, & Lutz; Peugh & Zapert). This paper presents the results of an

experiment testing two different mixed-mode designs representing two combinations of mail and web surveys. The survey focused on physician evaluations of the Mayo Electronic Medical Record (EMR). 500 physicians were randomly assigned to receive either an initial mail survey with a web survey follow-up to nonrespondents or its converse - an initial web survey followed by a mailed survey to nonrespondents. In both conditions, all physicians received a reminder one week after the initial mailing via the medium corresponding to the initial contact (web v. mail). Key research questions addressed in this paper include: Can mixing mail and web methods increase response rates among physicians? Which of the two combinations tested proved most effective? How does mixing modes impact the participation of different types of physicians? What effect do the different combinations have on the resulting evaluations of the EMR?

***Collecting Geographic Location Information in an RDD Survey*, David Grant, UCLA; dgrant@ucla.edu, Anthony Ramirez, UCLA; tonyram@ucla.edu, Sherman Edwards, Westat; ShermEdwards@westat.com, and John Rauch, Westat; Johnrauch@westat.com**
Geographic information systems (GIS) add a new dimension to survey research. However, the ability to incorporate and analyze GIS in survey data depends on collecting accurate information to geocode. Collecting this type of information can be challenging in population-based random-digit-dial (RDD) telephone surveys, where the geographic location of a residence must be voluntarily provided by a respondent, and subsequently geocoded to discrete latitude and longitude coordinates. This paper explores 1) several methods for collecting geographic information; and 2) compares success rates of geocoding data collected in two cycles of the California Health Interview Survey (CHIS). In the first cycle, names of the street of residence and the nearest cross-street were collected. In the second cycle, a series of questions dependent upon the known characteristics of the respondent's telephone number and the existence of a pre-matched address were used. Results demonstrate that most respondents were willing to provide detailed location information and that having multiple options for collecting geographic information can reduce administration time while improving accuracy of these data.

***The Utility of Extended Longitudinal Profiles in Predicting Future Health Care Expenditures*, Steven Cohen, AHRQ; scohen@ahrq.gov, Trena Ezzati-RIce, AHRQ; tezzatir@ahrq.gov, and William Yu, AHRQ; wyu@ahrq.gov**
Given the high concentration of health care expenditures in a given year among a relatively small percentage of the population, a prediction model that can accurately identify the persistence of high levels of expenditures is an important analytical tool. This type of modeling effort also enhances the ability to discern the causes of high health care expenses and the characteristics of the individuals who incur them. This feature also applies to prediction models that can accurately identify those individuals with persistently low or average levels of expenditures. The models that are presented have particular relevance as statistical tools to facilitate efficient sampling strategies that permit the selection of an over-sample of individuals likely to incur high levels of medical expenditures in the future. Furthermore, such modeling efforts are particularly attractive to assist in the targeting of disease management programs to high cost cases, which may facilitate reductions in the concentration of overall future year health care expenditures. The national Medical Expenditure Panel Survey (MEPS), sponsored by the Agency for Healthcare Research and Quality (AHRQ), is particularly well suited for supporting modeling efforts of this nature. By collecting nationally representative information on health care utilization and expenditures experienced over a two year period, the survey informs analyses that examine the persistence of health care expenditures over this two year window. Furthermore, as a consequence of the survey's integration with the National Health Interview Survey, an additional prior year of information on health status, utilization and socio-demographic characteristics is available to inform modeling efforts. In this paper, the performance of alternative prediction models to identify future high expenditure cases is evaluated. Particular attention is given to investigating the added value of an extended longitudinal profile in accurately predicting future health care expenditures.

***Nonresponse in Telephone Surveys: The Reporting of Response Rates and Outcome Measures*, Richard Seltzer, Howard University; rseltzer@howard.edu**
Non response rates in most telephone surveys have been growing since the 1950s. There is considerable controversy about the extent and effect of this problem. In 2003 I gathered outcome measures from eleven different survey firms covering 18 different surveys. I extended this research given the last pre-elections surveys in 2004 from 9 survey organizations. Despite standardization on how to report outcome measures by AAPOR - there is still too much latitude on how such rates are reported. I argue that data on nonresponse should be more broadly available and that further standardization needs to occur.

***Explaining Survey Quality to Data Users: Can we get them to care?*, Susan P. Love, U.S. Census Bureau; susan.p.love@census.gov, and Deborah H. Griffin, U.S. Census Bureau; deborah.h.griffin@census.gov**
The U.S. Census Bureau is committed to the production of high quality survey data to meet the demographic needs of America's communities, and the American Community Survey (ACS) was designed to produce such data. Using independent monthly samples and a multiple mode data collection strategy, the survey produces single and multi-year estimates of social, economic, and housing characteristics for a wide range of demographic groups, and for areas as large as the nation and as small as census block groups. We have identified and defined a set of quality indicators that describe the accuracy of the ACS data. Managers use these quality indicators to monitor the ACS and to identify areas requiring special attention, and they are released with the yearly survey data, making them easily accessible to data users. We want to encourage users to take these measures into account when using ACS data. This poster will summarize the key quality measures used by the ACS, their definitions and interpretations. We will provide access for conference attendees to the ACS website that allows data users to obtain the measures of survey accuracy at state and national levels. The website includes information on how to interpret these measures, and provides guidelines to help educate ACS data users on the potential limitations of the survey data. Feedback on the effectiveness of the website will be actively sought from conferees.

***The National Survey of America's Families, 1997-2002: A Project Summary of Response Rate Initiatives, Implementation, and Results*, Timothy Triplett, Urban Institute; ttriplet@ui.urban.org, Adam Safir, Urban Institute; asafir@ui.urban.org, and Natalie Abi-Habib, Urban Institute; nabihabi@ui.urban.org**
The decline in telephone survey response rates continues in spite of exhaustive efforts by survey data producers to countermand the trend. Understanding the relationship between nonresponse procedures and survey estimates is central to the evaluation of the severity of nonresponse bias and can provide a clearer picture of the future of telephone-based data collection. Our research presents a summary of the response rate initiatives and survey operations undertaken by the National Survey of America's Families (NSAF) across six years and three rounds of data collection. NSAF is a large-scale national survey of the economic, health, and social characteristics of children, adults under the age of 65, and their families. The survey was conducted in 1997, 1999, and 2002, yielding information on over 40,000 families and

100,000 persons in each round. For each round of data collection, a number of strategies were employed to improve declining response rates, including various systems of pre-notification mail-outs, incentive structures, changes in call scheduling, refusal conversion, and refusal sub-sampling. Our analysis focuses on the changes made in data collection procedures and call disposition results across rounds. We examine how changes made to protocols for incentive distribution, call schedules, maximum call attempts, and refusal conversion impacted response rates over time. We also assess the impact that these procedural and response rate changes may have had on sample composition across the three rounds. Finally, in summarizing our results, we make recommendations for improved telephone survey techniques and identify areas for future research.

***Changing Response Pattern over Time: Results from Statewide Telephone Surveys Conducted in 1999 and 2004*, Cyndi Garvan, University of Florida; cyndi@biostat.ufl.edu, Colleen Porter, University of Florida; cporter@phhp.ufl.edu, R. Paul Duncan, University of Florida; pduncan@phhp.ufl.edu, Qin Li, University of Florida; qli@biostat.ufl.edu, and vijaya komaragiri, University of Florida; vijay@biostat.ufl.edu**

The Florida Health Insurance Study (FHIS), was created in order to provide reliable estimates of Floridians without health insurance, both statewide and by geographic and demographic subgroups. Large-scale statewide RDD telephone surveys were conducted in 1999 (n=14,011) and in 2004 (n=17,435). Both telephone surveys were conducted by University of Florida's Bureau of Economic and Business Research (BEBR) Survey Research Center, with Sawtooth WinCATI software, using the almost-identical instrument, similar sample designs, comparable interviewer training protocols, and a mail followup procedure that used the same letter. Yet markedly differences pattern of response were noted between the two surveys. The pattern of disposition codes differs significantly by year (p-value of chi-square test < .0001). The results of disposition codes included the following: complete interviews (14.8%,1999; 12.8%,2004), partial interviews (0.2%, 1999; 0.0%, 2004), refusals and break-offs (8.7%, 1999; 21.0%, 2004), non-contacts (5.4%, 1999;2.0%,2004), unknown if household (17.4%, 1999; 29.2%,2004), other (3.6%,1999; 0.4%,2004). The observation of a markedly different pattern of response when so many other aspects of survey implementation remained constant provide a valuable opportunity for comparison that may yield greater understanding of the nature of survey response. To further explore these differences we examined associations of disposition codes and geographic and socioeconomic factors (e.g., race/ethnicity, income, age) measured at the telephone exchange level that are provided by a commercial vendor (GENESYS).

***Reaching Low-Income Populations for Telephone Interviews*, Ellen Marks, RTI International; emarks@rti.org, and Murrey Olmsted, RTI International; molmsted@rti.org**

Achieving equitable participation from a broad array of demographic groups is of concern to survey researchers. The issue is particularly important in studying low-income populations who are often hard to reach through telephone surveys. A wide variety of factors contribute to low response rates among this group including lack of telephone service, frequent moves, and wariness toward the survey. This paper summarizes the literature on reaching low-income populations by telephone and reports on special efforts to maximize their participation in a telephone survey conducted with parents or guardians of children enrolled in Head Start. We expect that response rates among low-income populations will improve when special efforts are made to remove barriers that impede their participation. For the study discussed in this paper, several measures were implemented to maximize cooperation and response rates, such as working with Head Start to get better telephone numbers for sample members with nonworking telephone numbers, sending field representatives to sample members' homes who could not be reached by telephone, offering a cell phone for respondents to use to conduct the interview, mailing Federal Express letters to refusal cases, increasing the incentive paid to respondents, and hiring specialized interpreters (e.g., Arabic and American Sign Language) to accommodate sample member needs. The presentation will report on the number and types of attempts to reach sample members and will assess the relative success of various methods used to encourage participation in the survey. Additionally, the presentation will address issues related to the relative costs and benefits of these activities, the effect on increasing cooperation rates, and demographic comparisons of easy- and hard-to-reach study participants. Implications for surveys of similar populations will be discussed.

***Advance Notice and Confidentiality Statements for Listed-Assisted Telephone Surveys*, Brian J. Meekins, Bureau of Labor Statistics; meekins.brian@bls.gov, and Roberta L. Sangster, Bureau of Labor Statistics; sangster_r@bls.gov**

Research on the use of advance letters without monetary incentive suggests a modest impact on cooperation for list-assisted telephone surveys. In a recent review, Sangster estimated the average impact for response rates for RDD surveys was between zero and three percent, while Cantor et al. (2002) noted the impact appears to be declining over time. However, the OMB was concerned about the response rate for the Telephone Point of Purchase Survey (TPOPS) and required the BLS to conduct an advance letter study. The use of the advance letter provided an opportunity to test the impact of a lengthy confidentiality statement. The statement of confidentiality was included in FAQ form on the back of the advance letter, and, for some, eliminated from the questionnaire. In general, assurances of confidentiality have been shown to slightly impact response, cooperation, and data quality, for sensitive topics (Singer, Von Thorn, and Miller, 1995).
However, in a recent list-assisted RDD study, Singer (2003) found that confidentiality statements can increase perceived risk and reduce response. The TPOPS is a large, nationally representative list-assisted, RDD survey. A treatment group of households with listed addresses received the advance letter, while a control group of listed households did not. A treatment group of those who received the letter skipped the confidentiality statement while a control group that received the letter received the confidentiality statement. This paper examines response and cooperation for unlisted households, listed households who did not receive a letter, and listed households that did receive a letter. In addition, it examines response for the different administration of the confidentiality statement. The timing of refusals (immediate, during the reading of the statement), effort expended in contacting or interviewing, and data quality are also examined. Certain subpopulations are found for which the treatments are more effective in encouraging cooperation.

***Does Imputation Bias Lead to Finding Significantly More Uninsured in the Current Population Survey's Estimates of Health Insurance Coverage?*, Michael Davern, University of Minnesota; daver004@umn.edu, Lynn Blewett, University of Minnesota; blewe001@umn.edu, Kathleen Thiede Call, University of Minnesota; callx001@umn.edu, and Holly Rodin, University of Minnesota; rodi0016@umn.edu**

Missing data in the form of item or entire survey module non-response is a common problem in survey research (Groves, Dillman, Eltinge, and Little 2001). Approximately 10 percent of the Current Population Survey (CPS) sample refuses to take the demographic supplement. These "full supplement" refusals have the variable values for the entire demographic supplement imputed. Although properly specified imputation can alter basic distributional summary statistics (e.g., means, rates and variances) from the statistics calculated using complete

cases only, it should not transform the relationships among variables. In other words, imputation should not create significant correlations between variables that were not there before the imputation, nor should it reduce the magnitude of significant correlations between variables that were there prior to imputation. With this in mind, there are two questions we attempt to answer in this analysis: 1) Is there a difference between the imputed cases and the non-imputed cases with respect to health insurance coverage? 2) Does hot deck imputation create a significant bias in health insurance coverage estimates? In the 2003 Current Population Survey's Demographic Supplement, 59.7 percent of 18-64 year old adults have commercial health insurance coverage if they have the full supplement imputed. However, 72.4 percent of the non-full supplement imputations have commercial health insurance coverage. Furthermore, full supplement imputations have a 26.7 percent uninsurance rate while all other 18-64 year old adults in the CPS have an uninsurance rate of 14.6 percent. We examine the relationships among key correlates to see whether this difference is due to the characteristics of the full supplement imputations or is due to the Census Bureau's missing data imputation routines introducing/reducing relationships among variables.

***Analysis of Behavioral Risk Factor Surveillance System Demographic Bias and Effect of Weighting on Risk Factor Estimates*, Herbert Stackhouse, Centers for Disease Control and Prevention; hls4@cdc.gov, and Pranesh Chowdhury, BCA; pchowdhury@cdc.gov**
The objective of the BRFSS is to collect uniform, state specific data on preventive health practices and risk behaviors that are linked to chronic diseases, injuries, and preventable infectious diseases in the adult population. Data are collected from a random sample of adults (one per household) through a telephone survey. This study evaluates demographic non-response biases in the BRFSS, assesses the effect of the BRFSS design weights and post-stratification on the survey's representativeness, and examines the results of different weighting strategies on calculated estimates of risk factors and prevalences. The analysis examines the effect of the current design weights and post-stratification in the BRFSS for the years 2000-2003 to evaluate the effect of the weights on the estimates produced and their variance. The project also evaluates different potential post-stratification methodologies and their effect on the same estimates. Data will be examined at the state level for trends over the four survey years. The results of this analysis are being compared to external population and public health data sources as a measure of accuracy of the survey and appropriateness of the current and proposed weights. The risk factors analyzed include smoking, overweight/obesity, health insurance, health status, physical activity, drinking, and nutrition. Demographic characteristics included in the analysis are race/ethnicity, age, sex, educational attainment, income, and employment. This analysis is invaluable for determining the success of the data collection operations of the BRFSS and for evaluating and potentially reducing demographic bias in the survey.

***Predicting Exit Poll Interviewer Response Rates*, Clint Stevenson, Brigham Young University; approximatelynormal@yahoo.com, Daniel Magleby, Brigham Young University; dm227@email.byu.edu, Quin Monson, Brigham Young Universiry, Kelly Patterson, Brigham Young Universiry, and Howard Christensen, Brigham Young Universiry**
In recent elections, the national exit polls have come under intense scrutiny. In 2004, the exit polling conducted by the National Election Pool encountered problems that may have been due to nonresponse error. Speculation has focused on the possibility that Kerry supporters were more likely to participate than Bush supporters. Previous studies on exit poll nonresponse have scrutinized response rates at the precinct level. We use a different approach and examine nonresponse at an interview level using a unique data set of every attempted interview in a 2004 statewide exit poll in Utah. We utilize data about interviewers and potential respondents that participated in the 2004 KBYU/Utah Colleges Exit Poll, conducted statewide every two years in Utah beginning in 1982. The Utah exit poll uses sampling methodology similar to national exit polls and has an excellent track record. To examine the question of interviewer characteristics on a successful response, we gathered data on interviewer characteristics, work experience, and surveys distributed on Election Day. We employ logistic regression to determine factors that contribute to success or failure of each interview attempt. We find that interviewer characteristics such as previous work experience (e.g. retail sales, door-to-door sales, waiter/waitress experience) have the strongest positive relationship with a successful survey response. We conclude the paper with a discussion of how to strategically select and position interviewers to maximize response rates. Our approach offers suggestions to improve exit poll survey quality by providing a model that will predict response success based upon interviewer characteristics.

***Public Perceptions of Poll Results - Do the Details Matter?*, Monika McDermott, University of Connecticut; monika.mcdermott@uconn.edu, and Samuel Best, University of Connecticut; sam.best@uconn.edu**
How does the public judge public opinion poll results? In election years such as this one in the United States the public is provided with a nearly daily supply of poll results about the election and the state of the country - results which are at least occasionally in disagreement. For this reason, understanding if and how the public discriminates among poll results is important to understanding what effect, if any, these results have on public attitudes and the election itself. While a few previous studies have asked survey respondents directly about their views of poll methodology, the answers cannot shed light on any causal relationships between people's perceptions of poll method and their resulting opinions about the polls. In contrast, this study uses and experimental design in both national and state level surveys in the United States to examine aspects of surveys that may affect public perceptions of their accuracy - including survey sponsor and sample size. By providing survey respondents with reported poll results and experimentally manipulating the details, this study will determine, for example, whether people put any more or less faith in the accuracy of results collected by universities, such as Harvard, than those collected by major media outlets, such as CBS News, or commercial organizations like Gallup. In addition, it will examine what effect people's resulting perceptions of poll accuracy have on their own attitudes. For example, are results perceived as accurate more likely to affect individuals' opinions than results perceived as inaccurate?

***Who Deliberates in the Public Square and Does It Matter?"*, Robert Goidel, Louisiana State University; kgoidel@lsu.edu, Charles Zewe, Louisiana State University; czewe@cox.net, and Steven Procopio, Louisiana State University Public Policy Research Lab; sproco1@lsu.edu**
Survey research has been frequently criticized for reflecting hastily drawn and poorly formed opinions. James Fishkin, for example, has argued that - measured in this way - public opinion surveys miss the normatively and substantively important deliberative component of public opinion formation. In this paper, we consider two questions related to deliberative public opinion. First, who shows up for deliberative opinion forums? And second, what difference does their participation make in terms of general attitudes toward the political process? To answer these questions we make use of unique set of data collected as part of a series of monthly programs, Louisiana Public Square, which aired on public television from July to December 2004. These programs covered a range of issues (e.g., public education, roads and highways, health care, religion and public life) and included participants selected using random digit dialing. These participants

learned about the issues, discussed the issues among themselves, and then asked questions of leading state policy-makers within each issue area. Data were collected both before and after the program, allowing us to: (1) Compare attitudinal and demographic differences among participants (pre-show and post-show surveys) and nonparticipants (pre-show survey only); and (2) Analyze attitude change among participants particularly with respect to opinions regarding the responsiveness of the political process to public concerns and the value of deliberation to democracy. We find that participants are often more intense about the issue in consideration but more cynical about state government's ability to handle the problem. After the program, participant perceptions of government responsiveness and trust increased.

***A Pre-validation Survey of Voter Turnout*, Michael McDonald, George Mason University; mmcdon@gmu.edu**
Survey researchers have long been concerned that the validity of survey measurement is questioned when percentages of responses to survey questions systematically do not match observable aggregate percentages. A frequent indicator of the reliability of political surveys is the seeming over-report of a respondent's act of voting. Comparison of turnout rates of respondents to aggregate election results shows that more people report voting on election surveys than the aggregate turnout rates. The validity of respondent self-reporting may also be validated by examination of actual individual voting records. The earliest surveys of political behavior found that more respondents report voting than official voting records indicate, a fact re-established though later validation effort by the National Election Survey. In an interesting twist to validation efforts, I investigate vote over-report in the 2004 election using a sample of registered voters than includes information about their vote in the 2004 election. This "pre-validation" design allows for a gentle probe into the reasons why some respondents misreport their vote, and helps us understand how to mitigate report bias.

***Predicting Sampled Respondents\' Likelihood to Cooperate in a Mail Survey: Part III*, A.T. Burks, Nielsen Media Research, Inc.; anh.thu.burks@nielsenmedia.com, Paul J. Lavrakas, Nielsen Media Research; paul.lavrakas@nielsenmedia.com, and Mildred Bennett, Nielsen Media Research, Inc; mildred.bennett@nielsenmedia.com**
There are many factors that affect the decisions of respondents to cooperate when they are sampled for a survey. To the extent a researcher can accurately predict a given respondent's likelihood to cooperate, i.e., predict the respondent's individual "response propensity" (cf. Groves et al, 2004), then in theory, the researcher could target differential approaches to recruiting each respondent rather than merely utilizing a "one size fits all" approach. Our previous findings from a large mail survey (n=10,000) conducted in January 2004 (cf. Lavrakas et al, 2004) and a larger national mail survey (n=69,000) conducted in May 2004 (cf. Burks, et al, 2004) showed that cooperation in the mail survey could be predicted at a level significantly better than chance by using data about the (1) mailed respondent/household gathered in a prior telephone survey, (2) variables associated with the calling history of the respondent's telephone number, (3) interviewer ratings about the respondent, and (4) Census variables at the level of the local zipcode where the respondent lived. Our next step in this line of research is to use Census data at the block group level to determine what gains can be achieved in predicting response propensity to a mail survey among a national sample of 18-34 year olds (n=1,000). Our previous two studies showed an ability to predict cooperation in a mail survey among this age cohort with approximately 60% accuracy. Bringing the level of aggregation for the geographic variables in our model to a smaller level (block group) is more precise in nature, and is expected to contribute to improving the predictive power of our statistical model. The implications of these findings will be discussed as they relate to the differential deployment of finite resources for incenting households and strategies to increase response rates.

***Partners in Preservation of Opinion Data: Developments of a Library of Congress Initiative*, Lois Timms-Ferrara, Roper Center; lois@ropercenter.uconn.edu, and Marc Maynard, Roper Center; mark@ropercenter.uconn.edu**
Congress has mandated the Library of Congress to create a set of public-private partnerships collectively known as the National Digital Information Infrastructure and Preservation Program. The purpose of these partnerships is to ensure "the long-term preservation of digital content and to capture current digital content that is at risk of disappearing." One result of this initiative is the formation of D-PASS, a partnership of the leading U.S. social science data archives to identify, acquire and preserve potentially "at-risk" data resources. These partners are the Roper Center (University of Connecticut), the Inter-university Consortium for Political and Social Research (University of Michigan), the Murray Research Center of the Radcliffe Institute for Advanced Study (Harvard University), the Odum Institute for Research in Social Science (University of North Carolina), the Virtual Data Center (Harvard-MIT Data Center and Harvard University Library), and the National Archives and Records Administration (NARA). The Library of Congress is a partner in all collaborations. This poster session will focus on the role of the Roper Center, the largest archives of opinion data, in this endeavor. The Center will maintain a close collaboration with the National Opinion Research Center to fill in the gaps in its collection of NORC surveys (1941-present) and with the National Archives to identify and acquire United States Information Agency surveys (1950s-present) not currently preserved at Roper or NARA. A third component of the Center's objective is to determine the availability for archiving those surveys that have questions in the iPOLL database, but are not represented by datasets in the archives. Although this grant would be for a period of only three years, the enabling legislation envisions a permanent commitment by the Library of Congress to digital preservation through public-private partnerships. The partnership may also become a foundation for inter-archives agreements with far-reaching consequences, such as a commitment to metadata standards in the private and public world of social science data. The AAPOR community is in a unique position to support this project by helping to identify "at-risk" studies and provide useful information in tracing the existence of relevant study documentation and data files. The benefit to AAPOR members from this venture, as in all data preservation efforts, is the availability of rich and varied sources of opinion data for better comprehension of the public in the aggregate.

***Sound Bytes: Capturing Audio in Survey Interviews*, Sue Ellen Hansen, University of Michigan; sehansen@umich.edu, Maria Krysan, University of Illinois at Chicago; krysan@uic.edu, and Mick Couper, University of Michigan; mcouper@umich.edu**
The use of multimedia-capable laptops for computer-assisted interviewing (CAPI and CASI) is changing the way survey interviews are conducted. Audio output is now quite common in surveys, being used for audio-CASI. But interest in audio input-the automatic capture of digital sound files from the interview-has been growing in recent years. Digital recording shows great promise both for methodological and for substantive goals. Substantively, they could supplement the response to open-ended questions, giving the analyst further insight not only into what the respondent said, but how they said it. They could also be used to explore nuances in the response to closed-ended questions, such as clarifications or qualifications that are not routinely captured in structured interviews. They could be used to evaluate interviewer performance in administering survey questions. In other words, they can vastly extend our insight in the dynamics of the survey interview and help us better understand the responses provided. In this paper we report on the use of a digital recording tool in a survey on racial attitudes in the Detroit area (n=734). The tool, embedded in the CAPI survey authored in Blaise, turned on or off automatically at

designated questions, recording the interaction in digital files. We will report on the success of this approach in terms of the rates of consent by respondents and the audibility of the interviewer and respondent. We used the recording of responses of closed-ended questions to explore the hesitations, clarifications or qualifications on the part of the respondent that are normally not recorded in CAPI surveys, and to examine possible interviewer effects in administering the questions. We also recorded several open-ended questions to evaluate the quality and completeness of the responses typed by the interviewers in the CAPI instrument.

***The Role of Economics, Religion and National Security in the 2004 Presidential Vote*, Leonie Huddy, SUNY; lhuddy@notes.cc.sunysb.edu, Stanley Feldman, SUNY; Stanley.Feldman@sunysb.edu, and Sarah Dutton, CBS News; sld@cbsnews.com**

Pre-election polls prior to the 2004 presidential election showed that voters held strong and highly partisan views on key domestic and international issues -- among them the economy, the war in Iraq and national security. In many cases, likely voters divided almost evenly on some of the most pressing issues currently facing the country. Yet, the impact of these issues on the 2004 presidential election remains unclear. Recent worsening economic conditions, especially the job losses of the past few years, suggest a heightened role for economic issues in presidential vote choice. But at the same time, the 2004 election fell under the powerful influence of national security issues. Conventional wisdom suggests that voters are reluctant to change presidents during wartime, yet the role of war and terrorism on the 2004 vote choice remains far from clear. How did the threat of terrorist attacks in the U.S. affect voters, especially those living in the areas that have been attacked before? And with an incumbent president who is a born-again Christian running for re-election, the role of religion may have been a more important factor for some voters in this election. This paper will explore in detail the role these issues played in voters' presidential choice. The contrast between economic issues and symbolic issues, such as religion, will be investigated. And the impact of national security concerns will be examined against both economic and symbolic concerns. The investigation will be based primarily on data from the 2004 exit poll, with supplemental data from CBS News Polls where applicable. In additional to standard analyses, multivariate analyses of specific issue questions will also be included. To the extent possible, comparisons will be made as to how these (or similar) issues affected previous elections.

CROSS-NATIONAL RESEARCH I

***How do citizens respond to photographs of politicians?: a cross-national analysis*, Robert Eisinger, Lewis & Clark College; eisinger@lclark.edu, and Thomas Petersen, Institut für Demoskopie Allensbach; tpetersen@ifd-allensbach.de**

While research about media influence continues to pervade the public opinion literature, there remains a dearth of scholarship about the effects of visual images in newspapers. Specifically, photographs of politicians appear on a daily basis, in virtually all newspapers. Yet there is no systematic study of how if at all, these photographs influence public opinion, how citizens react to these photographs, or if there is any pattern of what is portrayed in the photos. This paper questions if popular politicians portrayed more positively than unpopular politicians? We first locate photographs and poll data of political competitors in Germany and the United States respectively. We then ask respondents who are unfamiliar with the photographed politicians to rate the quality of the newspaper images. That is, Germans assess photos of U.S. politicians, and Americans rate photos of German politicians. We hypothesize that politicians who are ahead in the polls are not portrayed with more flattering photographs than their underdog competitors, arguing instead that 'flattering' photos are as likely to be printed of politicians who are behind in the polls, as those who are leading. The cross-national design of this study affords us to generalize beyond one political election, two political candidates or one respective region. Rather, we seek to theorize that photo editors make critical decisions about what to print, and that these images have genuine effects on how citizens perceive politicians. We simultaneously argue, however, that these images are not correlated with poll ratings. Popular and unpopular politicians are just as likely to be viewed in a flattering (or unflattering) way.

***Daily Contact as A Proxy Measure of Personal Networks: Data from 45 Surveys in Nine Societies*, Yang-chih Fu, Academia Sinica; fuyc@sinica.edu.tw**

It has been difficult to measure personal networks in surveys, largely because network boundaries constantly change and often remain vague. Most preexisting network proxies are sophisticated and tedious to administer. Under practical constraints, it usually becomes costly and intimidating to ask about personal networks even in well-administered face-to-face surveys. As a result, many surveys have been unable to come up with succinct items of networks, even though such information may greatly help one understand many important topics. To bridge the gap, this paper presents a simple and straightforward network proxy -- daily contact, which can be used as a stand-alone survey item and extended to further explore derivative issues. The item asks for an approximate number of persons the respondent makes contact with in a typical day. The answer includes six ordinal categories. Since it was first implemented in the fieldwork in 1992, the proxy has been modified to stipulate what kinds of situations qualify as "contacts", the response categories have been adjusted to accommodate various situations, and the answers have approximated normal distributions in most surveys. Over the past decade, this "daily contact" item has been incorporated into 45 surveys in nine countries/societies (including China, Great Britain, Hong Kong, Japan, Korea, Philippines, Poland, Taiwan, and U.S.A.) The paper will present and compare findings from 21 face-to-face surveys, 18 telephone surveys, and 6 self-administered surveys, with a combined sample size of over 60,000. Measured against widely used network proxies, not only does such an index of daily contacts correspond closely with the network size that the name generator produces, but it also distinguishes well the network span measured by the position generator, as well as the extent of social participation.

***Cross-National Contact Strategies and Effects*, Oliver Lipps, Mannheim Research Institute; lipps@mea.uni-mannheim.de, and Grant Benson, University of Michigan Institute for Social Research; gdbenson@isr.umich.edu**

There are a considerable amount of data on optimal contact strategies in the United States and select other nations, but very little truly comparable cross-national data. These data may be used to improve response rates, reduce survey costs, and inform whether there is potential respondent selection bias. The Survey of Health, Ageing, and Retirement in Europe, conducted simultaneously in 11 West European countries between April and October 2004, employed Computer-Assisted Personal Interviewing and a computer-assisted sample management system (SMS). Interviewers were provided a standardized training in their native language and were required to report the

details of each contact attempt into the SMS, including date, time of day and whether or not contact was made. This presentation will examine the impact of contact attempt history on non-response and respondent selection bias using logistic models. We will analyze this both in terms of (expected) household composition and key indicators central to the Survey of Health, Ageing, and Retirement.

***Using Anchoring Vignettes to Assess Cross Cultural Comparability of Consumer Ratings*, Patricia Gallagher, Center for Survey Research; patricia.gallagher@umb.edu**
There is evidence that Consumer Assessment of Health Plans (CAHPS) ratings vary by racial/ethnic group. It is not known whether this variation is attributable to differences in: expectations for the health care system; health care experiences; or the ways that respondents use offered response options. To learn more about the causes of these subgroup variations, we fielded a survey designed to test an approach to enhancing the cross-cultural comparability of CAHPS instruments that follows a model proposed by Gary King, Christopher Murray, et al. (2004)*. The technique involves the use of "anchoring" vignettes that present hypothetical examples of health care at various levels of efficacy. Respondents are asked to assess the care presented in the hypothetical scenarios. They are also asked to use the same response choices presented for the vignette assessments to self-assess the domain of care described in the vignettes. The responses to the vignette items can be used to identify differential item functioning (DIF) at the sub group level. The innovators of this method offer both parametric and nonparmetric approaches to analysis. We conducted a mail survey of parallel samples of Latino, African American, and non-Hispanic white Medicaid members. Sampled members each received a Canadian-style dual-language (English/Spanish) instrument that included core CAHPS items and the anchoring vignettes. Standard mail survey protocols were followed. A total of 368 questionnaires were completed, with more than 100 returns from each subgroup (RR=26% (RR1, AAPOR 2004)). Analysis of the data collected will allow the identification and quantification of differences in the ways that different cultural groups use CAHPS response scales. In addition, we will have the tools to calibrate self-assessments to adjust for interpersonal differences in the use of the response scales. *Enhancing the Validity and Cross-cultural Comparability of Measurement in Survey Research

SOCIAL CHANGE

***Changes in Attitudes Toward American Institutions and Occupations*, Richard Seltzer, Howard University; rseltzer@howard.edu, and Rhea Roper, Howard University; trinigyl@hotmail.com**
Others have analyzed changes in trust in American institutions by analyzing data from the NORC-GSS or the ANES. In this paper we extend this analysis by examining: 1. the NORC-GSS on American institutions-1973-present; 2. ANES on American institutions-1964-present; 3. Gallup on honesty in occupations -1976-present; 4. Gallup on confidence in American institutions 1966-present; 5. Harris on confidence in American institutions-1971-present; and 6. four different survey firms on attitudes toward industries-1969-present. Over 3,000 survey questions are analyzed and allow us to see how attitudes change over time, how institutions compare with one another, and whether results differ according to survey firm.

***Are Older Adults Losing Social Capital? 2004 and 1996 Compared*, Thomas Guterbock, University of Virginia; tmg1p@virginia.edu, and Edna Renee Macbeth, Univeristy of Virginia; erm2z@virginia.edu**
In _Bowling Alone_, Robert Putnam (2000) argued that social capital is in decline in the United States. He points to the aging of the "long civic generation" that matured during the Great Depression, and notes that the Baby Boom generation has been less involved in civic and social pursuits. In this view, there is an alarming decline in social capital that is attributable in part to cohort replacement. This paper tests (1) whether social capital is really in decline among older adults and (2) whether the Baby Boomers are laggard in their civic engagement. The 1996 AARP Survey of Civic Involvement is an RDD telephone survey of 1,500 adults, half of whom were age 50 and older. The 2004 AARP Beyond 50.05 Survey is an RDD telephone survey of 1,005 adults age 50 and older (including older cohorts among the Boomers), plus African-American and Hispanic oversamples. The two surveys cover similar topics and many items are replicated. When the data are compared, we find no significant drop in measures of overall social involvement, community attachment, or other civic engagement indicators for adults 50+. And when the aging Boomers are compared with those of like age in 1996, we see no clear trend of decline as a result of cohort replacement. Overall, age differences in civic engagement are not large among older adults, and the variations seem to be more closely tied to life-course changes than to generational differences. These results are good news for the health of America's social fabric. The results are also good news for older adults, because the 2004 survey shows strong evidence that continued civic engagement contributes strongly to successful aging.

***Deliberative Attitude Change: Changing Your Mind Within an Interview*, Patrick Fournier, Université de Montréal; patrick.fournier@umontreal.ca, André Blais, Universite de Montreal; andre.blais@umontreal.ca, Joanna Everitt, University of New Brunswick; jeveritt@usbsj.ca, Elisabeth Gidengil, McGill University; Elisabeth.Gidengil@mcgill.ca, and Neil Nevitte, University of Toronto; nnevitte@chass.utoronto.ca**
Recent research has uncovered evidence of a singular phenomenon: survey-induced deliberation. Our previous analyses suggest that that ordinary features of a questionnaire - the content and order of survey items - can influence the quality of survey responses. Using a split-sample experiment on the location of the vote intentions questions, we find that vote intentions captured near the end of a long and balanced interview exhibit less non-response and more enlightenment than those revealed at the start of the interview. Respondents who had the opportunity to overview their attitudes regarding vote choice before answering were more likely to express vote intentions, and to articulate vote intentions that are more closely related to their underlying interests/values and their eventual behaviour. This research warrants validation and further exploration. We propose to tackle each of the following questions: - Does the location of the vote intention questions affect non-response? - Does change in vote choice truly occur within the span of a questionnaire? - Are late-survey vote intentions better predictors of actual vote choice? - Are late-survey vote intentions more inline with people's interests and values? - Which individuals are changing their mind during the interview? The least informed? The ambivalent? To answer these questions, we employ a unique feature of the 2004 Canadian Election Study: the campaign vote intention questions were asked twice to every respondent, first early in the survey, and then at the very end. This design allows us to verify and supplement the previous findings based on a split-sample. We can ascertain how many people switch electoral preferences during the course of the interview, who these people are, and whether they move to adopt more enlightened opinions. These inquiries are important for understanding the quality of survey results, of citizen decision-making, and of the democratic process.

***Deliberation and Its Discontents: A Comparison of Survey and Deliberative Discussion Results, 1987-2003*, Frank Rusciano, Rider University; rusciano@rider.edu**

In recent years, a series of National Issues Forums, or deliberative discussions, around the United States have been conducted about public issues. Critics of this approach tend to argue that for a nation the size of the United States, deliberative discussions can only reach a miniscule portion of the adult population. To study the dynamics of mass political discourse relevant to politics, they argue, one must rely on the more "scientific" methods like the random public opinion survey. This approach also has its critics. Zahler (1992) notes how the survey method is itself an artificial "intrusion" into citizens' lives that forces them to formulate and express opinions on issues upon which they may not have well-formed attitudes. Ginsberg argues further that public opinion surveys have changed from a means by which to measure public opinion to a means by which to manage public opinion (1985). This paper responds by comparing the two approaches. It examines results from a series of deliberative discussions about the United States' role in the world, held at four-year intervals from 1987 to 2002. It notes the changes in discussants' opinions from pre- and post-discussion surveys. It then compares these changes to the differences in opinion on similar issues between citizens and opinion leaders in surveys done in the same years by Chicago Council on Foreign Relations. The results show that the deliberative discussants' responses tend to move significantly towards the opinion leaders' responses, and away from the citizens' responses, on comparable questions from the pre- to post-discussion surveys. The report concludes by assessing the meaning of these findings for future deliberative discussions; it proposes that the deliberative process both empowers citizens and makes them more knowledgeable about the issues, even on such a "complex" topic as foreign affairs.

INCREASING RESPONSE RATES

***(Inter) Net Gain? Experiments to increase web based response*, Steve Tourkin, Census; steven.c.tourkin@census.gov, Shawna Cox, U.S. Census Bureau; shawna.m.cox@census.gov, Randy Parmer, Census; Randall.j.Parmer@census.gov, and Andrew Zukerberg, U.S. Census Bureau; andrew.l.zukerberg@census.gov**

This paper focuses on using monetary incentives to increase overall response and Internet response when mail and Internet choices are offered. Previous research has indicated that offering an Internet option does not increase total response rates for mail out questionnaires, but that methods can increase Internet response over mail response. Internet questionnaires can reduce non-sampling error in surveys by automating skip patterns, as well as providing edit and range checks. Additionally, Internet surveys provide data processing advantages including faster turn around times and lower processing cost. Given the advantages of an Internet administration, Census and NCES wanted to encourage Internet response for the 2004-05 Teacher Follow-up Survey (TFS). At the same time, incentives were offered to increase response. An experiment designed to assess the impact of incentives on overall response and on Internet response was embedded into administration of the TFS. The experiment looked at 3 different Internet treatments: 1) initially providing only the internet option, 2) providing the Internet option initially and informing respondents that a paper questionnaire is forthcoming, and 3) no Internet option. Half of each of these groups was provided a $10 debit card incentive at the time of first contact. This paper will compare the relative impact of each method on response, and make recommendations for other surveys interested in encouraging Internet response and/or using pre-paid incentives.

***Does Type of Pre-Notification Affect Web Survey Response Rates?*, Michele Harmon, Westat; micheleharmon@westat.com, Elizabeth Westin, Westat; elizabethwestin@westat.com, and Kerry Levin, Westat; kerrylevin@westat.com**

Much is known about the effectiveness of pre-notification for mail and telephone surveys (Dillman, 2000). Some recent research has also demonstrated the effectiveness of pre-notification for web surveys (Kaplowitz et al., 2004). However, there is a paucity of research about the effect of different types of pre-notification on web survey response rates. This paper reports on a study that examines web survey response rates by type of pre-notification. The study sample was made up of 1,439 respondents that had ever applied for a grant from a Federal granting agency using a specific type of application process. As part of an evaluation, respondents were asked to complete a survey on their experiences with the grant application process. Data from the sample frame show that some of the grant applicants had received a grant award and some had not. The current study employed an experimental design in which the respondents were randomly assigned to one of three groups that varied by type of pre-notification. The three types of pre-notification included: 1) email from the sponsoring agency with letter attachment, 2) email from the research organization with letter attachment, and 3) paper letter from sponsoring agency. The electronic pre-notification letter attachments (in .pdf format) and the mailed paper pre-notification letter were identical and from a division director at the Federal granting agency. All remaining contacts with respondents were from the research organization that conducted the evaluation. This paper will provide results on the overall web survey response rates by type of pre-notification. Since it is thought that applicants who never received a grant award would be less likely to respond to the survey than those who had received an award, the impact of the type of pre-notification on response rates is also broken out by these two respondent groups.

***Did you get the message? Using e-mail and SMS for prenotification in Web surveys*, Wolfgang Neubarth, ZUMA, Center for Survey Research and Methodology; neubarth@zuma-mannheim.de, Michael Bosnjak, University of Mannheim; bosnjak@tnt.psychologie.uni-mannheim.de, Wolfgang Bandilla, ZUMA Mannheim; bandilla@zuma-mannheim.de, Mick Couper, University of Michigan; mcouper@umich.edu, and Lars Kaczmirek, ZUMA, Centre for Survey Research and Methodology; kaczmirek@zuma-mannheim.de**

Ample empirical evidence suggests that prenotification is an important method to increase response rates in mail surveys. However, for Web-based surveys, the results on prenotification appear to be mixed. One explanation may be the lack of awareness of the initial e-mail contacts. For instance, automatic spam filters and unidentified senders of the prenotices could lower the chance of the e-mail being noticed. To increase awareness when contacting participants for Web-based surveys, it might be more effective to send the invitations directly to the respondents' cellular phones. This method seems most promising with younger persons in Germany, because this segment has near-universal cellular phone coverage. To compare the effectiveness of different prenotice procedures, we experimentally varied the initial contact mode in a fully crossed two-factorial design with (1) three different prenotice conditions and (2) two invitation and reminder conditions. University freshmen were recruited for a Web-based access panel, by completing a short paper and pencil questionnaire in a classroom setting consisting of basic demographics, psychographics, and questions about lifestyles. A total of 560 students who revealed both their e-mail address and mobile phone number and agreed to participate were randomly assigned to one of the following experimental conditions: On the prenotice factor, one group received no prenotice at all, a second was contacted by a short text message (SMS) on their mobile phone, and a third group was contacted by e-mail, both announcing the prospective survey. For the survey invitation and reminder,

respondents were randomly contacted either by e-mail or by mobile text messaging. Our paper will explore the extent to which these experimental manipulations influenced the sample composition, different nonresponse patterns, and the responses to the substantive questions. Practical implications for contacting participants of Web-based surveys will be discussed.

***Analysis of Break-off Patterns in Web Surveys*, Syed Ahsan, NORC; ahsan-syed@norc.uchicago.edu, and Ronald Broach, National Opinion Research Center; broach-ronald@norc.uchicago.edu**
In order to gain competitive advantage through cutting edge technology, the Web survey is considered as a strategic tool in the data collection process. It offers great potential for collecting survey data in the most cost effective way, and the speed of collecting data makes it a viable option for surveys with a short field period. However, the research community has raised a number of interesting questions related to the successful implementation of web-based interview, including whether the web mode allows effective contact and follow-up procedures, and the impact of sensitive questions in break-off patterns, among other concerns. This paper will evaluate break-off patterns of respondents in the Survey of Doctorate Recipients (SDR) project from various perspectives. SDR project is sponsored by the National Science Foundation and the National Institute of Health, and the study collects data from science and engineering doctorate recipients who earned their degrees from institutions within the United States. Analysis will include an evaluation of break-off duration of complete cases in order to develop effective follow-up procedures. We will also review incomplete cases to determine if there is any specific break-off pattern in any particular section of the questionnaire. Selected variables will be used in the analysis to measure behavior across demographic profiles. Finally, the paper will introduce the concept of minimizing potential break-off through the intelligent use of user-interface and questionnaire programming techniques.

INTERVIEWER EFFECTS

***The effects of interviewer gender and race on sensitive health questions*, Gregg VanRyzin, Baruch College / CUNY; gregg_vanryzin@baruch.cuny.edu, Martin Frankel, Abt Associates and Baruch College, CUNY; Martin_Frankel@abtassoc.com, Farzad Mostashari, New York City Department of Health; fmostash@health.nyc.gov, and Stephen Immerwahr, Baruch College / CUNY; stephen_immerwahr@baruch.cuny.edu**
This paper examines how the interaction of interviewer and respondent characteristics (gender and race-ethnicity) influences the self-reporting of sensitive health behaviors (alcohol use, sexual activity, weight, and mental health). The decisions of public health officials and policy makers often depend on self-reported health behaviors gathered through telephone surveys of the general population. Yet it is known that interviewers, and more specifically the interaction of interviewers and respondents, can influence the responses gathered in such surveys. Although there is a large literature on these interviewer effects, the present study is unique in several respects. First, it focuses on health behaviors, rather than political or social attitudes toward race or gender issues, which have been the focus of much prior research. Second, the study focuses on New York City, one of the most diverse populations in the world, and thus includes fairly large proportions of white, black, Hispanic, and Asian respondents. The pool of interviewers was similarly quite diverse. Third, the sample is large enough (about 7000 suitable interviews in all) to allow for sufficient statistical power to detect various effects. Results suggest that, overall, interviewer gender and race-ethnicity have only minor and limited effects on self-reporting of sensitive health behaviors. Gender of interviewer has little effect on self-reported drinking, sexual activity, and obesity (although female respondents seem to report slightly less obesity to female interviewers). Gender does appear to influence self-reported mental health, however, with both male and female respondents reporting more mental health symptoms to female interviewers. The effects of the race-ethnicity of interviewer are few but interesting. For example, Hispanic respondents report more drinking to Hispanic and (to a lesser extent) black interviewers, while Asian respondents report more drinking to Asian interviewers.

***Interviewer effects in a RDD telephone pre-election poll in Minneapolis 2001. An analysis of the effects of interviewer race and gender*, Mario Callegaro, University of Nebraska, Lincoln. Program in survey research and methodology; mca@unlserve.unl.edu, Femke De Keulenaer, University of Antwerp; femke.dekeulenaer@ua.ac.be, Jon Krosnick, Stanford University; krosnick@stanford.edu, and Robert Daves, Star Tribune; daves@startribune.com**
The present paper analyzes the results of a RDD telephone pre-election poll conducted in Minneapolis a few days before the mayoral elections in November 2001. The candidates for mayor were the two-term mayor S. Belton, an African American woman, and the incumbent R. T. Rybak, a White man. This unique combination of race and gender of the candidates gave us the opportunity to test for both race and gender of interviewer effects. The dependent variable is measured with the question: "If the general election for Minneapolis mayor was held today, would you vote for S. S. Belton or R.T. Rybak?" and has three categories: the intention of voting for Belton, Rybak, or saying "don't know". A preliminary analysis showed an interaction effect between race of the interviewer and gender of the interviewer on voting intentions of the respondents, controlling for race and gender of the respondent. For example, we found that the odds for voting for Belton and not Rybak were the highest when interviewed by an African American woman. We also found that when interviewed by a White man the odds for making a choice were higher than for giving a DK answer. Contrary to other studies, we did not found that interviewer effects are dependent on matching between the respondent and the interviewer. In addition, we could not fully explain the interaction effect between race and gender of the interviewer. In this paper, the results are discussed in more detail in light of current theories of interviewer race and gender effects. In the current analyses, we will also test if the effect of interviewer race and gender can be accounted for by differences in the political interest and political affiliation of the respondents and by differences between the interviewers in factors such as interviewer experience, education, and workload.

***Gender-of-Voice Effects in an ACASI Study of Same-Sex Behavior*, Kristine Fahrney, RTI International; fahrney@rti.org, Jennifer Uhrig, RTI International; uhrig@rti.org, and Richard Wolitski, CDC; RYW1@CDC.GOV**
Audio computer assisted self interviewing (audio-CASI or ACASI) has become an increasingly popular method for administering surveys of sex and other sensitive behaviors. One of the major advantages of ACASI is that it removes the requirement that respondents divulge sensitive behaviors directly to another person. As a result, ACASI may reduce the extent to which response accuracy is compromised by interviewer effects. However, the literature on Computers As Social Actors suggests that computer users ascribe human traits to computers; even subtle cues such as the gender of the computer-generated voice may cause users to react to the computer as they would to another person. Based on these findings, we could infer that there may be the same pattern of gender-of-interviewer effects when using ACASI techniques as there are in traditional face-to-face interviewing. If this were true, the choice of an ACASI voice may be just as important as

the choice of interviewing staff in traditional face-to-face interviewing. Although there have been a few studies of the effects of gender-of-voice in telephone-ACASI and IVR, there is no published research on gender-of-voice effects in ACASI studies. The present study examines whether the gender of the ACASI voice affects the reporting of socially undesirable HIV-risk behaviors in a sample of men who have sex with men. A sample of 405 respondents were randomly assigned to the male or female voice condition. Although voice effects were demonstrated in only a few variables, the male ACASI voice was found to elicit more socially desirable responses than the female ACASI voice in terms of risky sexual behaviors (male voice yielded more reports of receptive anal sex with HIV-status-unknown partners) and numbers of sexual partners (male voice yielded fewer reported partners).

***Estimating Measurement Error Using Interviewer Feedback*, Jim Wolf, Indiana University; jamwolf@iupui.edu, Andy Hutcherson, IU Public Opinion Lab; email, Brianne Peyton, IU Public Opinion Lab, Charles Hulen, IU Public Opinion Lab, and Sharon Sidenbender, IU Public Opinion Lab**

Biemer and Lyberg (2003) describe a fairly exhaustive taxonomy of total survey error as being comprised of sampling and non-sampling error. The latter is made up of five types of error: specification, frame, nonresponse, measurement and processing error. Measurement error is made up of a variety of error types including those resulting from respondent, interviewer and the instrument being used. This research will focus on ways to use readily available information to aid in minimizing future errors associated with interviewers and instrument which, in turn, should reduce respondent error and overall measurement error. Interviewers at the Indiana University Public Opinion Lab are routinely asked a set of questions at the end of an interview to record their assessment of the quality of the data collected. These interviewer assessment measures will be compared to data quality measures to determine how reliable the interviewer assessments are. It is assumed that high rates of "don't know" and "refuse" responses are indications of poor quality data for a given question. Counts for each interview will be made of the number of DK and REF responses for all items as well as subcategories of types of items such as demographic, sensitive, political, and the like. Regional, statewide and national surveys on a variety of topics will be used in the analysis. Emphasis of the analysis will be driven by the following questions: 1. How reliable are interviewer assessments? (Are ratings of interview quality strongly correlated with nonresponse?) 2. Which assessment questions are the most robust? 3. Are certain nonresponse patterns unique to interviewers? The ultimate goal of the study is to develop a quality control procedure that can be readily put in place and implemented during pilot testing and the early stages of data collection to minimize nonresponse and increase data quality.

***Will They Talk to Us? Survey Performance Rates of a Cell Phone Sample*, Anna Fleeman, Arbitron Inc.; anna.fleeman@arbitron.com**

In recent years, cell phone ownership has increased substantially - as much as 75% in just four years according to the Cellular Telecommunications and Internet Association. Further, the percentage of cellular users who have "cut the cord" - those who do not have residential landlines and rely only on cell phones - has grown considerably. Current research estimates that cell phone only users may represent as much as 6% of the total population; these users are more likely to be male and under age 35, a typically hard-to-reach demographic group. This is of great concern to survey methodologists, who rely on sample frames constructed with residential telephone exchanges. As a result of the potential under-representation, namely of young men, Arbitron conducted a study of cell phone users. The primary purpose of the study was to determine whether respondents, if reached on their cell phones, would agree to participate in the Radio Ratings. We mirrored our standard methodology that consists of "placing" a self-administered seven-day radio listening diary with each person (age 12+) in the home. We also sought to identify the respondents' demographic characteristics as well as usability and refusal rates. Seventy-five hundred numbers in dedicated cell phone exchanges were hand-dialed in accordance with Federal laws. An eight-attempt calling methodology was employed, voicemail messages were left, and first time refusals were tried again. Respondents were promised a five dollar cash incentive for participating. In the contacted households, the placement rate of Arbitron's radio listening diaries was surprisingly high, indicating that respondents reached on their cell phone will indeed participate in survey research. Reported results will include sample performance rates and findings from a questionnaire completed by the study's interviewers. All give insight to the possibility of successfully including cell phones in survey research that relies on a residential phone frame.

THE DO NOT CALL REGISTRY

***Has the National Do Not Call Registry Helped or Hurt Survey Participation Rates?*, Michael Link, Centers for Disease Control and Prevention; MLink@cdc.gov, Ali Mokdad, Centers for Disease Control and Prevention; AMokdad@cdc.gov, and Dale Kulp, Marketing Systems Group; DKulp@m-s-g.com**

At the time the National Do Not Call Registry (DNC Registry) went into effect on October 1, 2003 more than 50 million telephone numbers had been registered. Since that time, the number has grown and registration remains on-going. The impact of the DNC Registry on survey participation rate is, however, largely unknown. Some believe the registry could make it easier for survey researchers to distinguish themselves from telemarketers. Other, however, believe that a significant portion of DNC registrants may not make such distinctions and prefer instead to reduce all unsolicited calls from marketers and interviewers alike. A third school of thought is that the DNC Registry may simply reflect attitudes towards surveys which are already formed and hence have little real impact on survey participation. In order to better understand the impact of the DNC Registry after its first year of operation, case outcomes from nearly 3.5 million telephone numbers called between October 1, 2002 and September 30, 2004 as part of the Behavioral Risk Factor Surveillance System (BRFSS) were analyzed. Six measures of survey participation for all 50 states and the District of Columbia were examined: case resolution, screening success, cooperation rate, initial refusals, refusal conversions, and response rate. While significant change was found in 177 of the 306 state-based survey participation rates examined, there were few clear-cut patterns that emerged which would allow us to reasonably attribute the changes to the launching of the DNC Registry in October 2003.

***The Impact of the National Do Not Call Registry on Social Survey Response*, Z. Joan Wang, PhD, REDA International, Inc.; zhengw@redainternational.com, Elham-Eid Alldredge, Ph.D., REDA International, Inc.; alldredg@redainternational.com, Jian Zhu, M.A., REDA International, Inc.; Jianz@redainternational.com, Michelle Cantave, REDA International, Inc.; mcantave@redainternational.com, and N. Craig Gilmore, REDA International, Inc.; cgilmore@redainternational.com**

In response to the increasing public demand for privacy protection from telemarketers, the Federal Trade Commission (FTC) created the National Do Not Call (DNC) Registry. More than 60 million phone number have been registered since June 27, 2003. According to the amended Telemarketing Sales Rule (TSR), as of October 1, 2003, it is illegal for most telemarketers or sellers to call a number on the DNC

list. Although public opinion and survey companies or organizations are exempt from this regulation, many survey researchers are concerned with the impact of the DNC Registry on the public's response to social surveys. Using comparable longitudinal surveys that were conducted before and after the establishment of the DNC list, the proposed study examines the change in survey response rate, the responding patterns of people who are on the DNC list and those who are not, and the characteristics of the DNC and non-DNC respondents.

***The Do Not Call Registry: Friend or Foe? The Effect of Do Not Call Lists on Survey Response*, Berwood Yost, Center for Opinion Research; byost@fandm.edu, Jennifer Harding, Center for Opinion Research, Franklin & Marshall College; jharding@fandm.edu, Christina Abbott, Center for Opinion Research; cabbott@fandm.edu, and Angela Knittle, Center for Opinion Research; aknittle@fandm.edu**

With more than 64 million phone numbers on the National Do Not Call Registry, it is surprising to find little research addressing the Do Not Call List's impact on survey response rates. While many researchers initially suggested the Do Not Call List would negatively affect response rates, we speculated that it has little or no effect on survey response. In this study, we compared response rates as well as other measures of cooperation for households that are on the Do Not Call List versus those that are not. We also explored: characteristics of households on the Do Not Call List; public awareness of Do Not Call-exempt-groups; perceived effectiveness of the Do Not Call List in reducing telemarketing calls; whether households that now receive fewer calls are more responsive; and whether it is beneficial to educate respondents on the industry's exempt status during the course of the interviewing process. This research is part of a statewide survey with a representative sample of Pennsylvanian adults. Respondents who initially refused were re-contacted after one month by phone/mail to obtain a complete understanding of the issues involved.

***Listed Numbers and Do Not Call Registration in a National Telephone Survey*, Jim Lepkowski, University of Michigan; jimlep@umich.edu, Howard Schuman, University of Michigan; hschuman@umich.edu, Rui Wang, University of Michigan; wrui@isr.umich.edu, and Richard Curtin, University of Michigan; rcurtin@umich.edu**

In data collected in August-October, 2003, in the national University of Michigan Survey of Consumer Attitudes telephone survey, respondents were asked "Is the number for the telephone that you are now using listed in a telephone directory or is it not listed?" and "Is the number for the telephone that you are now using listed in a state or the national 'do not call' registry to stop telemarketing firms from calling you?". Listed self-reports vary by age, gender, race, education, income, and region. Registration on a do-not-call list also varies by all but region. There is no or little difference on substantive items such as economic attitude indicators between listed and unlisted or registered and not registered households. There is an unexpected relationship between self-reported listed status and do not call registration, though. Those reporting not being listed are less likely to be in a do not call registry. A potential explanation is found in the examination of the joint relationship of listed status and do not call registration across demographic variables. There are important differences in being unlisted but not registered by race, across age groups, and by education and income. Self-reported telephone directory listing is not accurate. Objective address information obtained primarily from telephone directories is available for three-quarters of the respondents. Objective information is inconsistent with self-reports in an appreciable fraction of cases. A large fraction of Black and Hispanic respondents reported an unlisted telephone number when in fact they had a listed number. This finding suggests that studies using self-reported listed status may over or understate differences between race and ethnicity groups. The inconsistency may be due to a misunderstanding of the survey question or to some other reason.

CHANGING POLITICAL MINDS

***Handling The Truth: An Evaluation of Voter Capability to Discern Political Advertisement Accuracy*, Chris Borick, Muhlenberg College; cborick@muhlenberg.edu**

As has become custom in presidential elections, television advertisements once again were omnipresent in the 2004 race for the White House. And as in the past there has been continued concern regarding the truthfulness and factual accuracy of the many ads seen by the American electorate. In fact, there has been a concerted effort on the part of organizations such as Annenberg Public Policy Center's Fact Check.Org to evaluate the accuracy of the campaigns advertisement claims. While such efforts have helped provide an independent evaluation of political ads, it remains unclear if voters who actually view campaign commercials are able to discern inaccuracies as they are being presented to them. In this study we seek to determine if voters themselves display skepticism for candidate claims that have been shown to be inaccurate or misleading by the evaluations from Fact Check.org. To accomplish this goal we designed a study in which representative samples of voters (approximately 800) were asked to view campaign ads via the Internet, while continuously rating the advertisements believability via movement of a mouse. The aggregate reactions of the respondents (i.e. upward slopes when increased believability is identified) are then matched against Fact Check's evaluations of accuracy. Through this process we find mixed evidence regarding the ability of voters to discern inaccuracies in advertisements. More specifically we find that individuals from parties opposing the candidate airing the ad are significantly more likely to identify inaccuracies correctly than voters who are unaffiliated with a party, or who are affiliated with the party of the candidate sponsoring the advertisement.

***The Swift Boat Ads in Comparative Context: An Empirical Examination of Advertisement Effectiveness*, Chris Borick, Muhlenberg College; cborick@muhlenberg.edu**

During the 2004 presidential campaign few events drew as much attention as the advertisements by a group known as Swift Vets for Truth. During August this group began airing a series of television spots that questioned the validity of John Kerry's record from the Vietnam War. These ads, while airing in only limited markets, began to dominate the cable news programs, and were thus aired to national audiences for an extended period of time. Most importantly, the appearance of these ads corresponded with a noticeable drop in Kerry's position in the polls, and led to his campaign putting resources into an effort to refute the allegations. While the airing of the ads and Kerry's drop in the polls are intuitively related, what evidence is there that the television spots actually caused the drop in support for Kerry's presidential bid? In this paper we seek to examine this question through the presentation of the results of a series of empirical tests run during the campaign. More specifically, during the campaign we conducted regular tests on samples of American voters who were asked to view and rate television ads in a pretest-posttest format. During the campaign samples of approximately 800 voters viewed advertisements in a multi-staged design. First, voters were asked to complete a survey instrument in which they identified their current preference in the presidential election. Second, they watched a campaign commercial three times, rating the ads continuously (through

automated devices) in terms of aspects such as interest and believability. Finally, the voters were asked to complete the survey instrument once again in order to gauge any differences from their original positions. From these tests, we find strong evidence of the effectiveness of the Swift Boat advertisements in relation to other ads aired throughout the campaigns.

***Changing Minds in the 2004 Election?*, Jeffrey Jones, The Gallup Poll; jeff_jones@gallup.com, and Joseph Carroll, The Gallup Poll; joe_carroll@gallup.com**
This paper examines data from Gallup's pre/post-2004 election panel survey to assess changes in voter preference from immediately before to shortly after the election. The data include re-interviews of over 1,100 participants in Gallup's final national pre-election poll. While the conventional wisdom suggests that the polarized nature of the election meant most had made up their minds early in the election process (indeed, most respondents say they made up their minds before the conventions started), panel surveys generally show a greater degree of change at the individual level than is found by comparing results of successive cross-sectional surveys. The panel design allows us to estimate the percentage of voters who changed their mind in the final days of the campaign. We can compare the estimates to our 1996 post-election survey, an election also involving an incumbent president running for re-election, that showed roughly 1 in 6 voters changing their minds in the final days of the campaign. The panel design also allows for an examination of short-term change in other key attitudes such as party identification and ratings of George W. Bush before and after the election.

***Using the Internet to Learn About Presidential Candidates and Issue Positions in the 2004 Presidential Primary and General Election Campaigns*, Kenneth Winneg, Annenberg Public Policy Center, University of Pennsylvania; kwinneg@asc.upenn.edu**
At the 57th Annual AAPOR Conference, we showed how the Internet was used as a source of campaign information during the Democratic primary campaign using data from the 2004 National Annenberg Election Survey. In this research, we will expand our analysis to evaluate Internet use during the general election campaign, which unofficially began in early March, 2004. This research examines two main areas. First, it looks at how the Internet functioned in 2004 as a place for citizens to participate in and exchange ideas about the presidential campaign. Second, we examine how this function changed from the primary, when we were looking mainly at Democratic voters, to the general election, where we can consider the entire voting age population for analysis. Again, we will use data from the 2004 National Annenberg Election Survey Rolling Cross Section to address these issues and answer the following additional questions: * How did voters use the Internet to learn about and participate in the presidential campaign? * How strong was the relationship between Internet use and learning? * What was the role of the Internet relative to other means of campaign participation? * Who are the people that went online to discuss politics? * Does the profile differ by party identification and candidate support? * As we saw in the primaries, is there still a divide between the politically engaged and unengaged?

RELIGION AND PUBLIC OPINION

***Trends in Public Opinion among Religious Fundamentalists, 1972-2002*, Anne Driscoll, University of California, Davis; akdriscoll@ucdavis.edu, and Katherine Heck, University of California, Davis; keheck@ucdavis.edu**
Religious fundamentalists have emerged as a group of interest to politicians, political parties and survey researchers. It is often assumed that the attitudes of Christian fundamentalists are monolithic and differ sharply from those of religious moderates and liberals. We test these assumptions for several issues that have long been important in U.S. politics. Using data from the General Social Survey spanning thirty years, we examine trends in attitudes toward several topics according to whether respondents belong to a fundamentalist, moderate, or liberal denomination. We examine attitudes toward sexual behavior and privacy (premarital sex, homosexual sexual relations, abortion, divorce), gun control, and government spending in various areas (military, foreign aid, health, environment, welfare, education, law enforcement, drug addiction). Preliminary results suggest while fundamentalists differ from others on issues related to sexual behavior, they are remarkably similar to moderates and liberals on other important issues. The attitudes of both fundamentalists and liberals toward premarital sex and abortion have remained steady over the last thirty years, while the beliefs of religious moderates increasingly resembled liberals. Believing that homosexual sex is always wrong has declined among all groups since the mid-1980s. Attitudes of fundamentalists towards gun control converged toward those of moderates and liberals over time, such that there was little difference in levels of support for gun permits across these groups by 2002. A similar pattern is seen for attitudes towards welfare spending. In other areas of public spending, fundamentalists have consistently resembled those with moderate or liberal religious beliefs. Fundamentalists themselves were not uniform in their attitudes. Even with abortion, an issue taken up strongly by deeply religious Americans, about one-fourth of fundamentalists believed consistently abortion should be legal for any reason; and about a quarter agreed with the Supreme Court decision banning prayer in public schools.

***The National Spiritual Transformations Study: An Overview*, Tom W. Smith, NORC, University of Chicago; smitht@norc.uchicago.edu**
As part of the 2004 General Social Survey, the first national study of personal religious and spiritual transformations was conducted. Among the key findings: 1) while spiritual transformations are common, many were fairly minor in their impact on people's lives, 2) the single most frequent trigger of a spiritual transformation was an illness or accident suffered by the person or someone close to him/her, and 3) religious entities (e.g. God, Christ, Buddha) were not explicitly mentioned in most accounts.

***The Public Acceptance of Evolution and the Big Bang*, Jon Miller, Northwestern University; j-miller8@northwestern.edu, Linda Kimmel, Northwestern University; l-kimmel@northwestern.edu, and Rafael Pardo, Foundation BBVA; rpardo@fbbva.es**
The 2004 presidential campaign reminds us of a growing cultural divide in the United States. One front on which this cultural conflict emerges is the acceptance of science. Americans have been strong believers in the promise of science and technology to improve the quality of our lives, especially in the application of modern science and technology to medicine. At the same time, there has been a declining level of public acceptance of the concept of evolution in the United States over the last 20 years, which places the United States in sharp contrast with Europe and Japan. This paper will utilize data from more than 20 years of national surveys in the U.S. to document the level of public understanding and acceptance of the concepts of evolution and the big bang, both of which provide a direct contradiction to a literal interpretation of the Genesis story in the Bible. The paper will also utilize cross-national studies of public acceptance of these ideas in the European Union and the Japan. The paper will provide a description of the results that will be useful to individuals with and without a quantitative background, and the analysis will include a set of structural equation models to identify the major factors associated with the acceptance or rejection of evolution and the big bang.

***The Impact of "Church Night" on Southern Polls with Short Field Time*, Richard Clark, Carl Vinson Institute of Government; clark@cviog.uga.edu, and Craig Smith, School of Public and International Affairs, UGA; csmith@cviog.uga.edu**
All days of the week are not alike in terms of the population likely to be found home and willing to participate in telephone polls on a given night. This is not problematic unless an anomalous day accounts for a large portion of the ultimate sample. With the short field periods of media polls and pre-election polls, this can be problematic. In the Southern states of the U.S., specifically Georgia-where for many people their church is the center of their community-people regularly gather in the middle of the week for a community meal, bible study, choir practice, or merely to share company with people of the same faith. Our polling in Georgia shows that as many as 31 percent of adults regularly attend church functions on Wednesday night, commonly known as "Church Night". Consequently, the sample frame of available respondents on Wednesday nights is not representative of the general population. This paper examines the impact of "Church Night" on media and pre-election polls. Using data from the Peach State Poll, a quarterly general population survey in Georgia, and the National Anneberg Election Survey from 2000, this paper examines the potential problems when surveys in the South rely on Wednesday nights for a large portion of their sample. In pre-election polls using a rolling sample-wherein a Wednesday evening can account for a third of the total sample-or in media polls-wherein Wednesday can account for up to half the sample-the phenomenon of "Church Night" may have some impact on the generalizability of the results.

A SPACE FOR "PLACE" IN RACIAL ATTITUDES RESEARCH: EXPLORING CONTEXT EFFECTS

***Congregational Context and Racial Attitudes*, Michael Emerson, University of Notre Dame; memerso1@nd.edu**
In this paper I ask whether people's religious congregational context-not attending a religious congregation, attending a uniracial congregation, or attending a multiracial congregation-is associated with people's racial attitudes. Using data from the 1999-2000 Lilly Survey of Attitudes and Social Networks, I compare people in these three contexts along racial, and for comparative purposes, other social attitudes. Consistent with theoretical expectation, I find that congregational context is associated with people's racial attitudes, but not consistently with other types of attitudes. I use interview data drawn from 160 interviews to assess whether congregational context is the cause or the result of people's racial attitudes. I also explore what factors associated with congregational context influence racial attitudes.

***Neighborhood Racial Composition and Perceived Racial Discrimination in the Black Women's Health Study.*, Matthew Hunt, Northeastern University; m.hunt@neu.edu, Lauren Wise, Boston University; lwise@slone.bu.edu, Marie-Claude Jipguep-Akhtar, Howard University; mjipguep-akhtar@howard.edu, and Lynn Rosenberg, Boston University; lrosenberg@slone.bu.edu**
Racial discrimination against blacks is a pervasive problem in the United States. While much research has examined the effects of micro-level interaction (i.e., interracial contact) and macro-level context (e.g., metropolitan-area racial composition) on the racial attitudes of whites, less is known about the effects of such interpersonal and contextual ("place") factors on the perceptions and experiences of blacks. We used data collected from 42,481 U.S. black women during the 1997 follow-up wave of the Black Women's Health Study to investigate the association between neighborhood racial composition ("percent black" at the block-group level in 2000 Census data) and perceptions of racial discrimination. Perceived racial discrimination was measured using self-reports of the frequency of discrimination in "everyday" settings (e.g., being treated as if you are dishonest) and "lifetime" occurrences of discrimination on the job, in housing, and by police. There was a linear inverse relationship between neighborhood percent black and perceived discrimination (i.e., higher percent black associated with lower levels of discrimination), a pattern that held across education levels. For "everyday" discrimination, neighborhood racial composition was most strongly associated with reports of others acting as if they are "afraid of you" and "think you are dishonest." Regarding "lifetime" occurrences, neighborhood racial composition was most strongly associated with discrimination in housing. While the observed inverse association is inconsistent with a study suggesting perceived discrimination is highest in mixed-race neighborhoods, our results support the findings of several studies examining the effect of racial composition of blacks' social networks and workplaces on blacks' psychological well-being. Virtually all participants in the BWHS have completed high school or a higher level of education compared with 85% of black women nationally. Thus, while the BWHS is not a representative national sample, our results may be applicable to the vast majority of black women.

***Educational Attainment and Perceived Discrimination among African Americans: Testing the Racial Awareness and Racial Context Hypotheses*, Mosi Ifatunji, University of Illinois at Chicago; mifatunji@yahoo.com**
Since the civil rights movement and its resulting legislation, African Americans continue to report experiences with racial discrimination. Racial discrimination, or the perception thereof, has been considered in terms of its relationship to lower psychological well-being, and as a barrier to social and economic mobility. This paper focuses on the positive correlation between educational attainment and perceived racial discrimination among African Americans. There are two competing explanations for this finding. The "racial awareness" hypothesis posits that the respondents' racial identity profile mediates the relationship between education and perceived discrimination. The "racial composition" hypothesis posits that the racial makeup of the social context moderates the relationship between education and perceived discrimination. In an effort to understand the effects of racial awareness and racial composition on perceptions of racial discrimination, this study evaluates all constructs in one model. Racial awareness is operationalized using racial identity profiles. Racial composition is measured in terms of the level of racial integration at the workplace. Racial discrimination is considered along five dimensions: lifetime, recency, everyday, institutional and prevalence. Data for this study comes from a multistage area probability sample of the Detroit metropolitan area conducted in 1995. Oversampling was used to attain the 586 face-to-face interviews with African American respondents. Expected findings include support for the racial awareness as mediator hypothesis, such that the relationship between education and perceived discrimination is spurious, and support for the racial composition as moderator hypothesis, such that the relationship between education and perceived discrimination varies by level of integration.

***"Trust in People" Among Black and White Americans -- A Multi-Level Analysis*, Marylee Taylor, Pennsylvania State University; taylor@pop.psu.edu**
A notable race gap exists in "trust in people." Black Americans are substantially less likely than whites to believe that other people are helpful, fair, and trustworthy. Racial inequality on this dimension of psychological well-being thus echoes racial inequalities in tangible life outcomes. This research explores the meanings of self-reported distrust among blacks and whites and evaluates alternate understandings of the race gap in trust. Within each racial group, is the belief that other people are not helpful, fair, or trustworthy linked to broader images of an evil world and corrupt human nature? To lack of confidence in government; in business, educational and medical institutions; in the

media? Is distrust in people related to fear for one's physical safety? Is it associated with economic pessimism? With weak ties to neighborhood and community? Differences between blacks and whites in the meaning of trusts -- discrepancies potentially born of the unique social position of black Americans -- are of special interest. Among blacks, it may be experiences within their own often-distressed communities or perceptions about white Americans' prejudice and discrimination - or both - that fuel distrust of people. To more fully understand the character of distrust within each racial group and reasons for the race gap in this important social psychological outcome, the impact of community context as well as personal characteristics must be assessed. Individual characteristics such as age, gender, and socioeconomic status will have a part in these analyses. But contextual predictors will be featured: region but also such locality-level characteristics as population race composition, average education and economic levels within the black and white communities, and crime rates. Data come from a macro/micro file in which responses to recent General Social Surveys are linked to census and crime statistics for metropolitan areas and non-metro counties.

POLITICAL ISSUES I

***After the California Recall: Governor's Ratings, 2004 Election Choices, and the Implications for Direct Democracy*, mark baldassare, public policy institute of california; baldassare@ppic.org**
California's historic recall election in October 2003 resulted in the removal of an incumbent Democratic governor and his replacement by GOP candidate Arnold Schwarzenegger. This paper analyzes a series of largescale public opinion surveys conducted by the Public Policy Institute of California during Governor Schwarzenegger's first year in office in seeking to answer the following questions: (1) how did the public's attitudes toward their governor's performance in office, trust in state government, and perceptions of the state of the state change over time?; (3) to what degree did attitudes change toward the recall process and direct democracy?; (4) is there a relationship between the governor's ratings and attitudes toward use of the recall process?; and (5) is there a link between the governor's ratings, attitudes toward the recall, and 2004 election choices? The paper discusses the implications of the California recall for the future use of direct democracy.

***Overlapping Decisions in a Presidential Campaign: Battleground Wisconsin in 2004*, G Donald Ferree Jr, University of Wisconsin; gferree@ssc.wisc.edu**
Beyond such factors as partisan identification and special issue positions, elections involving an incumbent are in part referenda in which many ask themselves whether or not the incumbent has done a good enough job to warrant returning him or her to office. If the answer is "yes", there is little or no need for many to consider the qualifications of a challenger. But if the answer is "no", this creates an opening for, but no guarantee of support for, those seeking to replace him or her. Polling conducted by the University of Wisconsin Survey Center as part of its ongoing Badger Poll series demonstrates clearly how these two strands can move separately from one another. In particular, data from a half dozen surveys during 2004 illustrate how the key issue of Iraq and terror posed special difficulties for the challenger, since his natural constituency was divided on a number of facets of this debate in ways in which the President's was not, and also underscore the importance of the challenger's presenting a clear image of what he or she would do in justaposition with the incumbent's actual record. In the end, while Wisconsinites had strong doubts about the performance of George W. Bush in a variety of areas, John Kerry was never able to convince many voters that he would be a clear, acceptable alternative, and the data make clear the extent of skepticism about campaign messages which further made it harder for this case to be put. The end result was the second straight election in which Wisconsin was up for grabs up to the last minute, ultimately remaining a "blue state" by an extremely small margin. The paper also examines key groups within the electorate and traces their positions during the campaign.

***As Easy as 1-2-3? An Assessment of Ranked-Choice Voting in the San Francisco 2004 Election*, Lisel Blash, Public Research Institute/San Francisco State University; lblash@sfsu.edu, and Francis Neely, Political Science/San Francisco State University; fneely@sfsu.edu**
Exit polls are generally used to predict election results before an election is called. However, as a form of intercept survey, exit polls can also be used as an evaluation tool to measure "customer satisfaction" with changes in voting systems. The utility of using an exit poll rather than an "after the fact" phone or mail survey is that it is possible to measure impressions of the new technology or technique prior to those impressions being "contaminated" by the substantive results of the election. This method also guarantees a higher response rate and less screening as it is possible to survey only those who have actually voted. Using student volunteers, we surveyed 3,688 voters in 28 precincts in San Francisco on their first experience using Ranked Choice Voting (RCV) to elect members to the Board of Supervisors. RCV requires that voters rank candidates in order of preference rather than vote for one. In this study, we evaluate the ease of use of the ballot form and self-reported understanding of the new system. We are especially concerned about the impact on "at risk" populations such as the elderly, new immigrants, and those with low levels of education. We will address survey findings in light of additional information provided by detailed vote tabulations. In particular, we will look at the relationship between "satisfaction" with the new system, understanding it, and using it to maximum effect in order to determine what types of questions are most useful in evaluating impact on enfranchisement.

***Nonresponse in exit poll methodology: A case study in Mexico*, Rene Bautista, Gallup Research Center- University of Nebraska-Lincoln; rbautis1@bigred.unl.edu, Mario Callegaro, University of Nebraska, Lincoln. Program in survey research and methodology; mca@unlserve.unl.edu, Francisco Abundis, Parametria SA de CV, Mexico City; fabundis@parametria.com.mx, and Jose A. Vera, Parametria SA de CV; javera@parametria.com.mx**
Recently, survey researchers have been exploring nonresponse error in exit poll methodology. This topic turns into a crucial discussion when pre election polls suggest "too close to call" scenarios and when accurate estimates matter. Three state-level exit polls were conducted in Mexico under the authors' supervision (Veracruz and Tlaxcala states, September 2004; Guerrero state, February, 2005). Two elections were won by less than 1%, and the latter was won by almost 13%. Probability samples of 100 precincts were used in each case with an average of 30 interviews per precinct. Due to voters' weak abilities to read and write, interviews were divided into three parts. First, interviewers administered demographic and political questions face to face. Next, respondents filled out a secret ballot and dropped it in a ballot box. Finally, interviewers administered additional questions regarding economic and political judgments. As a part of the nonresponse research, interviewers recorded refusals in each exit poll. Interviewers registered, by observation, nonrespondents' age and

gender. Guerrero's interviewers filled out a questionnaire regarding problems with electoral officials or political parties, difficulties following systematic skips, distance with respect of polling place exits, experience as interviewer and demographic characteristics. Preliminary results suggest that men voters are more likely than women voters to refuse, particularly in rural areas. These findings are congruent with previous findings in the US. Interviewer's education, gender and experience are critical factors in explain nonresponse rates. Elements such as problems with electoral officials and political parties are not significant predictors of nonresponse. A precinct-level analysis is performed to compare exit poll data versus official results as a way to measure nonresponse impact on exit poll estimates.

INCENTIVES I

***The Effect of Incentives on Two Physician Mail Surveys: A Response Rate Comparison*, Jeanine Christian, Battelle Centers for Public Health Research and Evaluation; christianj@battelle.org**

Garnering cooperation for surveys from physicians presents some interesting challenges for survey research professionals. A common approach for surveying physicians is the use of mail surveys and incentives. Battelle conducted two such studies -- one study was of psychiatrists who treat schizophrenic patients and the other was a study of primary care physicians, dermatologists, urologists and Ob/Gyns who had the potential to treat patients with the genital human papillomavirus (HPV). Both studies were conducted using modified versions of Donald Dillman's "Tailored Design Approach." Additionally, in the psychiatrist survey a $15 money order was included along with the initial package of survey materials as incentive for completing the survey. Clinicians in the HPV survey received a $50 bill included with the initial package of survey materials. The poster will present summary descriptive statistics comparing the two survey's samples and response rates. A brief summary of the protocol utilized in each survey will also be presented.

***Individualized Treatments within a Household: Can Targeted Incentives Raise Young Male Response?*, Christina Frederick, Arbitron Inc.; christina.frederick@arbitron.com**

The Arbitron Radio Ratings methodology uses a list-assisted RDD sample consisting of a 2-stage survey. First, a member of the household is asked to complete a short telephone interview to collect demographic information and the mailing address. Then, each person in the household age 12 and older is asked to complete a self-administered 7-day diary. The diaries for the household are mailed together in one package, addressed to the person who answered the phone survey (henceforth known as the consenter). One metric used to measure the quality of the ratings results is proportionality; that is, how well does the age/sex distribution of returned diaries reflect the distribution within the total population. Households with an 18-24 year old male are particularly difficult to reach, and there is some evidence that the young men themselves are less motivated to return a 7-day diary. Thus, this group may be under-represented in the radio ratings. During Fall 2004, Arbitron conducted a test targeting a special mailing addressed to the young men (aged 18-24) in the household. All the diaries for the household, however, remained in a single package, which was addressed to the consenter. The targeted mailing included a promised incentive for the young man in the household. We hypothesize that the targeted mailing will interest the young male in the survey and encourage him to complete his diary. In addition, we will examine if using differential incentives within a household in an attempt to lower non-response bias adversely affects the response of the other members of the household. Findings with regard to young male response and radio listening behavior, characteristics of responding households, and the effect on proportionality will be presented.

***Value and Timing Strategies in Prize Draws: A Further Examination of the Immediacy Effect in Web Surveys*, Tracy Tuten, Virginia Commonwealth University; ttryan@vcu.edu, Mirta Galesic, Joint Program in Survey Methodology; mgalesic@survey.umd.edu, and Michael Bosnjak, University of Mannheim; bosnjak@tnt.psychologie.uni-mannheim.de**

Prize draws, frequently used as an incentive in Web surveys, out-perform prepaid and promised monetary incentives (Bosnjak & Tuten, 2003). Tuten, Galesic, and Bosnjak (2004) illustrated the presence of an immediacy effect with prize draws in Web surveys. Willingness to participate and complete response were significantly higher when participants were told that the prize draw outcomes would be announced immediately upon submission of the survey than when participants were told that prize draw winners would be announced in a few weeks. The current study seeks to build upon that earlier work by examining the optimization between prize value and immediacy or delayed gratification. According to the standard discounted utility model, the utility of future events is discounted by a constant rate, which motivates people to prefer immediate over delayed rewards. A modification of this model is hyperbolic discounting model, which accommodates the empirical observation that discounting rates decline over time. Applied to prize draws in Web-based surveys, the standard discounted utility model suggests that people will prefer immediate notification of prize draw winners over delayed notification until the point at which the present value of the prize draw amount in the delayed condition exceeds the present value of the prize in the immediate condition. In other words, because people tend to discount values paid in the future, immediate notification prizes can be of lesser value than those in the delayed condition while still achieving higher response and completion rates. In an experiment completed in Fall 2004, we investigated the optimal prize values in immediate and delayed conditions via a Web-based survey among members of a panel using a 3*3experimental design with three notification variations and three levels of monetary incentive. Results indicate the optimal prize values in immediate or delayed conditions to minimize nonresponse and incomplete participation.

***Comparative Analyses of Parallel Paper, Phone, and Web Surveys With and Without Incentives: What Differences Do Incentive and Mode Make?*, Danny Olsen, Brigham Young University; Danny_Olsen@byu.edu, Vaughn Call, Brigham Young University; Vaughn_Call@byu.edu, and Steve Wygant, Brigham Young University; Steve_Wygant@byu.edu**

This study investigates the impacts of survey mode and incentives used by institutional researchers in higher education as they conduct research. Findings from parallel paper, phone, and web administrations of a student advisement survey will be presented with a view toward the following research questions: 1) Do the modes yield comparable results? 2) How do the modes compare on logistical matters such as turnaround time, cost, efficiency, and response rates? 3) What impact do incentives have on response rates? One common issue that nearly all institutional researchers in higher education and that researchers across the country face is determining the most effective means whereby they can acquire needed data. In a higher education setting the portfolio of these means may vary from campus to campus. Numerous ways to gather data exist. Many use traditional paper and pencil surveys, some use phone surveys, and more and more are utilizing web technologies to survey their campus populations. This study will facilitate a number of purposes. From a content standpoint the data collected will inform the university as to student attitudes and experiences concerning both academic and career advising. Using these data the university advising council will make modifications as needed in order to create a university instrument to be administered annually. Methodologically, this study will help answer a number of critical research questions: 1) Do the modes yield comparable results?

2) How do survey modes compare on logistical matters such as turnaround time, cost, efficiency, and response rates? 3) What impact do incentives have? 4) Do multiple invitations for a web survey help or hinder the process? Essentially the methodological dimensions of this study are intended to provide insight into the tradeoffs between paper, phone, and web surveys as administered in a homogeneous campus environment.

TRANSLATION

***How survey questions mean*, Janet A. Harkness, ZUMA - Centre for Survey Research and Methodology; harkness@zuma-mannheim.de**
Research on differences in understanding questions across cultural groups is used to explain how linguistic pragmatics and semantic theory can help us to understand how questions "mean" and to better decide on testing question wording and design. While cross-cultural examples help highlight differences in how questions are understood, the arguments and conclusions advanced apply equally to assumed mono-cultural questionnaire design. The paper explains in what sense survey questions in general -- not just complicated or difficult questions -- do not have unique and fixed meanings. It explores how cultural background shapes researcher decisions on question formulation and design as well as respondents' readings of questions. A key distinction is made between intended meaning (what questionnaire designers intend as the meaning of a question) and perceived meaning (what respondents perceive as the meaning of questions). Examples illustrate how these relate to the notion of shared or common ground (micro and macro social context) essential to successful communication. The paper suggests that rather than elaborating further on question explanations, we need to focus on ensuring that the meaning intended by researchers is also the most salient reading for our targeted population(s). At the same time, cultural diversity within and across survey populations poses special challenges for using identical questions with different cultural groups and across languages.

***Lost in Translation: Quality issues in multilingual research*, Lauren Doerr, NORC - University of Chicago; Doerr-Lauren@norc.uchicago.edu**
The literature supporting the importance of conducting certain surveys in non-English languages is growing: response rates, data quality or generalizability of findings can suffer by limiting the language of administration to English. However, poor translation and interpretation procedures can severely compromise survey quality, and lead to inaccurate comparisons between linguistic and cultural subgroups. Past literature has done much to establish an effective and rigorous methodology for written survey translation; although many of the same risks and benefits apply to oral interpretation, little work has addressed this growing trend. The use of interpreters can greatly widen capabilities for completing interviews in languages other than English, but well-designed training and procedures are necessary to maximize consistency of administration. This talk addresses potential data quality pitfalls with the use of interpreters in multilingual studies as well as recommendations to evade them.

***Translation Effects in the Demographic Health Survey*, Adam Gluck, ORC Macro; Adam.h.Gluck@orcmacro.com**
While there has been much research into the issues of question order and wording choice for English language surveys, there is a limited body of research about the impact of translation on survey design. Increasingly, researchers are finding that they must frequently translate their survey instruments into different languages to reach their populations of interest. This paper studies the relationship between language of interview and attitudinal and opinion questions, in the Demographic Health Surveys collected by the U.S. Agency for International Development. Preliminary research has indicated that there may be a divergence of opinion related to the language of the interview. We will use regression and likelihood ratio analysis to determine whether the language of the interview is a driver of respondent's answers, or if some other factors such as education, gender, age, or ethnic affiliation are the root cause of differences of opinions.

***Managing Survey Translation: Knowledge Base for Researchers who are not Multi-Lingual*, Man-chi Mandy Sha, NORC; sha-mandy@norc.org**
Translating English surveys in one or multiple languages has become an increasingly important part of survey design aimed to minimize measurement errors due to non-response or non-coverage of a population. However, the process may be daunting or sometimes even preclude survey researchers who are not multilingual. The Culture Methodology Program at NORC continues to work to demystify the myth by developing processes by which researchers with limited multi-language skills can effectively manage translations. The process is not about what words translate better, but rather knowing what to expect and understanding what needs to be done in the process of managing different types of survey translations. They include: documents (advance or fielding materials for respondents), training materials for bilingual interviewers, and most importantly, the questionnaire. A questionnaire translation involves translating new questions, existing questions that require updates, or questions adopted from previously translated surveys. We will focus on the two most complex types of questionnaire translation management: 1) performing updates to existing questions, using our experiences from the Survey of Consumer Finances that has been conducted every three years since 1983, and 2) adopting translation from previously translated surveys. The latter is an applied response to past AAPOR president Schulman's call to examine the "copy and paste" practice in English questionnaire design, which means copying established survey questions and pasting them into new surveys without carefully considering contextual changes over time.

CELLULAR PHONES I

***Characteristics Related to Cell Phone Status: Why Generation Y Should be Targeted*, Laura Hancock; laurahancock@comcast.net**
In assessing the feasibility of including cell phone numbers in telephone survey sampling frames, researchers could examine more closely the characteristics of the telephone owning population. Comparisons could be made between those who own only cell phones and those who still have landline telephones. If the two groups indeed differ demographically, then the portion of the population owning only cell phones are not being well represented in the current RDD telephone surveys. If the two groups proved to differ demographically, then the cell phone only population should be targeted in future telephone surveys. In other words, if the two groups differ, then including cell phone numbers in survey sampling frames may be a worthwhile endeavor. This paper explores the theory that the cell phone only population differs demographically from the rest of the telephone owning population. The hypotheses that level of income and age are

related to cell phone status (cell phone only ownership vs. landline telephone ownership) are tested using data from the 2004 JPSM Survey Practicum survey. The survey was conducted using two separate samples; a cell phone number sample and a landline telephone number sample. For analyses in this study, a combined sample was used, which consisted of a total of n =1,324 completed interviews. Results from analyses of this data show that income and age are both significantly related to cell phone status.

***Comparison of Cell-Only and Landline Telephone Users Among 18-24 Year-Olds*, Robin Albee, Assessment Resource Center, University of Missouri-Columbia; albeer@missouri.edu, and James Cole, University of Missouri; colejs@missouri.edu**
As people continue to abandon traditional landline telephones in favor of cellular telephones, public policy researchers relying on random-digit dialing increasingly face the specter of potentially biased samples. One vanguard of this "cell-only" phenomenon can certainly be seen as young adults, aged 18-24. How does this crucial demographic differ, if at all, by type of telephone usage? And what does this suggest in terms of improving the quality of survey research? This research uses a web-based survey to survey undergraduate students at the University of Missouri-Columbia. Analysis focuses on differences among the 3 types of students: landline-only, cell-only, and landline and cell. In addition to demographic variables, students are probed about assorted issues, including questions relating to survey taking inclinations and preferences.

***Can You Hear Me Now?: Differences in Vote Behavior in the Cell and Landline Populations*, Tarek Albaghal, University of Maryland; talbaghal@survey.umd.edu**
Recent research indicates that survey data collected using standard random-digit dialing (RDD) or list-assisted telephone interview methods may encounter problems that bring into question the generalizability of results. One of the major problems facing survey methodologists and social scientists employing this data is the penetration and wide use of cellular opposed to landline telephones as a main source of communication. Initial research shows that there are important differences in a number of respondent characteristics between those answering surveys over cellular and landline telephones. It is therefore likely that these respondents also differ on a number of behavioral and social aspects as well. To dismiss these factors would undermine an array of research involving a number of social issues. This paper

examines the comparative mode effect and ownership differences of these modes on self-reported voting behavior. There are theoretical reasons to believe cellular telephone respondents and those owning will differ in vote report. Understanding the potential impact of these populations impact survey methodologists and political scientists alike when developing theory on such behaviors. A dual frame list-assisted survey collected data from respondents on either landline or cellular telephones. Two aspects are examined, whether mode of response itself is predictive of response, and whether different telephone ownership patterns are related to voting behavior. Question order - a well-known correlate to response distribution - of retrospective and prospective voting behavior questions was alternated to determine differential sensitivity between modes. Logistic regression shows that cellular telephone response is not a significant predictor to self-report of voting, although question order does have some impact as expected. More importantly, however, ownership of only a cellular telephone is a significant predictor of reported voting behavior. The model incorporates significant correlates from the political science literature, which are significant indicators. Implications and further research needed are examined.

***Ownership and Usage Patterns of Cell Phones: 2000 - 2005*, Peter Tuckel, Hunter College; ptuckel@hunter.cuny.edu, and Harry O'Neill, NOP World; honeill@nopworld.com**
This paper has three major objectives: (1) to measure the incidence of individuals who have jettisoned their land-line phones as well as those who are "heavy users" of cell phones, (2) to assess the possible biases that both "exclusive" and "heavy users" of cell phones might introduce into the conduct of regular telephone surveys, and (3) to discuss the implications of the findings with respect to carrying out surveys via the cell phone. The data for this study will be based on four nationwide surveys of face-to-face interviews that have been carried out between 2000 and 2004 and one similar survey to be carried out in 2005. The sample size for the 2005 survey will consist of approximately 1,000 respondents. Respondents in the 2005 survey will be asked about the following topics: (1) the frequency with which their household uses a land-line phone to answer voice calls, (2) ownership of a cell phone, (3) the percentage of all personal calls they make and receive via their cell phones, (4) the frequency of keeping their cell phones turned on during the day when not making calls, and (5) the likelihood of their household abandoning land-line phones and relying exclusively on cell phones to both initiate and receive voice calls. The 2005 survey will also investigate whether respondents view their cell phones as mainly a private or a public mode of communication by asking them how willing they would be to have their cell phone numbers listed in a directory of cell phone numbers. Finally, the 2005 survey will also ask cell phone owners about their general willingness to participate in a telephone survey via their cell phones. For comparative purposes, respondents will also be asked about their general willingness to participate in a telephone survey via their land-line phones.

PANEL SURVEYS

***Looking for Mr. Right: Effective Locating Techniques for Panel Surveys*, Angela Herrmann, NORC; herrmann-angela@norc.uchicago.edu, Karen Grigorian, NORC; grigorian-karen@norc.uchicago.edu, and Syed Ahsan, NORC; ahsan-syed@norc.uchicago.edu**
For many panel surveys, overall response rate is declining. A diminished response rate negatively affects the accuracy of that round's data as well as the total quality of the longitudinal dataset. An important step in addressing the growing response rate problem is to decrease the number of non-contact cases. To that end, we will examine the effectiveness of various locating techniques used during the 2003 Survey of Doctorate Recipients data collection effort. The SDR is a panel survey administered biennially to a national sample of some 40,000 doctorate-holders in the fields of science and engineering. During the 2003 round of SDR, over 30% of the sample required locating. The production center staff deployed a wide variety of locating techniques, including database searches with locating vendors, calling the past round's contacting information, calling past employers, calling Directory Assistance, contacting alumni associations, contacting professional associations, and conducting Internet searches of publicly available websites. The staff carefully tracked each locating step tried and the ultimate success or failure of that technique. Further, the staff tracked the type of data that was obtained at each step. Comparing the outcome of various locating techniques with sample member characteristics, i.e., age, sex, ethnicity/race, length of time since last survey response, employment industry and employment type, yields a wealth of information about what locating steps are effective for which sample members and suggests a protocol to follow to maximize the effectiveness of locating resources.

***Ensuring High Cooperation Rates in a 4-year Longitudinal Study of People with Multiple Sclerosis*, Todd Robbins, Abt Associates, Inc.; todd_robbins@abtassoc.com, Michele Laramie, Abt Associates, Inc.; Michele_laramie@abtassoc.com, Sarah Minden, Abt Associates, Inc.; Sarah_minden@abtassoc.com, Debra Frankel, Abt Associates, Inc.; debra_frankel@abtassoc.com, Carin Cartwright-Chunga, Abt Associates, Inc.; carin_chunga@abtassoc.com, Brander Sieber, Abt Associates, Inc.; brander_sieber@abtassoc.com, and Paul Howard, Abt Associates, Inc.; paul_howard@abtassoc.com**

In this paper we describe in detail, the various steps and activities taken that continue to ensure participants' retention and cooperation in a 4-year longitudinal study of people with Multiple Sclerosis. The Sonya Slifka Longitudinal Multiple Sclerosis Study follows a nationally representative sampling of almost 2000 persons with Multiple Sclerosis (MS). This project collects a comprehensive set of personal and disease-related information in a series of bi-annual telephone interviews. Our strategies to ensure cooperation and retention include: a periodic study newsletter mailed to all participants; a participant 'preferences' database describing participant preferences such as best times to call, accommodations required etc.; a study diary/calendar; extensive interviewer training regarding Multiple Sclerosis and techniques for interviewing persons with disabilities; frequent interviewer debriefings; and a toll-free project number. The resulting high completion rates as a percent of reachable sample are as follows: 93.9 at six months; 96.7 at annual; 97.7 at eighteen months; 97.6 at twenty-four months; and 97.3 at thirty months. Our experiences with the various tools and strategies we use to ensure high cooperation in this study can serve to inform other researchers on ways to ensure high rates of cooperation in longitudinal studies of people with disabilities and help to mitigate the problems associated with survey non-response.

***Using process data to predict attrition from a panel survey: a case study*, Femke De Keulenaer, University of Antwerp; femke.dekeulenaer@ua.ac.be**

A major concern with panel surveys is the increasing loss of respondents that occurs as the panel ages. In this paper we study attrition from the Panel Study of Belgian Households (PSBH). The PSBH is a longitudinal data set beginning in 1992 with approximately 11500 individuals, living in about 4400 households. Since then, data collection was continued on a yearly basis. The PSBH has adopted certain procedures to minimize panel attrition as far as possible. Nevertheless, by the time of wave eleven, more than 50 percent of the original sample members had left the sample due to loss of contact, inability to respond, or refusal to continue participation in the panel. One advantage of studying attrition is that the first waves of the survey are a source of information on possible causes of later wave cooperation. We discuss socio-demographic and socio-economic correlates of attrition in an attempt to describe individuals who may be at risk of attrition. In addition to survey data, the first waves of the panel also contain process data that could be used to predict attrition. For this reason we will also include characteristics of the previous interviews in our analyses, such as the number of attempts it took to get an interview, the mode of data collection, the language in which the interview was conducted, and the amount of item nonresponse. Survey data and process data from previous waves that are related to future participation can be used to alter the design during the course of the panel in order to improve panel survey quality by lowering the amount of attrition. In addition, survey data and process data could be used to tailor or adopt the procedures used to minimize panel attrition to fit the particular characteristics of the respondent.

***Increasing Response to Mailed Follow-up Surveys: A Test of Two Motivators*, Sue Mann, Fred Hutchinson Cancer Research Center; smann@fhcrc.org, Diana Lynn, Fred Hutchinson Cancer Research Center; dlynn@fhcrc.org, and Arthur V. Peterson, Jr., Fred Hutchinson Cancer Research Center; avpeters@fhcrc.org**

Randomized controlled trials (RCTs) require high outcome survey response to obtain a valid intervention evaluation. To test ways to increase survey response, the Fred Hutchinson Cancer Research Center's Youth Smoking Studies Group embedded a methodology experiment within the follow-up procedures for an adolescent smoking cessation RCT with 2,894 trial participants (TPs) from 50 randomly selected Washington state high schools. The standard procedures for the RCT's 3-year follow-up of TPs (at age 19) include (1) a TP 'location' request to parents and (2) a 'survey' request to TPs. Each of these requests includes an initial mailing, multiple non-responder mailings (Day 16 and 30 for parents; Day 16, 30 and 47 for TPs), and non-responder phone calls (starting Day 47 for parents, Day 60 for TPs). The initial 'survey' mailing to TPs includes a $10 pre-paid cash incentive; non-responders at Day 30+ are promised $20 cash upon receipt of their survey. From experience, we know that TPs of early-responding parents (28 days) are most responsive to the 'survey' request, while TPs of late/never-responding parents (>28 days) are least responsive. To maximize response from these two groups, we tested two new motivators in a small randomized trial among a randomly selected sub-sample of the RCT: (1) including $2 cash in the initial parent mailing and (2) using an 11"x5" "pillow" box container for the initial 'survey' mailing to the TPs of late/never-responding parents. The $2 incentive increased mailed parent response over controls at Day 16: 41.7% (25 of 60 mailed with $2) vs. 20.3% (25 of 123 mailed without $2) and at Day 28: 46.7% (28/60) vs. 29.3% (36/123). Results for the "pillow" box test and for the effect of both motivators on overall survey response (in progress) will be available for presentation at the 2005 AAPOR conference.

PRE-ELECTION SURVEYS I

***Media Coverage & Nonresponse: A Time Series Examination of the 2004 Election Cycle*, Natalie Jomini Stroud, University of Pennsylvania; tjomini@asc.upenn.edu, and Kate Kenski, University of Pennsylvania; kkenski@asc.upenn.edu**

Past research suggests that there is a relationship between survey response and topic salience, namely that individuals responding to a survey are likely to find the survey topic more salient than nonrespondents. For election surveys, nonresponse resulting from lack of salience can impact findings because respondents interested in politics may differ systematically from nonrespondents who are less interested in politics. The agenda-setting model suggests that media coverage should heighten interest. Thus, as media coverage of political campaigns increases over the course of an election, refusals should decline. Using data from the 2004 National Annenberg Election Survey (NAES), which has been conducted nearly continuously between October 7, 2003 and November 1, 2004, this study investigates the issue of nonresponse in a RDD telephone survey across the 2004 election cycle. This study investigates daily changes in the refusal rates using time series analysis. A content analysis of the frequencies of presidential campaign stories mentioned in the New York Times and three network news broadcasts was matched against time series from the NAES to demonstrate that increases in press coverage of the election were negatively related to the survey refusal rate.

***Estimating Non-Response in Presidential Polls*, Eugene Ericksen, Temple University; eugene.ericksen@nera.com, and Sarah Butler, Temple University; sarah.butler@nera.com**
Perhaps one of the most challenging aspects of survey research is the understanding the effects of non-response. This is a particularly important problem for surveys used in litigation, especially if there is reason to believe that the characteristics of respondents and non-respondents differ. The problem is made more difficult by existing litigation standards that call for the achievement of very high response rates that are beyond the reach of any survey research firm or academic center. We would like to argue that these standards should be replaced by a standard requiring a sensible method of non-response adjustment. As researchers, we are often unable to ascertain whether individuals not surveyed differ significantly from individuals who have agreed to provide their opinions. The two most recent U.S. presidential elections provide valuable sources of data for comparing and evaluating methods of estimating the impact of non-response. This paper will examine data from public opinion polls preceding presidential elections, with a focus on how likely voters were defined and how differential non-response was adjusted for. We will compare the methodologies of twenty-three polls (including ABC, Gallup, Harris Interactive, The Los Angeles Times and others) that reported results in the public media and whose companies have made at least some of the details of their methodologies available. The paper will also examine how the estimates of non-response for these polls varied during the months leading up to the election. The analyses from 2004 and 2000 will hopefully allow us to offer a more effective method of estimating non-response and will allow us to make suggestions about possible methods of testing the impact of non-response designs in future survey work.

***A Comparison of RDD and Listed Sampling in a National 32-Night Tracking Poll*, Daniel Merkle, ABC News; dmmerkle@aol.com, Gary Langer, ABC News; gary.e.langer@abc.com, and David Lambert, TNS; Dave.Lambert@tns-global.com**
This paper explores a number of important methodological issues in pre-election polling and how they affect survey estimates. The data are from ABC News' 2004 pre-election tracking poll, the only media-sponsored daily tracking poll conducted for this election. More than 21,000 respondents were interviewed by RDD over 32 consecutive nights leading up to the election. We begin by looking at how party identification varies across the many waves of tracking. Are there large, short term swings in party id as some have suggested? And are these changes real and meaningful or simply random variability? Related to this, the paper explores the hot topic of using party identification as a weighting factor. Does it make sense to do this, and how much of a difference does it make in the estimates? We also investigate whether using a sampling frame of only listed numbers is as good as RDD, as some have argued. Each phone number in the tracking poll was checked against a database of listed numbers and flagged as listed or unlisted. We compare survey estimates from the RDD sample to those from only the listed numbers across a wide range of demographic and attitudinal variables. The tracking data also allow us to look at the issue of day of week effects in pre-election polls. Are certain nights of the week consistently better for Democrats or Republicans as some have suggested? And finally, how important are callbacks? We look at survey estimates by the number of attempts needed to reach respondents.

***How SurveyUSA weights its election poll data*, Joseph Shipman, SurveyUSA; jshipman@surveyusa.com, and Jay Leve, SurveyUSA; jleve@surveyusa.com**
SurveyUSA compares unweighted (unpublished) data and weighted data from 30 statewide election polls for the November 2004 general election. Weights are derived from Census data compiled into a 4-dimensional sample grid, with dimensions for gender, age, race, and geographic region. The impact of the weighting method on Republican-Democrat spreads and the accuracy of vote projections is examined. Unweighted and weighted demographic distributions are compared. Weighting adjustments related to other variables are discussed.

***Partisan Identification in the Pre-election Environment*, Frank Newport, Gallup Organization; frank_newport@gallup.com**
This paper addresses two conceptions of partisan identification - a variable that has become one of the more controversial measures in the pre-election survey environment. One view conceptualizes the distribution of partisan identification across the American population as a relatively stable variable subject only to slow over time shifts. Partisan identification is thus viewed as a stable characteristic of individuals, with the concomitant assumption that most individuals have "a" partisan identification similar to age, gender, height, or race. This perspective assumes that aggregate patterns of partisan identification are similar to patterns of such ascribed characteristics as age, gender and race. The second view assumes that partisan identification is more fruitfully viewed as an attitudinal variable, subject to significant short-term variability, particularly in a pre-election environment. The relevance of these two ways of viewing partisanship is significant. The first perspective leads one to expect survey waves in an election year to show minimum variability in partisan identification, under an assumption that the marginal distribution of partisanship in each survey should match some specified and stable population parameter. The second perspective accepts partisan variability in survey waves, and assumes that partisan identification in the population is a dependent variable to some degree subject to the same factors that influence political variables such as vote intention and perceptions of candidates. The paper analyzes assumptions implicit in the ways partisan identification is measured, including the hypothesis that although partisanship may vary in the short term in an election environment, it becomes more predictably stable as Election Day approaches. The paper relies primarily on over 30,000 Gallup surveys conducted in 2004 in which partisan identification was measured.

MEDIA USE

***Media use, political talk and social capital*, Hernando Rojas, University of Wisconsin-Madison; hrojas@wisc.edu, and Dhavan Shah, University of Wisconsin-Madison; dshah@wisc.edu**
In recent years, academics have popularized the term social capital to describe how basic features of community life provide the means for citizens to cooperate on joint problems. Social capital has been expressed theoretically as an associational capability grounded in norms of trust and reciprocity that makes collective action possible. However, the conditions for social capital and its relationships to civic and political engagement have been empirically assessed mostly within democratic contexts. This study examines social capital in a society in turmoil, one that resolves fundamental political conflict through a combination of democratic practices and violence. To investigate these relationships, we rely on survey data collected in the city of Bogota in 2003. The data was collected on behalf of city government, as an assessment of civic culture in Bogota. The Civic Culture Study used a stratified sampling technique to recruit a representative sample of adult residents of Bogota, and generated 1,433 face-to-face completed responses. Results suggest that demographics such as age and education are related to political/civic participation; Informational uses of media, political conversation, political knowledge and political interest were also found to be significantly related to participation. However, once institutional trust, social trust, political satisfaction and

associational ties are included in the model, a distinct pattern emerges: there is no relationship between institutional trust, social trust and political satisfaction with participation. On the other hand, associational ties emerge as the single most important predictor of participation in the model. Beyond the effects of having multiple ties to associations, the role played in the association is also significant to spur civic action. Our data indicates that democratic norms considered so important for democratic functioning in stable and orderly societies do not function similarly in societies in turmoil. Instead, real world ties, not perceptions of social norms, shape participation.

***The Pitfalls of Self-Reported News Exposure*, Markus Prior, Princeton University; mprior@princeton.edu**
The purpose of my study is to examine the accuracy of survey-based self-reports of news exposure. Previous research has shown that people's self-reports can differ considerably from independent assessments of media use such as newspaper circulation (Price/Zaller 1993). These past research efforts do not go beyond anecdotal evidence, however. My goal is to assess the accuracy of self-reported TV news exposure more systematically. I use Nielsen data as a benchmark against which survey self-reports can be compared. The Annenberg National Election Study (ANES), conducted as a rolling cross-section with independent random samples for each day of the year 2000, includes an appropriate question about respondents' exposure to the nightly network news "in the past week." This question was asked in all ANES interviews and can thus be used to calculate weekly survey-based estimates of the network news audience for the entire year. These survey estimates are compared to weekly Nielsen audience estimates. I also compare ANES and Nielsen estimates of debate audiences in 2000. Results show severe overreporting of news exposure. Survey estimates of network news exposure follow trends in Nielsen ratings relatively well, but exaggerate exposure by a factor of three. To make matters worse, the extent of this bias is not constant across time or individuals. For example, self-reported exposure to debates declines as time passes (but remains higher than Nielsen estimates suggest.) Faulty self-reports emerge not as a result of flawed survey design, but because people simply cannot recall or estimate their own news exposure.

***Putting survey response data into context of everyday life: Data modeling with an electronic measurement system of media exposure*, Beth Webb, Other; beth.webb@arbitron.com, and Jay Reid, Arbitron; jay.reid@arbitron.com**
Arbitron has recently developed a new multi-media measurement method - called the Personal People Meter (PPM). Television ratings obtained using the PPM have been considerably higher than those measured by Nielsen Media Research's (NMR) set-top diary. To further explore this ratings difference, Arbitron and Nielsen Media Research (NMR) placed stationary PPMs in 19 households that were already outfitted with TV set-top meters in order to monitor how often these PPM's picked up exposure from TV's in the same room versus a different room. Although it is currently unknown whether people using the diary/set-top system credit such instances as "watching or listening", an audio-based measurement system would likely credit such instances. In order to know whether the higher PPM TV ratings are due to a difference in the definition of audience between the two measurement systems we must determine how much the PPM credits out-of-room TV exposure. Since we did not know where the people were located when the TV's were tuned, we cross-tabulated separate data sources to derive estimates of how likely people were in the same or different rooms as the TV's. The data sources consisted of: a survey conducted with the household to determine what times they were usually away from home (e.g. at work or at school) and the NMR people meter data. Using this data, we assigned to each person in the household, their probability of being in each room of the household. These probabilities were overlaid onto the media detections of the stationary PPM's in each room of the household to determine that if the person were wearing a meter, how many minutes of detected TV exposure would have occurred while the person was in the same room as a TV and how many would have occurred while in a different room than the TV.

***The PPM/Diarykeeper Study: Same-Sample Comparisons of Radio Listening Behavior*, Fran Harmon, Other; fran.harmon@arbitron.com, and Beth Webb, Other; beth.webb@arbitron.com**
For the past 50 years, Arbitron has used a seven-day self-administered radio diary to measure radio listening. Concerns about respondents' ability to accurately report radio listening in an increasingly complex media environment, as well as the desire to develop single source measurement of multiple media, led us to develop a personal metering device (about the size of a pager) which automatically detects exposure to encoded media with minimal effort on the part of the respondent. This device also contains a motion detector so that compliance could be monitored. We recruited a panel that included approximately 850 persons age 12+ in the Philadelphia metro to carry these meters, and compared the resulting radio usage results for Spring 2002 to those of our syndicated diarykeeper sample of approximately 4500 persons in the Philadelphia metro over the same time period. This study details the differences in radio usage that were detected between the two measurement methodologies. In addition, we conducted a second study to confirm that the differences we saw were not due either to the differences in sampling methodology between the diary service and the personal meter panel, or in the characteristics of persons who would agree to participate in one measurement method versus the other. In this study we recruited a sample of former personal meter panelists who were part of the Philadelphia 1500 test panel during the Winter 2003 survey period. They were recruited by telephone and asked to keep a standard syndicated radio diary for a one-week period during February 2004. A total of 460 usable diaries from 254 households was returned. In this paper, the differences in radio usage found using different measurement methods within the same sample at different points in time will be compared to the differences found using the different measurement methods in independent samples.

THE EUROPEAN SOCIAL SURVEY: IMPROVING SURVEY QUALITY IN A CROSS-NATIONAL CONTEXT

***From design to implementation: methodological innovation on the ESS*, Roger Jowell, City University; r.jowell@city.ac.uk, Rory Fitzgerald, City University; r.fitzgerald@city.ac.uk, Gillian Eva, City University; g.eva@city.ac.uk,**
Longstanding obstacles to rigour and consistency in the conduct of social surveys across a disparate group of nations and cultures, create an inherent conflict between the twin imperatives of equivalence on the one hand and due respect for cultural norms on the other. The unique organisational infrastructure of the ESS exemplifies how such clashes of interest may be resolved. From the outset, the ESS has set new precedents for improved cross-national survey practice, not least the importance it has placed on well-documented technical specifications and the necessary combination of top-down and bottom-up forms of research governance. This paper presents an overview of the methodological developments of the ESS and examples of how the peculiar relationship between central supervision and local control actually works on the ground. In particular, the authors describe how the unique 'federal structure' adopted by the ESS makes it possible to maintain a strong central measurement strategy that is nonetheless influenced by the experience of local co-ordinators and roving advisory panels, all working in tandem to achieve a common goal. Evidence for the success of this strategy is presented in the form of examples of local innovations from national co-ordinating teams, which have been developed in response to the challenges presented by the strict methodological specifications of the ESS.

***Refusal conversion and the estimation of non-response bias in the European Social Survey (Round 1): An analysis of contact forms combined with substantive data*, Jaak Billiet, University of Leuven; jaak.billiet@soc.kuleuven.ac.be, Inneke Stoop, Social & Cultural Planning Office, The Netherlands; i.stoop@scp.nl, and Rory Fitzgerald, City University; r.fitzgerald@city.ac.uk**
Conclusions in international comparative survey research are sensitive for non-response bias when the fieldwork in the participating countries is characterized by substantial differences in non-response rates. For that reason, in round 1 of the ESS, the National Coordinators in each of the 22 participating countries were urged to reach a target response rate of 70%. To help fieldwork organisations to reach this target rate, they were encouraged or recommended to implement a set of current best methods concerning the type, number and time of calls, refusal conversion attempts, and an extensive contact form for each sample unit. This paper analyses the efforts made in order to increase response rates and evaluates the results. Strategies for refusal conversion led to a substantial improvement of response rates in certain countries while in other countries they had little impact. Even so, large differences in response rates between countries were observed (ten countries achieving response rates close to 70% or even higher and four falling below the 45% level). These large differences in response rates are likely to produce bias in some substantive variables. The second part of the paper analyses "traces of bias" by combining contact form information with substantive data from the ESS. Countries with large numbers of converted refusals are used to analyse the relationships between the type of respondent (cooperative, soft refusal, or hard refusal) and relevant substantive variables such as trust in politics, trust in other people, political participation, interest in politics, and ethnic prejudice. The purpose was to discover if the converted respondents were indeed closer to actual non-respondents than to other respondents. The paper ends with a discussion about the pros and cons of refusal conversion in the context of survey burden and survey costs.

***Methodological advances on the ESS: A mixed mode future?*, Caroline Roberts, Centre for Comparative Social Surveys, City University; C.E.Roberts@city.ac.uk, Peter Lynn, ISER, University of Essex; plynn@essex.ac.uk, Robert Manchin, Gallup Europe; Robert_Manchin@gallup-europe.be, and Agnes Illyes, Gallup Europe; agnes_illyes@gallup.hu**
In striving for cross-national equivalence, the ESS currently insists on face-to-face interviewing as the exclusive mode of data collection. However, countries vary in their preferred modes of data collection and in their experience of conducting surveys using personal interviews. Moreover, the cost of carrying out survey research in this way is becoming increasingly prohibitive. Meanwhile, the use of the internet is spreading to parts of the population it has not previously conquered, with the result that various distinguished time-series whose data have so far been collected face-to-face or by telephone will have to consider alternatives. Should they risk switching modes (or adding new ones), or should the realistic fear of confounding their time series prohibit them from doing so? The ESS has responded to these challenges with a programme of research investigating the feasibility of permitting - or even encouraging - mixed mode data collection in future rounds of the survey. This research - conducted in collaboration with Gallup Europe - is examining the different categories of question, which seem to be particularly sensitive to the mode in which they are administered. The paper will present the results of the first two phases of experimental work and the conclusions reached with respect to the future of the ESS, revealing how modifications to question wording may indeed mitigate some of the more serious mode effects.

A new source of US-European comparisons on citizen engagement, **Marc M. Howard, Georgetown University**
This paper introduces a major new American survey, "Citizenship, Involvement, Democracy" (CID), which will be carried out in May/June of 2005. The survey, which has been developed within Georgetown University's Center for Democracy and the Third Sector (CDATS), will consist of in-person interviews with a representative sample of Americans, and it includes extensive questions about important themes and concepts related to civil society and democracy. The American CID survey began as a loose collaboration with a European survey with the same name, but it has now become more closely connected with the ESS, which included a CID module in 2002. The U.S. version of the CID includes many questions from the 2002 ESS, as well as a host of new and innovative questions - particularly related to the themes of informal social networks and democratic values - that are connected to lively debates about civic engagement and democracy. The survey promises to generate an unusually rich perspective on American civic participation in both the public and private realms, and to generate comparative data that will allow researchers to investigate the similarities and differences between the U.S. and 22 European countries in terms of the forms and mechanisms of civic engagement.

DIFFICULT AND RARE POPULATIONS

***Interviewing Persons with Hearing Impairments*, Elisha Smith, Mathematica Policy Research; esmith@mathematica-mpr.com, Susan Mitchell, Mathematica Policy Research; smitchell@rti.org, and Debra Wright, Mathematica Policy Research; dwright@mathematica-mpr.com**
When interviewing persons with disabilities it is important that every effort be made to ensure that all persons, regardless of disabling condition, are able to participate in the survey process. The National Beneficiary Survey (NBS), sponsored by the Social Security Administration's (SSA) Office of Disability and Income Security, collects data from nationally representative samples of SSA disability beneficiaries. In 2004, data were collected by means of a CATI survey with CAPI follow-up of nonrespondents. To conduct the NBS, we made a series of modifications to standard instrumentation and data collection procedures to ensure representation of the broadest spectrum of the beneficiary population. This paper focuses on the modifications we made for people with hearing impairments, a growing segment of the population. The literature on interviewing persons with hearing impairments is sparse. In this paper, we summarize the literature and update what is known by describing our experiences conducting interviews using assistive technologies. These technologies include amplifier telephones, text typewriters (TTY), telecommunications relay service, and instant messaging. In addition, we discuss in-person interviewing using sign language interpreters. Finally, we report the proportion of the hearing-impaired population that responded by each mode, and the overall response rate. Our experience suggests that interviewing persons with hearing impairments does not present extraordinary challenges. The barriers to response can be overcome through modifications to data collection procedures, and the use of assistive technologies.

***Hospice, HIPAA, and Hope: Survey Research with the Terminally Ill*, Terri Kovach, Wayne State University, Detroit MI -- Sociology; kovach@tdi.net**
Quantitative studies have provided insight into the needs and desires of those who are dying as well as those who care for the dying. Yet survey research with the terminally ill is particularly challenging. There are a number of ways for error or bias to occur while researching this population. This is a narrative description of the experience of developing and administering a theoretically-based survey instrument to measure suffering in terminally ill patients. The setting is a Veterans Administration hospital and outpatient clinic. This paper will describe issues involved in instrument development and administration in relation to population definition, recruitment decisions in the Health Insurance Portability and Accountability Act (HIPAA) environment, interviewer characteristics, item order, and protections for this highly vulnerable population. It also suggests that participation in research protocols offers the terminally ill an opportunity for personal growth and contribution to the larger community . Research into quality of life and suffering at end of life is developing more sophisticated quantitative strategies. Understanding the specific challenges involved in survey development and administration with the terminally ill will further the research agenda for end-of-life studies.

***Amish Women's Health Survey: Methodological Considerations with Unique Populations on Sensitive Topics*, Berwood Yost, Center for Opinion Research; byost@fandm.edu, Christina Abbott, Center for Opinion Research; cabbott@fandm.edu, Jennifer Harding, Center for Opinion Research, Franklin & Marshall College; jharding@fandm.edu, and Angela Knittle, Center for Opinion Research; aknittle@fandm.edu**
Specialized interviewer training is necessary to promote high response rates and minimize refusals, particularly when sensitive issues are addressed. This need becomes pronounced when the population under study is unique and adheres to social norms that differ considerably from the population as a whole. A recent research project used specially trained interviewers to conduct face-to-face interviews with women living in traditional Amish homes (aged 18-45) on women's health issues, particularly those related to pregnancy outcomes. This survey is part of a larger effort to reduce disparities in the rates of preterm birth and low birth weight babies between female populations by improving their preconceptional health. This research, apart from generating information on medically underserved populations with high-risk pregnancy factors, will examine the development and use of culturally sensitive interviewer training when dealing with unique subpopulations. Both findings and methodology will be discussed.

***A Brain is Terrible Thing to Lose: Locating U.S.-Educated Foreign Nationals Intending to Live Abroad*, Lauren Seward, NORC; seward-lauren@norc.org, Angela Herrmann, National Opinion Research Center; herrmann-angela@norc.org, and Tom Hoffer, NORC; Hoffer-Tom@norc.org**
Surveys of post-graduation outcomes typically restrict their samples to individuals residing in the U.S. Because locating individuals living outside the U.S. is widely assumed to be difficult and expensive, otherwise-eligible individuals who have moved outside the U.S. or who indicated definite plans to do so are often categorically excluded from the sample frame. While reasonable in many respects, the assumption of ineligibility is rarely tested and the risk of sample undercoverage is thus unknown. As part of the 2003 Survey of Doctorate Recipients, a biennial panel survey of doctoral scientists and engineers sponsored by the National Science Foundation, a systematic effort was undertaken to check the eligibility of non-U.S. citizens who planned to live outside the U.S. following receipt of the doctorate from a U.S. university. This paper presents results of that effort, including detailed descriptions of the relative effectiveness of various locating methods. These findings and the procedures identified will provide useful information to a broad range of surveys dealing with in- and out-migration of sampled subjects in an increasingly mobile world.

GAINING COOPERATION

***Analysis of Behavioral Risk Factor Surveillance System Partial Completes and Terminations*, Herbert Stackhouse, Centers for Disease Control and Prevention; hls4@cdc.gov, and Ziya Gizlice, NC State Center for Health Statistics; ziya.gizlice@ncmail.net**
In 2002, the number of BRFSS final disposition codes increased from 14 to 31. Interviews with a disposition of 110 (complete) or 120 (partial complete) are included in the dataset for calculation of estimates of prevalence rates and risk factors. Numbers given a disposition code of 210 (terminations within questionnaire) are not included in the dataset for producing estimates because they do not include necessary and complete demographic information. The objective of this analysis is to examine the factors associated with variance in the numbers of partial completes and terminations. For example, in 2002, the percentage of all complete or partially completed interviews that were partial completes varied from 0 to 13 percent among states. This study will examine the number and rates of partial completes and terminations from the BRFSS for the years 2002-2004 to determine the factors associated with selected respondents who begin interviews failing to fully complete an interview. Data will be examined at the state, contractor, annual, and national levels. This analysis will examine factors affecting the distribution of 120 and 210 disposition codes by state, including questionnaire length, questionnaire order, demographic variation (e.g. age, sex, race) by individual respondents or in a state's population, interviewer data quality, reliable assignment of disposition codes and other factors associated with non-response. A major part of this analysis is an item by item examination of the termination/hang up rate for each question in the BRFSS core.

***Avoiding Refusal Training: comparative work and indication of long term impact*, Mark McConaghy, member; mark.mcconaghy@ons.gov.uk, and Siobhan Carey, Member; siobhan.carey@ons.gov.uk**
Avoiding Refusal Training (ART) pilots to improve unit level response have shown positive results (Groves and McGonagle, O'Brien and Mayer). In a paper presented last year ONS showed that interviewers undergoing a similar type course in the UK showed significant improvement in co-operation of around 9 percentage points (McConaghy and Carey). Two tantalising questions remain; firstly, could any form of training intervention have the same impact and secondly how sustainable would this impact be in the longer term? To test the hypothesis that a learning event, less focused on developing practical skills, could produce the same improvement, ONS ran a second pilot with a fresh set of interviewers. The placebo form of training consisted of a workshop event with discussion in small groups to promote a knowledge exchange. As with ART, interviewers were monitored for three months following training and a shadow sample was issued to a group of control interviewers for the same geographical areas. Results suggest that there was some improvement in response for the treatment group but this was less than that observed for ART - with a gross improvement of around one third of that observed for ART. Analysis of both experiments indicates that the rigorous 'drilling' of interviewers in role-plays and 'rapid fire' sessions does indeed have a greater impact than more passive learning activities. The ART group of interviewers were followed up after the initial trial to assess if their

performance improvement was sustained. In the period after the trial, the performance of those who benefitted from ART training has remained higher than that of their colleagues who received no additional training and higher than those interviewers who took part in the workshop event. This answers the second important question that ART training does provide a skills improvement which is sustainable over time.

***A Successful Conversion or Double Refusal: A Study of the Process of Refusal Conversions in Telephone Survey Research*, Kana Fuse, The Ohio State University; fuse.2@osu.edu, and Dong Xie, University of Central Arkansas; dxie@uca.edu**
Refusal conversions have been an important part of the data collection process to improve response rates. However, there have been very few empirical studies concerning the process of refusal conversions. This study investigates various aspects of the refusal conversion process and identifies correlates of a successful conversion. With this knowledge, we can train interviewers to adopt strategies to convert cases more effectively. Data are based on information collected from three types of forms filled out by interviewers at the OSU Center for Survey Research in 2002 across three distinct telephone surveys. Interviewers filled out a "Refusal Report Form" when the respondent/household member refused to do the survey at the initial contact. During the refusal conversion phase, interviewers completed a

"Refusal Conversion Report Form" if they converted a case or a "Double Refusal Report Form" if they were refused again. This yielded 314 cases (157 refusal conversions and 157 double refusals). Chi-square tests and t-tests indicated that if the initial refusal was due to an immediate hang-up or time-related concerns/excuses, it is more likely to result in a double refusal at the second call attempt. Results also suggested that conveying information that justifies the legitimacy of the survey and revealing the previous call history were effective in refusal conversions. Moreover, we found that greater degree of interaction between the interviewer and the respondent/household member is associated with a successful conversion. Subsequent analysis will focus on establishing models to predict refusal conversation outcomes (successful conversion or double refusal) through logistic regression. We investigate interaction effects between the degree of interaction and the type of information about the survey revealed on the outcome. We also examine whether respondent's age and gender affect the outcome. We are particularly interested in how respondent's gender and age interact with other predictors to produce a successful conversion.

***Evaluating Methods for Increasing Physician Survey Cooperation*, Murrey Olmsted, RTI International; molmsted@rti.org, Emily McFarlane, RTI International; emcfarlane@rti.org, Joseph Murphy, RTI International; jmurphy@rti.org, and Craig Hill, RTI International; Chill@rti.org**
Conducting surveys with physicians is different from conducting surveys in the general population. Physicians have demanding work schedules, participating in a survey is often not seen as a priority, and they are frequently approached about taking part in surveys or other research, which can make them even more reluctant to participate. When confronted with a survey, many will refuse to respond, while others will agree to participate only after multiple persuasion attempts. In addition, physicians typically have a number of "gatekeepers," such as receptionists, administrative staff, or nurses who protect them from unwanted intrusions. Over the years, a number of effective methods have been developed to increase the level of survey participation among many populations of study. These methods include pre-notification letters, incentives, reminders, survey form design, endorsement letters, sending additional surveys, shortening survey length, and others. Each of these methods has been demonstrated to have a positive impact on the level of participation in surveys with the general population. However, the success of many of these methods has not been adequately tested with special or elite populations such as physicians who typically have low rates of survey participation. A key question for researchers considering methodological improvements is the cost to benefit ratio. Essentially, the question is which method will yield the best response at the lowest cost. To address this issue, researchers at RTI International conducted several experiments within a national survey of board certified physicians conducted for U.S. News & World Report. This paper assesses the cost and utility of differential reminders (post cards vs. 1st class letters) and length of survey (short vs. long forms) on survey response. The presentation will describe relevant literature, cover the findings from the experiments embedded within this year's study, and discuss possible future directions in survey research with physicians.

INTERNET AND WEB SURVEYS I

***Using the Web to Survey College Students: Institutional Characteristics That Influence Survey Quality*, Scott D. Crawford, Survey Sciences Group, LLC; scott@surveysciences.com, Sean Esteban McCabe, University of Michigan; plius@umich.edu, and Karen Kurotsuchi Inkelas, University of Maryland; ki21@umail.umd.edu**
The Web has gained acceptance as a tool for survey researchers. It has also been found to have introduced a whole new gamut of potential cost and error trade-offs, and re-introduced long known trade-offs. As Web survey quality is explored and the Web is used to study new ideas and populations, there is at least one population where the Web survey has been found to be a very effective tool: college students. However, there remains significant work to fully understand how to most effectively use the Web when conducting survey research among college students. In our experiences with several multi-campus student surveys, the same implementation process and questionnaire has resulted in significantly different rates of response and survey completion from one campus to another. During the spring of 2004, the National Study of Living-Learning Programs (NSLLP) was conducted among undergraduate students at thirty-four college campuses. Response rates (AAPOR RR2) for each campus varied from 17% to 52%. Simultaneously, school administrators at each campus were surveyed about their institutions adoption of Internet technologies, student use of computing facilities, the general culture of email use at their campus, as well as other activities outside of the control of the research team that individual schools may have done to promote the NSLLP. In this presentation, we will describe what we found to be key characteristics about a college environment that results in high response rates to a Web survey. We will also describe what other institutional efforts were found to be effective at increasing response rates. We will discuss how our results may be used to tailor a Web survey data collection to fit specific campus characteristics.

***What They See Is Not What We Intend-Gricean Effects in Web Surveys*, Ting Yan, Joint Program in Survey Methodology; tyan@survey.umd.edu**
Prior studies have shown that many incidental features of survey questions create pragmatic effects-inferences about the meaning of the questions that lead to unwanted measurement error. As cooperative communicators who observe Grice's Cooperative Principle and its associated maxims, survey respondents sometimes read between lines and use every piece of information (including purely formal features like the numbers assigned to the response categories) to understand and answer survey questions. This study examines the impact of one

formal feature of web survey questions-the physical arrangement of the questions on web pages-on survey respondents' inferences and responses. We presented a set of questions in a web survey either 1) one question per screen, 2) all on the same screen, or 3), in a grid on a single screen. I hypothesized that if respondents applied the maxim of relation during the survey response process, they would infer from the use of a grid that these seemingly irrelevant questions were related to each other. As a result of such a 'relatedness' inference, we would observe higher intercorrelations among the questions when they were in a grid. By contrast, the 'relatedness' inference would be weaker when question items were presented one question at a time on separate screens. To dampen and strengthen the possible effects of the physical layout of the questions, I varied the introduction to the items. One introduction indicated that the questions were taken from the same source (thus, encouraging respondents to apply the relation maxim); a second introduction indicated the items came from different sources so as to discourage the application of the relation maxim. Implications for using grids and instructions in designing web surveys will be discussed.

***Visual Context Effects in Web Surveys*, Mick Couper, University of Michigan; mcouper@umich.edu, Frederick Conrad, University of Michigan; fconrad@isr.umich.edu, and Roger Tourangeau, Joint Program in Survey Methodology; rtourangeau@survey.umd.edu**

There are many examples of context effects in survey measurement. Context can be shaped by the order of questions, the format of response options, manipulation of the survey environment, and so on. For Web surveys, the inclusion of visual images is a trivial design issue, but may have consequences for the responses obtained. Understanding the role that images may play in shaping respondents' answers is an important step in designing effective online surveys. We have conducted a series of experiments examining how context may be shaped by the use of images in Web surveys. For example, we examined the effect that pictures of a healthy woman exercising versus a sick woman in a hospital bed may have on responses to a question on self-rated health. In an effort to understand how pictures may influence respondent's answers, we have replicated the experiments in three different surveys, varying such factors as the size of the image (large versus small), the placement of the image (in the header versus alongside the question versus on a prior screen), and the position of the question (first question versus later in the survey). We have also conducted experiments on questions of mood or depression, and on behavioral questions such as travel. We report on the results of these experiments, and discuss the implications for Web survey design.

***Interactive feedback can improve quality of responses in web surveys*, Frederick Conrad, University of Michigan; fconrad@isr.umich.edu, Mick Couper, University of Michigan; mcouper@umich.edu, Roger Tourangeau, Joint Program in Survey Methodology; rtourangeau@survey.umd.edu, and Mirta Galesic, Joint Program in Survey Methodology; mgalesic@survey.umd.edu**

Because of its interactive character the web may promote more accurate survey data than do other, more static, modes. One way that interactivity can increase data quality is by providing feedback to respondents about their answers. For example, if questions require that multiple answers add-up to a fixed total, e.g. 24 hours or 100 percent, then inviting respondents to revise answers that do not sum to this figure could increase the number of appropriate sums. We report a study that compares respondents' estimates of percent time spent on each of 10 randomly ordered categories of internet usage when they did and did not receive feedback. 1788 web respondents were randomly assigned to one of three conditions: No feedback, Server-side feedback indicating that the submitted responses did not tally to 100%, and Client-side feedback in which a running tally was incremented as each response was entered prior to submission of all responses. Client-side feedback was also accompanied by a server-side message if necessary. The server-side feedback was displayed once, after which the system accepted all answers. The results show a clear advantage for Client-side feedback. More answers tallied to 100% in the Client-side than Server-side condition with the smallest number in the No Feedback condition, and responses were faster in the Client-side than Server-side condition suggesting that the running tally promoted revision before respondents submitted their answers. Client-side feedback seemed to promote more thoughtful revisions as well: when the Server-side feedback indicated the sum did not equal 100%, respondents who had not been given the running tally changed answers to the first category most often - regardless of what it was - but respondents who were presented the running tally changed answers equally often for all categories. We conclude by discussing the promise of interactive features in general for survey data quality.

MULTIPLE AND MIXED MODES

***Data Quality Issues in a Multimode Survey*, Claire Wilson, Mathematica Policy Research; cwilson@mathematica-mpr.com, Debra Wright, Mathematica Policy Research; dwright@mathematica-mpr.com, Tom Barton, Mathematica Policy Research; tbarton@mathematica-mpr.com, and Paul Guerino, American Institutes for Research; pguerino@air.org**

Declining response rates to telephone surveys have prompted survey organizations to be creative and open-minded concerning implementing new survey procedures and considering alternative data collection strategies. One strategy that may be helpful to slow or reverse the trend in declining response rates is to offer sample members a variety of ways to respond to the questionnaire-essentially allowing the respondent to choose the mode with which he/she is most comfortable. For example, adding a mail option, a Web option, or both to a telephone survey creates an alternative method for obtaining responses from willing participants, thereby allowing more resources to be devoted to reluctant sample members and late responders. This paper describes some of the practical issues encountered in designing and implementing a multimode survey and presents data on the impact of using multiple modes on data quality. The National Survey of Recent College Graduates (NSRCG) is a biennial survey of science and engineering bachelor's and master's degree recipients sponsored by the National Science Foundation. Conducted as a CATI survey since 1974, in 2003 the NSRCG was converted to a multimode survey using mail, Web, and CATI. In this paper, we discuss some of the challenges faced in converting the NSRCG to a multimode survey, such as revising question wording and response formats to suit each mode, managing multiple instrument versions, and maintaining a common database. We then examine measures of data quality by mode, including item nonresponse and response error (as measured by rates of editing). We conclude with lessons learned and suggestions for researchers implementing a multimode design.

***Dual-Frame, Dual-Mode Designs to Improve RDD Survey Response Rates: An Initial Feasibility Study*, Jill Montaquila, Westat; jillmontaquila@westat.com, Mary Hagedorn, Westat; maryhagedorn@westat.com, Michael Brick, Westat; mikebrick@westat.com, Shelley Brock Roth, Westat; shelleybrock@westat.com, Priscilla Carver, Westat; priscillacarver@westat.com, and Chris Chapman, National Center for Education Statistics; chris.chapman@ed.gov**

Random digit dial (RDD) survey response rates have been declining despite increasing levels of effort in the form of extended calling protocols, advance and refusal mailings, and respondent incentives. While some recent research suggests that late responders and nonrespondents are not very different from survey respondents, concerns about the negative effect of nonresponse bias on survey quality remain. If significant improvements in survey response rates are to be achieved, or telephone nonresponse bias is to be more thoroughly evaluated, new approaches must be considered. This paper presents the findings of an initial feasibility study using a dual-frame, dual-mode approach with in-person follow-up to complete interviews with telephone nonrespondents. The feasibility study was conducted during the field test of a large national RDD program and focused primarily on the household screening stage, at which most RDD nonresponse occurs. Conducted in four selected counties, the feasibility study used two designs, called forward and reverse designs. In the forward design, an address sample was selected in the sites and the addresses were matched to telephone numbers where possible. In the reverse design, a list-assisted RDD sample was selected and the telephone numbers were matched to addresses where possible. In both designs, the first phase of data collection was conducted by telephone. Advance and refusal mailings were used at the telephone interviewing stage. In-person data collection was attempted with address sample cases with no telephone number match, and with cases from both samples that refused the telephone survey or were not completed after many telephone attempts. The presentation will cover the design of this study, the operational issues encountered, and the results of the in-field efforts to complete interviews. The potential of this approach for response improvement and for the assessment of telephone nonresponse bias will be considered.

***Web Survey Methodologies: A Comparison of Survey Accuracy*, Jon Krosnick, Stanford University; krosnick@stanford.edu, Norman Nie, Stanford University; nhnie@stanford.edu, and Douglas Rivers, Stanford University and Polimetrix, Inc.; doug@polimetrix.com**

The estimation accuracy of Web surveys is assessed using an identical instrument fielded by seven Web survey vendors as well as a similar instrument fielded using an RDD telephone sample. The items used for comparison include demographic and other variables, for which high response rate survey measurements exist, and some other items typical of commercial market research, for which lower response rate benchmarks are available. The results show considerable variation in accuracy across items and vendors.

CELLULAR PHONES II

***The Prevalence and Impact of Wireless Substitution: Updated Data from the 2004 National Health Interview Survey*, Stephen Blumberg, National Center for Health Statistics, CDC; sblumberg@cdc.gov, Julian Luke, National Center for Health Statistics, CDC; jluke@cdc.gov, and Marcie Cynamon, National Center for Health Statistics, CDC; mlc6@cdc.gov**

The National Health Interview Survey (NHIS) provides data for current estimates of the prevalence of persons who have substituted a wireless telephone for their residential landline telephone. Conducted by the National Center for Health Statistics of the Centers for Disease Control and Prevention, the NHIS is an annual face-to-face survey that collects comprehensive health-related information from a large sample of households representing the civilian noninstitutionalized household population of the United States. At the 2004 AAPOR conference, we reported 2003 estimates that approximately 3.1% of U.S. civilian noninstitutionalized adults did not have a landline telephone but had a wireless telephone. Only 1.9% of adults did not have any telephone service (landline or wireless). For the 2005 AAPOR conference, we will update these analyses with data from 2004, and focus these analyses on the changing demographics of this dynamic group. This presentation will include the percent of households with only wireless service and the percent of persons living in households with only wireless service. Household prevalence estimates will be presented by household size, family composition, income, home ownership, and region. Person-level prevalence estimates will be presented by age, sex, race/ethnicity, education, and employment status, as well as by household size, family composition, income, home ownership, and region. To better understand the implications of wireless substitution for RDD household health surveys, this presentation will also include NHIS estimates on the health and health care access of persons with and without landline telephones. Particular attention at AAPOR 2005 will be paid to the health and health care access of persons living in poverty because more than 13% of them do not have landline telephones.

***Non-interviews in Mobile Phone Surveys*, vasja vehovar, universtiy of ljubljana; vasja@ris.org**

Paper addresses the issue of noncoverage and nonresponse patterns in mobile phone surveys. One aspect of this problem is the actual availability. Person may have a mobile device, but it can be most of the time turned off. We have similar situation also with fixed telephones: they may exist in households, however they are used only for the Internet access. Further specific are arising from the patterns of usage or from the corresponding attitudes of the users. Here, the main purpose of the paper is to identify the key segments of mobile phone users that have specific nonresponse behavior. The empirical data are based on Slovenian Labor Force Survey 2002-2004, where the fixed and mobile phone usage is systematically monitored. First, the apparent stagnation in the share of mobile-only household in 2004 (after the growth in 2003) is analyzed. Next, the experience with follow-up telephone surveys for later waves of LFS is studied: the persons that revealed the mobile phone are compared with the ones that revealed the fixed telephone. The differences compared to the corresponding experience from Scandinavian countries are discussed. Within the context of LFS, the frequency of mobile phone usage was analyzed with questions replicated in Slovenian LFS from the 2003 US Current Population Survey. Available comparisons are drawn. Finally, the key nonresponse segments are analyzed from a specially designed 2004 survey that studied the attitudes and usage patterns of mobile phone users.

***Quality Assessed: Cellular Phone Surveys versus Traditional Telephone Surveys*, Charlotte Steeh, Georgia State University; cgsteeh@gsu.edu**

The main question addressed in this paper relates to the quality of the data collected via mobile telephone. We ask if it is possible to overcome the difficulties that wireless telephones present in order to collect data of substantive value. Three data sets will be compared. The first two consist of results from identical RDD surveys conducted nationally in 2003, the first carried out by cellular telephone and the second by conventional telephone. The third set comes from the replication in 2003 by the Pew Research Center of their landmark 1997 response rate experiment. To judge by response rates-a standard measure of quality, we would not have much confidence in the matching

landline and cellular studies, especially the cellular survey. Both suffer from response rates in the twenty to thirty percent range. The rigorous Pew survey achieved a comparatively high rate of fifty-one percent. Thus it will serve as the gold standard by which we assess the other two surveys. The analysis will focus on the relationships among variables. Preliminary research suggests that few marginal differences in attitudinal and behavioral measures across the three surveys reach statistical significance. However, no study has extensively examined differences in relationships either by mode or by response rate. The matched landline and cellular surveys replicated attitudinal questions from the Pew response rate experiment. For each of these items we first describe the variables that theory and research indicate have the greatest explanatory power. We then apply an appropriate multivariate model to the Pew data. In the last step, we fit the same multivariate models to data from both the cellular and landline surveys. Our purpose is to explain how well these results conform to the theoretical models, paying particular attention to what any differences may reveal about the quality of data collected via mobile phone.

PRE-ELECTION SURVEYS II

***Re-examining Our Assumptions in Pre-Election Polls*, Randall Thomas, Harris Interactive; rthomas@harrisinteractive.com, David Krane, Harris Interactive; dkrane@harrisinteractive.com, John Bremer, Harris Interactive - HI Europe; jbremer@hieurope.com, Humphrey Taylor, Harris Interactive; htaylor@harrisinteractive.com, and Renee Smith, Harris Interactive; rsmith@harrisinteractive.com**

We examined a variety of both likely voter and weighting models from both Election 2000 and Election 2004 to determine factors that would have improved prediction in both online and telephone interview modalities for both elections. In the 6 weeks prior to the 2004 U.S. general election, we fielded 7 different national online studies (each with a minimum of 2000 respondents) and 2 national phone studies (1000 respondents each wave). We had 3 different likely voter screening criteria in the 7 different online studies. While most any likely voter screen or weighting led to improvement in our 2000 general election online predictions (our final prediction was among the most accurate nationally), few models would have actually led to an improvement in the accuracy of our 2004 general election online predictions because the raw, unweighted online data were so close to the final election outcome. We also examined different mixed-mode combinations that may help provide greater confidence in our inferences when pooling data from multiple modes of interviewing.

***A Dispositional Model of Election Polls*, David Dutwin, ICR/International Communications Research; ddutwin@icrsurvey.com, and Melissa Herrmann, ICR/INternational Communications Research; mherrmann@icrsurvey.com**

Presidential election cycles generate a great deal of "seasonal" interest in the quality of political opinion polls. Typical areas of investigation include the refinement of likely voter screens, the improvement of exit poll methodology, and comparisons of various firm's pre-election polling data. The present paper proposes an alternative analysis by taking a major national post-election poll and constructing a model of partisanship on a wide range of dispositional information. Included in the model are the various "avenues to completion," for example, a completed interview that was the result of a refusal conversion attempt. Other dispositions that can lead to completion include callbacks, both scheduled and unscheduled, privacy managers, language barriers, no answers, busy signals, and answering machines. All of these dispositions can furthermore interact with another key variable, the number of call attempts (the survey made as many as 15) it takes to reach completion. The paper, therefore, constructs probability models for party identification and voting behavior, as well as for key demographic variables and other indicators, based on these critical dispositional variables. The implications of the findings are far ranging, by not only providing information on the systematic nature of respondent demographics by call disposition and number of attempts, but also insight into best practices and survey quality regarding the reliability and validity of political public opinion polls in the context of a Presidential election.

***A New "Interval" Measure of Election Poll Accuracy*, Jay Leve, SurveyUSA; jleve@surveyusa.com, and Joseph Shipman, SurveyUSA; jshipman@surveyusa.com**

Measures of election poll accuracy, starting with Mosteller's pioneering work in 1949, have always involved "point estimates" of error, using a single set of predicted vote percentages. A known drawback of any such measure is that it must ignore "undecded" voters or make an arbitrary assumption about how to allocate them. A new measure is proposed, which uses an "interval estimate" to account for every possible allocation of undecided voters, rather than choosing to ignore undecided voters or allocate them arbitrarily. Advantages and disadvantages of this measure are discussed.

***The Performance of the Polls in the 2004 Presidential Election*, Michael Traugott, University of Michigan; mtrau@umich.edu**

The 2004 presidential election campaign provided a venue for a wide variety of polling, and it was not without its controversies. In the end, the final estimates of the pre-election polls, the bread and butter of the polling industry, were very good at suggesting it would be a close race with Bush the likely winner. In historical perspective, the overall performance was one of the best since 1956. There were issues leading up to the final estimates, however, that created controversy, especially about the likely voter methodology used by different organizations. The main issue for pollsters in 2004 was the debate about the leaking of early exit poll results and their use to characterize the election outcome early in the evening of Election Day. This created enough controversy to diminish the evaluations of the performance of the pre-election polls. The procedures for collecting and analyzing exit poll data - and for the dissemination of the results to the sponsoring parties - will be reviewed and receive continued attention in the next several months. Changes in these procedures will undoubtedly take place before the 2006 and 2008 elections are held. The American electoral system is entering a period of dynamic technological change with the passage of the Help America Vote Act (HAVA) of 2002. As the voting procedures and devices change and more opportunities to vote early arise, pre-election and exit pollsters will face new challenges for the standard telephone surveys of voters at home and face-to-face interviews with voters leaving the polls. More public disclosure of methods and their consequences will be required to maintain public confidence in the profession, as well as in the basic foundation of the American electoral system and its transparency to public.

HEALTH SURVEYS II

***Quality of Children's Height, Weight, and BMI information in the MEPS*, Frances Chevarley, AHRQ; fchevarl@ahrq.gov**
This presentation/poster summarizes item non-response for height, weight, and Body Mass Index (BMI) information for children ages under 18 in the Medical Expenditure Panel Survey (MEPS). A file with height and weight information that is used to produce the BMI, although not available from the MEPS public use files because of confidentiality concerns, is available in the AHRQ Data center. An initial evaluation of the Height in Feet data suggests higher levels of non-response for young children and this decreases by increasing age of the child. Because of the desire to be able to interpret and use the MEPS BMI information for children, further evaluation is warranted in assessing the quality of the existing BMI information and of the height and weight variables that are used to produce the BMI for children, and for determining suggestions to improve these data.

***Comparing Administrative and Survey Data to Identify Adults with Chronic Conditions*, Jennifer Dunne, Center for Survey Research; jennifer.dunne@umb.edu, Patricia Gallagher, Center for Survey Research; patricia.gallagher@umb.edu, and Vickie Stringfellow, Center for Survey Research, University of MA Boston; vickie.stringfellow@umb.edu**
OBJECTIVE We will compare survey data to administrative claims data to examine the rate at which adults with chronic conditions are identified by responses to a parsimonious set of self-identifiers. Identifying respondents with chronic conditions is useful as a means for making case mix adjustments when comparing health plans. Individuals with chronic conditions are likely to be higher users of health care and there is evidence that they are more likely than low users to give negative reports and ratings of their health care. METHODS We collected survey data in the Massachusetts Medicaid population using the CAHPS® (Consumer Assessment of Health Plans) instrument to measure health care experiences. A probability sample of adult enrollees was drawn (n=1124) and 564 respondents ultimately completed the survey. Survey data were collected by mail or telephone using an instrument that included a four-item identifier for adults with chronic conditions. Administrative data used in the analysis includes International Classification of Diseases (ICD-9-CM) codes for the entire sample for the 24 month period prior to the date the sample was drawn. We will examine the rate at which adults with chronic conditions are identified by comparing screener question results with diagnoses codes in the claims data. RESULTS Through these analyses, we will be able to address the following research questions: Which diagnoses of chronic conditions are identified using the chronic condition screening questions? Which diagnoses of chronic conditions are missed using the chronic condition screening questions? How do the proportion of those identified as having a chronic condition compare with recognized lists of the prevalence of chronic conditions?

***An Assessment of the Quality of Health Data Collected from Multiple Sources: Registries, Provider Records, and Household Reports*, Meena Khare, NCHS/CDC; mxk1@cdc.gov, Linda Piccinino, Abt Associates Inc; Linda_Piccinino@abtassoc.com, and Michael Battaglia, Abt Associates Inc.; Mike_Battaglia@abtassoc.com**
Incomplete or missing data due to unit or item-nonresponse in a survey generally impact the survey quality. Also, collecting the same information from multiple sources in a survey is likely to yield differences in reporting of key outcome measures. The CDC-sponsored National Immunization Survey (NIS) is designed to measure and monitor vaccine-specific coverage rates among children aged 19-35 months in the United States. NIS, a large on-going telephone survey, consists of a list-assisted random-digit-dialing (RDD) telephone survey of the parents/guardians of the eligible children, followed by a mail survey of the children's immunization providers to obtain immunization histories from medical records. Approximately 45% of the children with complete household interviews have vaccination data from a written shot card. In 2002, we implemented a pilot study to investigate completeness and quality of the immunization histories from four immunization registries. The objective for this paper is to examine the completeness and quality of the vaccination data received for the sampled children from each of the three sources-- registries, provider records, and written shot cards. We compare vaccine-specific shot dates, number of doses and coverage rates for the vaccination series 4:3:1:3 [4 or more doses of diphtheria and tetanus toxoids and pertussis vaccine (DTaP), 3 or more doses of poliovirus vaccine, 1 or more doses of measles measles containing vaccine (MCV), 3 or more doses of Haemophilus influenzae type B vaccine (Hib)]. We also assess the feasibility of developing 'best value' estimates by supplementing provider-reported vaccination data from other two sources.

***Assessing the Impact of Medicaid Bias on Counts of the Uninsured*, Robert Goidel, Louisiana State University; kgoidel@lsu.edu, Dek Terrell, LSU; mdterre@lsu.edu, Steven Procopio, Louisiana State University Public Policy Research Lab; sproco1@lsu.edu, and Douglas Schwalm, Louisiana State University; dschwalm@lsu.edu**
Surveys attempting to measure the number of insured consistently find that respondents underreport Medicaid coverage. In a study of Minnesota Medicaid recipients, Call, et. al., for example, find a number of reasons for the undercount and conclude that the undercount has a negligible impact on estimates of the uninsured for Minnesota. However, they also note that the study may not generalize to other locations with very different populations. In this study, we consider Medicaid underreporting in Louisiana and its implications for survey estimates of uninsured adults and children. Data for the analysis are based on a random sample of Medicaid recipients in Louisiana, allowing comparisons of reported insurance status for these Medicaid recipients. The paper will also consider the implications of Medicaid underreporting on state-level and national-level estimates (CPS data) of Louisiana's uninsured population. As a final note, with its unique charity hospital system and high poverty level among Louisiana natives, Louisiana provides an interesting and unique context for examining the problem of Medicaid underreporting.

RELIGION AND POLITICS: VALUES AND DIVISIONS IN THE 2004 ELECTION

***Religious Affiliations and the 2004 Presidential Election*, Marjorie Connelly, New York Times; connelly@nytimes.com**
Using the 2004 national exit poll that was conducted for the National Election Pool by Edison Media Research and Mitofsky International, this paper will examine how different religious groups voted for President in 2004, the differences within religious groups, and how that may have changed compared with previous elections. For all the attention paid to evangelical Christians, President Bush owes his re-election to a formula that includes conservative Catholics and Protestants, and an improved showing among Hispanics and Jews. Although John Kerry was the first Catholic nominated by a major party for president since 1960, most Catholic voters chose his opponent. Mr. Bush was supported by more than half of all Catholics, a significant change from 2000, when Al Gore won more Catholic votes than Mr. Bush did. The president also did better among Hispanic voters: from 35 percent in 2000 to 44 percent in 2004. While Hispanic Catholics backed

Mr. Kerry, Hispanic Protestants and evangelicals strongly supported Mr. Bush. African-Americans remained, regardless of religion, firmly in the Democratic column. While the Jewish vote is small - 3 percent of the electorate - the president increased his share of the Jewish vote from 19 percent in 2000 to 25 percent this year. The paper will examine the vote preferences of different groups since 1972, along with analysis of differences within religious groups.

***Evangelical Voters: Bedrock of the Republican Party*, Scott Keeter, Pew Research Center; keeters@people-press.org**
President Bush was boosted to re-election by the strong support of religious voters, especially white evangelical Protestants, a group that has become the bedrock of the Republican Party. In the 2004 election, they were the largest single demographic group among Bush voters, constituting fully 36% of his total. By comparison, African-Americans - the most loyal of Democratic constituencies - constituted only about one-fifth (21%) of Kerry's voters. Bush received 76% of the vote among white evangelicals, up 8 percentage points from 2000. This was the highest level of support for Bush among any religious group, and represented the largest increase in his vote share compared with 2000. Our paper will discuss differences between evangelical and other voters in issue concerns and criteria used in voting, and will address the controversy over the measurement of "moral values" as a voting issue. But even though voter turnout was up significantly in 2004, there is no indication that white evangelicals boosted their level of participation more than other groups in the population. According to a comparison of exit polls and Pew surveys in 2000 and 2004, white evangelicals constituted the same percentage of the electorate in both years: 24%. Overall, there was remarkable stability from 2000 to 2004 in the religious composition of the electorate. Our paper will also examine trends in partisan identification among evangelicals. Evangelical affiliation with the Republican Party has grown significantly over the past 17 years, first among southern voters and more recently among evangelicals nationwide.

***Religion, Ideology, Partisanship and "Moral Values" in the 2004 Election*, Gary Langer, ABC News; gary.e.langer@abc.com, and Jon Cohen, ABC News; jon.cohen@abc.com**
The nexus of religious, ideological and political identity gained new attention after the 2004 election, as analysts and commentators alike grappled with the meaning of "moral values" and the claim by religious conservative leaders to have been the decisive force in the re-election of George W. Bush. Are "moral values" on the rise? Are evangelical Christians wielding new clout in the electoral process? We will undertake a detailed analysis of the extent to which religious belief and practice informed choices in the presidential contest, examining differential turnout and vote preferences among religious, ideological and partisan political groups compared with past elections, and conducting regression analyses to tease out the strength of these factors in vote and issue preferences. We'll also examine the inclusion and meaning of "moral values" as an option on the issues list of the 2004 national network/AP exit poll, and delve into the political, ideological, demographic and attitudinal makeup of various religious groupings. Our data sources will include the 2004 national and state exit polls (N=13,660 and net N=76,504, respectively), exit poll data from the 2000 presidential elections, and data from the 32-day pre-election ABC News tracking poll (N=21,265, RDD).

***Religious Identification, Congregational Membership and Political Party Preference*, Ariela Keysar, Brooklyn College; akeysar@aol.com, and Barry Kosmin, Institute for Jewish Policy Research, London; bkosmin@jpr.org.uk**
The United States is unique among western democracies in both its religious diversity and its more religious public. This was very evident in the 2004 presidential campaign, yet to fully appreciate the 'religious effect' a rich data set is required. Using a large national sample of over 50,000 adults, the American Religious Identification Survey of 2001, we can provide detailed data and analysis of the role of religion in American politics. The sheer number of respondents to ARIS 2001 allows for in-depth exploration of patterns among both large and small religious groups, comparison of members and non-members, and looking at racial/ethnic groups as well as geographical clustering. Religiosity interacts with race differently across religious groups. To illustrate, the Catholic vote can be segmented into three parts. First, "cultural" Catholics, who identify themselves as Catholic but are not church members, tend to continue the historical tradition of favoring the Democrats. Second, "practicing" Catholics now tend to favor the Republicans. Third, Hispanic Catholics are more likely than non-Hispanic Catholics to favor Democrats, though here as well church members are more Republican than non-members. Asian-Americans are the least religious ethnic population and have the largest proportion of persons (35%) not identifying with a religion or with a political party. Asian Christians show a slight preference for the Democrats whereas those of Eastern religions have very low levels of support for the Republicans. These patterns warrant attention in states with large Asian population.

Saturday 10:00 am-11:30 am - AAPOR ABSTRACTS

DATA QUALITY II

***Improving survey quality in the measurement of social capital*, Abigail Dewar, ons; abigail.dewar@ons.gov.uk, and Amanda Wilmot, Office for National Statistics; amanda.wilmot@ons.gov.uk**
The Office for National Statistics, UK, developed a set of harmonised questions about social capital (a concept akin to community spirit) designed for Computer Assisted Personal Interviewing (CAPI) on national government surveys. There is interest in collecting data on social capital at the local government level. However, most local government surveys are carried out using postal self-completion methods. Different methods often produce different results and reduce the comparability between national and local data on social capital. This paper reports on two aspects of this ongoing research project. Firstly, on lessons learnt from the conversion of the established CAPI based social capital question set to a paper self-completion. Secondly, on the cognitive testing of the questions using a combination of concurrent think-aloud and retrospective probing techniques. Respondents were also videoed completing the questionnaire.

***Development of the UK Continuous Population Survey from a data collection perspective: improving survey questions and the quality of data output.*, Alison Blackwell, Office for National Statistics; alison.blackwell@ons.gov.uk, and Abigail Dewar, ons; abigail.dewar@ons.gov.uk**
The Continuous Population Survey (CPS) is an innovative project to integrate the five key government surveys on which the UK Office for National Statistics (ONS) leads, and replace them with one modular survey. A comprehensive integration of the entire survey process is proposed: from the sample design, creation of a unified field force of interviewers administering a common modular questionnaire using different data collection modes, to the processing and production of outputs from a single common source. Reduction of non-sampling

error through improved question design within a mixed mode collection environment (telephone and face-to-face) is one element of the development process discussed in this paper. The development of core questions on caring activity, educational attainment and income were identified for redesign and tested using a combination of qualitative and quantitative techniques. Investigations were carried out on existing questions to assess conceptualisation, question wording, unit and item non-response, system mode effect and proxy response. This provided an understanding of whether the questions were acceptable to interviewers, respondents and proxy responders (acceptability), whether respondents understood and completed question tasks successfully (reliability/validity), and how accurate the outputs were when comparing different modes and in comparison to other measures of caring, education and income. Redesigned questions, based on Dillman's unimode approach, were tested live in the field using a split sample design on the ONS Omnibus and Annual Population Surveys, to quantitatively evaluate the revisions made. This paper reports on the findings from these two stages: analysis of existing data, identifying areas of possible measurement error, and question validation, including examination of order effects and questionnaire flow of the redesigned questions, giving evidence of how improvements have been made.

***Data Quality and the Use of Standardized Child Assessments in Survey Research*, Michelle Ernst, NORC - University of Chicago; ernst-michelle@norc.org, and Mike Pergamit, NORC at the Univerisity of Chicago; pergamit-michael@norc.org**
Beginning with the Children of the National Longitudinal Survey of Youth-1979 cohort (CNLSY79) in 1986, large-scale surveys began to incorporate standardized assessments. Formerly used only in clinical settings or schools, these assessments are now administered in a household setting by lay field interviewers. The administration of standardized assessments to children by lay field interviewers raises data quality concerns. Standardized child assessments have rigid administration protocols. Deviation from procedure can greatly affect a child's response. Furthermore, administrative complexity varies across assessments. While some assessments consist of a very simple and straight-forward administrative protocol, other assessments rely much more on the skills of the individual conducting the administration. It is hypothesized that an administratively complex assessment with strong published psychometric properties may not maintain those properties when administered by interviewers in large-scale studies. This paper proposes examining the published psychometrics for three assessments (the Woodcock-Johnson, the PPVT, and the PIAT) and comparing the published psychometrics with its reliability and validity within single, longitudinal studies (the NLSY79 and the PSID). By using multiple years of assessment data from the Children of the NLSY79 (PPVT/PIAT) and the Panel Study of Income Dynamics Child Supplement (Woodcock-Johnson), we have access to a large number of assessments conducted by a large number of interviewers. We can compare distributions between interviewers as well as looking at the same interviewer over time. These two data sets provide a rich source of data on assessments that allows us to examine many differences in administration. It will also be possible to examine how the psychometric properties of different assessments stand up in a large-scale survey as a function of the complexity of the assessment. If interviewer variability is greater in administrations of the complex tests, this argues for greater consideration of administrative procedures when choosing assessments for large-scale survey research.

***Measuring Quality in Observational Data Collection*, Jaana Myllyluoma, Battelle CPHRE; mylly@battelle.org, and Diana Buck, Battelle CPHRE; buckd@battelle.org**
Data collection via direct observation of behavior or objects is a useful method of capturing information in a real-world situation. While well-established methods exist to measure the quality of survey data collected through interviewing or abstraction, data collected via observation by its nature is subject to observation error that requires non-standard methods to capture and quantify. This paper reports on the methods and results obtained from an assessment of reliability of unobtrusively observed information related to pricing, promotion, placement, and advertising of tobacco products in retail environments. Multiple observers were assigned to visit the same retail outlets during the same week in order to minimize actual changes in the retail environment and objects to be observed. For continuous variables, such as 'Total Number of Built-In Shelving Units,' inter-observer correspondence was calculated using correlation and intra-class correlation (ICC) analyses. For categorical variables, such as presence or absence of a specified item, Kappa statistics were generated. Findings from this study indicated good to excellent reliability for almost all total item measures per store. Good to excellent reliability was obtained when item placement was described using specific anchors, such as "next to candy displays". Reliability was poorer when observers were asked to judge placement distance, such as "within 12 inches." In addition, whereas items that required observers to discriminate one feature, such as a brand, exhibited good to excellent reliability, items that required observers to discriminate multiple features, such as a branded sign that also had the words "special price" exhibited poorer reliability. It is not surprising that items with higher cognitive burden tend to have lower inter-observer reliability. These results have implications for the design of observation protocols - including the level of detail to be ascertained, how the observers are trained, and how the data are analyzed.

TELEPHONE SURVEYS

***Study of Ported Telephone Numbers Using the NeuStar Database*, Michael Yang, University of Chicago; yang-michael@norc.net, Manas Chattopadhyay, University of Chicago; chattopadhyay-manas@norc.net, and Ned English, NORC, University of Chicago; english-ned@norc.net**
To comply with the Telephone Consumer Protection Act of 1991 (TCPA), survey organizations need to remove wireless telephone numbers, including those ported from wireline to wireless, from their telephone samples/frames. This causes undercoverage to the extent that wireless users cannot be reached through a conventional wireline number. This study attempts to understand the nature of the undercoverage by providing a general description of the households that ported telephone numbers from wireline to wireless based on the NeuStar database. To profile the households who have ported their numbers from wireline to wireless, we appended relevant census data to the ported numbers. Census data are available only at the aggregate levels corresponding to census geographies. Our strategy was to describe the distribution of the porting population by the relevant characteristics of the census geographies where they are physically located. First, we matched the wireline-to-wireless ported numbers to postal addresses using a vendor. Second, we obtained the geographical coordinates for every matched address through a process known as geocoding. Third, we matched the geographical coordinates to census geographies. Fourth, we abstracted relevant census data from the census summary files for matched census geographies. After the census data were obtained, we examined the distribution of the ported numbers with respect to geographic region, race, age, income, education, occupation, employment, poverty status, and citizenship status. We focused on ported residential numbers that were being removed from the RDD frame by describing the geographical distribution of ported numbers by census region, division, and urban and rural status, and by examining the demographic, socioeconomic, and other characteristics of the households that have ported their telephone numbers from wireline to wireless.

***Purging Out-Of-Scope And Cellular Telephone Numbers From RDD Samples*, Michael Battaglia, Abt Associates Inc.; Mike_Battaglia@abtassoc.com, Meg Ryan, Marketing Systems Group; MRyan@M-S-G.com, and Marcie Cynamon, National Center for Health Statistics, CDC; mlc6@cdc.gov**

At the 1995 AAPOR conference Battaglia et al. presented one of the first papers on the use of automated procedures to remove business and nonworking telephone numbers from random-digit-dialing samples to increase interviewer productivity while at the same time improving the resolution rate. Since then, advances have been made in methods for purging out-of-scope numbers from RDD samples. Sample purging results from the National Immunization Survey will be presented for 78 geographic areas covering the entire U.S. These sample-purging techniques set aside residential directory-listed telephone numbers on the assumption that all are residential numbers. The latest sample purging methods also process residential directory-listed numbers through the automated procedure, resulting in some of these numbers being classified as out-of-scope. Based on a recent experiment, validation (false negative) results are given for this component of the sample. Another aspect of sample purging that is of growing importance in RDD samples is cellular telephone numbers. FCC regulations specify that cell numbers not be dialed using automated dialing devices. Cellular telephone numbers fall into three categories: 1) numbers in dedicated cellular telephone exchanges, 2) cell numbers in mixed-use exchanges, and 3) land-line numbers that have been ported to cellular service. Procedures for identifying and removing cellular telephone numbers from RDD samples will be discussed.

***Testing the Impact of Caller ID Technology on Response Rates in a Mixed Mode Survey*, Norman Trussell, Nielsen Media Research; norman.trussell@nielsenmedia.com, and Paul Lavrakas, Nielsen Media Research; pjlavrakas@tvratings.com**

Many factors effect survey response rates and one is thought to be a respondent's awareness of the survey organization's name. In RDD telephone surveys, if a respondent has Caller ID, it may help a survey company to display its name. In 2004, we conducted a test involving Caller ID technology in Nielsen's RDD diary recruitment calling process. The objective was to determine the impact of displaying our name and phone number on the respondent's Caller ID display on the (1) telephone contact rate, (2) telephone response rate, and (3) subsequent mail survey response rate. Recent events have reinforced the importance of exploring the use of this technology, including: 1) the increasing popularity of home Caller ID use, 2) the increasing use of Caller ID as a screening device, 3) the proliferation of telemarketing calls, 4) the DNC list, 5) the recently enacted federal law that requires telemarketers to make a Caller ID number available (which respondents may be unaware that researchers are not subject to), and 6) new services which do not allow "out of area" numbers to even ring the phone unless the caller provides additional information. The telephone stage of the experiment was conducted in mid 2004. The Caller ID treatment was randomly assigned to 108,368 RDD telephone numbers throughout the entire U.S. Numbers in this sub-sample that reached households with Caller ID (estimated at approx. 30 - 40% of all U.S. households) had the words, "Nielsen Ratings" displayed. Another randomly assigned 242,173 RDD numbers served as the control group for which the Caller ID treatment was not used. The test results showed a significant increase in telephone response rates of 2.5 PP, but there was no significant impact on the response rate at the subsequent mail survey stage. These and other findings will be discussed.

***Using Multiple Sources of Contact Data to Improve Quality in Telephone Surveys*, Jason Markesich, Mathematica Policy Research; JMarkesich@mathematica-mpr.com, Anne Ciemnecki, Mathematica Policy Research; ACiemnecki@mathematica-mpr.com, Nuria Diaz-Tena, Mathematica Policy Research; NDiaz-Tena@mathematica-mpr.com, and Anne Hower, Mathematica Policy Research; AHower@mathematica-mpr.com**

Survey respondents often are selected from program participants' records. However, because the quality of contact information on these records varies, they may or may not contain the telephone numbers of participants selected to participate in the survey. To supplement missing or poor contact information, Mathematica Policy Research, Inc. (MPR) will use a commercial vendor (such as Genesys or Accurint) or some other source of administrative records, usually one from a government program other than the one under study. Traditionally, such supplementation has been used for respondents who prove unlocatable during the field period. More recently, however, we have been able to supplement contact data for all sample members at the beginning of the field period. For the second round of the Targeted Beneficiary Survey (TBS), and prior to the start of data collection, we combined data from the Centers for Medicare & Medicaid Services' (CMS's) Medicare Enrollment Database (EDB) file, data that do not contain telephone numbers, with information from commercial vendors and the Social Security Administration (SSA) records. (CMS executed an interagency agreement with SSA to append telephone numbers to the TBS survey sample.) This paper will use the TBS as a case study to demonstrate that, by using a commercial vendor and/or secondary source of administrative data early in the field period to supplement contact data, the quality of the telephone survey can be improved, along with the reduction of potential bias and the level of effort required. In addition, we will, one, analyze the quality of the telephone numbers supplied by each source, two, present response rates by source of telephone number, and, three, assess the level of effort of the telephone interviews according to the presence or absence of a telephone number.

SCHOOL AND YOUTH SURVEYS

***Re-examining Approaches to Achieving High Response Rates on Web-based Surveys of Post-secondary Students: Results from a Follow-up Survey*, Bronwyn Nichols Lodato, National Opinion Research Center at the University of Chicago; Nichols-bronwyn@norc.org, and Rashna Ghadialy, National Opinion Research Center; ghadialy-rashna@norc.net**

This paper is an update and expansion of a 2003 AAPOR Conference paper "Achieving High Response Rates on Web-based Surveys of Post-secondary Students" (Nichols & Ghadialy). Using results from NORC's baseline, web-only survey of post-secondary minority students, the authors showed that Don Dillman's Tailored Design Method (TDM) (Dillman, 2000) was very effective in achieving high response rates on web-based surveys. In 2004, the first-follow-up survey in this longitudinal series was conducted by NORC. In this updated paper we will present how effective TDM was in increasing response rates by reviewing prompting efforts and survey results. We will look at the average number of contacts needed to yield one completed web survey, and discuss prompting innovations employed via mail, telephone and email to reach various population types to encourage their participation. When examining study outcomes, the paper will compare the demographic characteristics of respondents who responded in both the baseline survey and the first follow-up to those respondents who only participated in the baseline survey. Furthermore, we will present the sample attrition rate for the first follow-up study and compare these rates against other longitudinal surveys of post-secondary students.

First, a pre-notification was sent by mail, followed by a letter directing the hunter to the Web site where the questionnaire was located. A follow-up postcard was sent by mail. The final contact for this group was a mailing that included a paper copy of the questionnaire. In the letter directing hunters to the Web site, a postcard was included asking the respondent to return it if he/she did not have access to the Internet. Those returning the postcard received a mail version of the questionnaire. Demographics for the two groups of 2,000 were evaluated to compare response rates across age and gender. In addition, similarities between groups in their responses to questions were analyzed to evaluate whether responses from the Web group differed from those in the mail group. The group receiving all contacts by mail had the highest response rate in this study. The results also suggest that the older age group were more slightly more likely to respond by mail rather using the Web.

***Effectiveness of E-mail and Paper Mail Notifications for Internet Surveys*, Paul Ruggiere, University of North Texas, Survey Research Center; paulr@unt.edu, and D'Arlene Ver Duin, University of North Texas, Survey Research Center; dverduin@scs.unt.edu**
Successful use of e-mail to notify potential respondents of an Internet survey partially depends upon the completeness and accuracy of the e-mail database. In the case of a customer satisfaction survey, when a client's database contains e-mail addresses on only half of its customers, selection bias may occur if only the customers with e-mail addresses are asked to participate. A state agency proposed a customer satisfaction survey using its list of superintendents, business managers and principals within the school districts it serves. Approximately half of the records in the agency's database included an e-mail address for the contact person on record. All records included a postal mailing address. The agency estimated that nearly all of the potential respondents had Internet access in the office even though the agency did not have a complete listing of e-mail addresses. To be sure that all customers had the opportunity to participate, potential respondents were notified about the Internet survey in one of two ways. Respondents with e-mail addresses on file were sent an e-mail notification and respondents without an e-mail address were mailed a letter notifying them about the online questionnaire. A follow-up notification was sent about two-weeks later to most non-responders using the same method as used in the first wave. A subset of randomly selected non-responders who were e-mailed the first notification were sent a second notification via postal mail. Thus, three groups resulted from the two wave notification process: (1) e-mail/e-mail (2) e-mail/paper (3) paper/paper. The response rate effectiveness of these three methods was tracked over time. Differences in response rate and question response by customer type will be examined for each of the methods. Implications for using Internet surveys with paper notifications will be discussed.

***Transforming a Paper Survey into a Web-based Survey: Respondent Experiences*, Annette Luyegu, Mathematica Policy Research; luyegu@yahoo.com**
As researchers devise more efficient and economic ways of collecting data, and as Web-based surveys become more refined and popular, survey methodologists need to advance the current data collection techniques to insure validity and reliability of data collected through the Internet. In order to make self-administered Web-based surveys more user-friendly, it is critical to recognize and understand the respondents' experiences as they complete the survey. For the computer-savvy respondent, a Web-based survey can be a pleasant experience; however, an individual with little or no knowledge in computer use could find the Web-based survey particularly frustrating. In 2002, the data collection exercise for the federally funded 21st Century Community Learning Centers Program annual performance report transformed an existing paper survey to a Web-based interface, and more than 5,000 after-school center directors were required to submit their data online. By not relying on "volunteer-only" respondents, this survey was unique from most data collection practices. This paper discusses experiences reported by the respondents as well as observations made by the author on the patterns of response, for example, the proportion of errors by question type, and the overall response rates of the paper survey and the Web-based survey. It will be interesting to see if survey methodologists are addressing the real challenges faced by respondents as they complete Web-based surveys.

INCENTIVES II

***A Test of Monetary Incentives for a Large-Scale Establishment Survey*, Paul Biemer, RTI International; ppb@rti.org, Christopher Ellis, RTI International; ellis@rti.org, Kimberly Robbins, RTI International; krobbins@rti.org, and Angela Pitts, RTI International; apitts@rti.org**
Incentives have been shown to increase response rates in household surveys but few studies have tested their effectiveness in establishment surveys. Since establishment survey respondents are employees of the sample organizations, participation in the survey may not be the choice of the Point of Contact (POC) in the establishment. Further, the decision to participate may be based more on economic than on altruistic or social exchange considerations. This paper presents the results of an incentive experiment embedded in the Occupational Information Network (O*NET) data collection program - a large-scale survey of employees in all occupations and industries. The split-ballot experimental design was conducted over a two-year period and involved about 30,000 employees in more than 10,000 establishments in 115 occupations. In the incentive group, establishment POCs were offered a prepaid $20 incentive, in addition to other non-cash incentives for their participation in the survey. The control group POCs received only the non-cash incentives which are part of the standard O*NET data collection protocol. Analysis was conducted to examine a) the effect of the incentive on response rates and b) the relative costs of the incentive versus the control protocols, since monetary incentives have the potential to at least partially offset their inherent costs through greater efficiencies in the data collection process and higher response rates. Results from this study will help survey research practitioners make informed decisions regarding the cost and data quality merits of this type of incentive in business surveys.

***Exploring a New Establishment Survey Incentive to Improve Response Rates*, Amy Luo, Ernst &Young; amy.luo@ey.com, and Glenn White, Ernst &Young; glennwhite@ey.com**
Exploring a new establishment survey incentive to improve response rates Obtaining higher response rates to surveys is becoming increasingly more difficult. This is happening in both household and establishment surveys. There has been much research focused on exploring the effectiveness of different forms of incentives in household survey, such as cash, gift cards, certificates etc. and the 'optimum' amount of the incentive. In practice, we have found that these incentives are not as effective for establishment surveys as they are for household surveys. We have experimented with a new incentive to boost the response rate of establishment surveys - providing an individual benchmark survey report to each participating establishment. The benchmark report shows how an individual company's response compared to the overall responses. To evaluate the effectiveness of this method we compare the response rates from establishment surveys with and without benchmark incentive reports. Second we compare the responses from those who indicated they would like to receive the benchmark survey report to those who did not indicate a desire for a report but completed the survey.

***An Incentive Experiment For a Survey of School Districts and Schools*, David Cantor, Westat; davidcantor@westat.com, Thomas A. Fiore, Westat; tomfiore@westat.com, Beth Sinclair, Westat; sinclad1@westat.com, James Demery, Westat; jamesdemery@westat.com, and Ellen Schiller, ABT Associates; Ellen_Schiller@abtassoc.com**

Decisions to participate on an organizational survey involve both group and respondent dynamics. Very little is known about the effectiveness of incentives to influence these dynamics. This study experimented with incentives on a survey of school districts and schools. As part of a survey for the State and Local Implementation and Impact of the Individuals with Disabilities Education Act (SLIIDEA), conducted for the Department of Education, an experiment was conducted in 2002 to explore these issues. A nationally representative sample of school districts and schools were surveyed to evaluate the implementation and outcomes associated with the 1997 Individuals with Disabilities Education Act. The experiment was conducted at two levels. At the first level, districts were assigned to one of three conditions designed to compensate the organization for participating: 1) no incentive, 2) $50 paid to the district and 3) reimbursement on request for costs to fill out the questionnaire (with a $500 hidden ceiling). Condition 2 was seen as directly providing a monetary benefit to the district, while condition 3 was seen as a gesture of good faith by the survey sponsors to relieve the burden of participation. The second level of incentives was for individual schools and also had three conditions: 1) no incentive, 2) $50 paid to the school and 3) $10 cash sent with the initial questionnaire mailing. Condition 2 was meant to compensate the school. Condition 3 was intended to be ambiguous as to whether the target was the individual completing the survey or the school. This paper will present the results of this experiment with respect to the response rates and the costs related to the use of an incentive. Costs will consider not only the incentive, but also any savings realized by reducing the amount of effort needed to obtain a completed questionnaire.

***Offering Cash Incentives to Students in a School Setting: Effects of an Incentive on Student Participation Rates*, Debbie Herget, RTI International; herget@rti.org, Tiffany Lytle, RTI International; lytle@rti.org, Ellen Stutts, RTI International; ess@rti.org, Peter Siegel, RTI International; siegel@rti.org, Kimberly Ault, RTI International; ault@rti.org, Mani Medarametla, RTI International; medaram@rti.org, James Rogers, RTI International; jerogers@rti.org, and Daniel Pratt, RTI International; djp@rti.org**

Obtaining high rates of student participation in school-based studies has become increasingly difficult in recent years due to such trends as increasing numbers of schools requiring active consent, student and school overburden, and the rise in number of high-stakes assessments. Additionally, high student participation rates are more difficult to achieve for studies that select student participants at random from a roster as opposed to intact classrooms. The success of these studies typically requires the assistance of school personnel to advertise the study and handle logistical arrangements. The Education Longitudinal Study of 2002 (ELS:2002) is a school-based research study that uses a random student selection process. ELS:2002 faced challenges in securing student participation in its base year data collection. As part of the design of the field test of the first follow-up study, RTI implemented an incentive experiment to increase student response rates. Offering an incentive was believed to not only make students more likely to participate but also facilitate school cooperation. Student respondents in half of the field test schools were offered a cash incentive and student respondents in the other half of the schools were offered a gift card in the same amount. It was anticipated that cash would result in a higher response rate than gift cards; this was the result in the field test. Based on the results of the field test, cash incentives were offered for the main study. Gift cards were offered only for schools that did not allow cash incentives for students. This presentation will report on the utility of using cash and gift card incentives to increase student response rates and the implications for other school-based studies. Analysis will include measures of data quality (in terms of item nonresponse) as well as quantitative data about response rates.

PREDICTING ELECTIONS

***Polls Apart?: Factors affecting Vote Choice for President*, Randall Thomas, Harris Interactive; rthomas@harrisinteractive.com, Mitchell Sanders, Harris Interactive; msanders@harrisinteractive.com, Renee Smith, Harris Interactive; rsmith@harrisinteractive.com, Robin Repass, Harris Interactive; rrepass@harrisinteractive.com, and Alyssa Johnson, Harris Interactive; amjohnson@harrisinteractive.com**

Many have suggested that America experienced a significant division of opinion in the 2004 U.S. election. We asked people about their attitudes toward a wide range of political issues (abortion, military spending, illegal immigration, etc.) and examined differences in attitudes between Democrats and Republicans in 5 primaries of 2000 prior to the election. For the 2000 general election, we examined differences in attitudes between voters for Gore and Bush in a survey immediately following the election. In 2004, we also administered the same battery of attitude items prior to the general election in two waves. Across both election cycles, voters for Bush were significantly different than voters for Gore or Kerry in 18 of 20 issues. In addition, as in the 2000 election cycle, we found significant differences in perceptions of the candidates and liking for them. As an example, Kerry voters saw Kerry as much more moral, intelligent, and a better leader than Bush while Bush voters saw Kerry as much less moral, less intelligent, and less of a better leader than Bush. The differences we obtained between Bush and Gore voters were somewhat smaller in 2000 than the differences we obtained between Bush and Kerry in 2004, supporting some observations that this most recent election cycle may have been characterized by greater polarization in attitudes.

***A Probability Theory-Based Test of the Reliability of 2004 Election Polls*, Joel Bloom, University of Oregon; jbloom@uoregon.edu, and Jennie Pearson, University of Oregon; jpearso2@gladstone.uoregon.edu**

After the elections of 2004, the reliability of election polls is again coming under fire. The appropriate question is not, however, whether some polls differed from election results, but whether the gap was larger than we would expect given the laws of probability. A number of factors suggests that election polls should be less reliable than issue polls, primarily that election polls attempt to sample from an unknown population - those who will vote; house differences in "likely voter" screens add to these discrepancies. In an earlier work ("Reliable Compared to What? Empirical Tests of the Accuracy of Election Polls, 2002," AAPOR 2003), I analyzed 232 polls of Senate races in swing states, and applied a reliability test based on simple probability theory. I found that while the average reported margin of error for these polls was 4.4% they had an actual average margin of error between 6.1% and 7.4%. In this paper I will update that analysis for a larger sample of presidential and Senate polls, and add a correction for real movement in voter preferences. I anticipate that findings here will be similar to those of my 2002 analysis and that the meta-sample of polls will show an apparent actual margin of error higher than the reported average. As a sidebar, in the previous work, I included an analysis comparing surveys conducted by SurveyUSA, finding no substantial differences. In the current analysis, I will also add a large number of Rasmussen surveys. I anticipate that my results will again show little difference between the newer (and somewhat controversial) method of computerized interviews used by these firms and more traditional live telephone surveys. I will also include a similar analysis of Zogby's web-panel polls, which I anticipate will be less reliable than traditional telephone surveys.

***Survey of Political Surveys [SOPS]*, Nat Ehrlich, Office for Survey Research, Mich. State University; nathaniel.ehrlich@ssc.msu.edu, and Larry Hembroff, Michigan State Univ.; larry.hembroff@ssc.msu.edu**

Prior to the political conventions, we contacted the directors of surveys at ABC and CBS/Times asking for their cooperation in compiling findings of national presidential election polls, including all questions asked, voter categories, and a complete explication of their methodology; they agreed to provide the data. Our analysis will focus on two points: the overall predictive accuracy of the different methodologies and datasets, and the relationship to accuracy of the demographic and attitudinal data [e.g. race, sex, and age; favorable/unfavorable impressions of the candidates, approval of the incumbent, importance of various issues]. We have begun the analysis and can report that one of the contributing agencies conducted 27 tracking polls between the close of the Republican convention and the election; the average margin of victory for President Bush, over all the polls and based on registered voters, precisely matched the actual national margin of 3.1%. In addition, a multivariate analysis indicates that the popular vote is a joint factor of evaluation of the performance of the incumbent and the difference in the proportion of respondents who say they have favorable impression of the major candidates.

***Outcomes and Lessons Learned from Polling Voters about the Voting Experience*, Lonna Atkeson, VoteWatch; atkeson@unm.edu, Rich Clark, VoteWatch; clark@eviog.uga.edu, Julio Cesar Campero, Aguirre International; jcampero@aiweb.com, Joan Crowley, VoteWatch; jcrowley@nmsu.edu, Joan Decker, VoteWatch; joand@coldwellbanker.com, Susan Gabbard, Aguirre International; sgabbard@aiweb.com, Marie Hansberry, VoteWatch; mthansberry, Trish Hernandez, Aguirre International; thernandez@aiweb.com, Steven Hertzberg, Votewatch; steven@votewatch.us, Kristin Juffer, VoteWatch; drkjuffer@earthlink.net, Kerry Krisman, VoteWatch; kkrisman@cas.ucf.edu, Carlos Leon, VoteWatch; joand@coldwellbanker.com, Kate Mulqueen, VoteWatch; kate@nm.net, Robert Santos, NuStats; rsantos@nustats.com, Fritz Scheuren, NORC; scheuren@aol.com, Diane Stoner, VoteWatch; dstoner55@comcast.net, Carmen Sum, Aguirre International; csum@aiweb.com, Nicole Vicinanza, Aguirre International; nvicinanza@aiweb.com, Cheryl Wiese, VoteWatch; wiese.c@ghc.org, and Jen Ziemke, VoteWatch; jjziemke@students.wisc.edu**

During the 2004 Presidential election, Votewatch a non-profit, non-partisan election monitoring organization conducted an exit poll in New Mexico, one of the battle-ground states. The goal was not to predict the election outcome but to obtain information on the experience of voters and more generally on voting process. The survey's 962 respondents were randomly selected from Election Day voters in the Albuquerque area via two-stage probability sampling. Respondents were asked about their voting history and polling place, interactions with poll workers and voting technology, confidence that their vote counted, under voting, the presidential vote and basic demographic information. The method was unique in that it used volunteer, trained survey professionals, including AAPOR members, supervising citizen volunteers to do voter intercepts. The results of the poll are presented, and we discuss how this gives us a better understanding of the voting process in Albuquerque. The paper concludes with implications to US voting systems and how they might be scientifically monitored in future elections.

***Vote Over-Reporting: Testing the Social Desirability Hypothesis in Telephone and Internet Surveys*, Allyson Holbrook, University of Illinois at Chicago; allyson@uic.edu, and Jon Krosnick, Stanford University; krosnick@stanford.edu**

Scholars studying political behavior have long been troubled by the fact that survey respondents typically report having voted at a rate higher than citizens in fact turned out on election day. Many observers of this phenomenon have presumed that it reflects intentional misrepresentation by respondents who did not vote and would be embarrassed to admit it. Previous attempts to reduce social desirability bias have not successfully reduced over-reporting of turnout, though. However, no one has tested on a large scale whether reporting turnout anonymously would reduce over-reporting. This is a strong test of whether social desirability plays a role in vote over-reporting, as anonymity has been shown to reduce social desirability bias in other contexts. People are less likely to under-report report socially undesirable attitudes and behaviors (e.g., anger toward affirmative action policies and falsifying tax returns) and less likely to over-report socially desirable attitudes and behaviors (e.g., recycling) when they know their responses are anonymous and cannot be directly linked to them. In order to test this hypothesis more directly, we implemented two techniques that allowed respondents to report anonymously whether or not they voted: the "list" technique and randomized response. We report the results of four studies involving 9 separate national samples of adults (one telephone and eight internet). Our results suggest that the list technique can be successfully implemented in both telephone and internet surveys. However, using the list technique was successful at reducing turnout reports only among respondents interviewed via the telephone, suggesting that social desirability concerns lead people to intentionally distort their direct self-reports in telephone surveys, but not in internet surveys. However, respondents were apparently unable or unwilling to implement the randomized response technique properly in either internet or telephone surveys, suggesting a limit to its utility in surveys conducted in these two modes.

WAR AND FOREIGN POLICY

***Combat and Confidence: The Foreign Policy Attitudes of U.S. Soldiers Engaged in the Global War on Terror*, Jason Dempsey, United States Military Academy & Columbia University; jason.dempsey@usma.edu**

This research project will analyze the results of a large-scale random-sample mail survey of active-duty Army personnel conducted April-June 2004. This survey is the first of its kind to ask soldiers how they view the U.S. role in the world and to ask what they believe should be the primary purpose of the armed forces of the United States. The survey is constructed to allow for comparison of attitudes between officers and enlisted personnel, as well as controlling for deployment history, military occupational specialty, citizenship status, and race/ethnicity. This research is significant because it will be the first large-scale study ever to analyze the worldview and motivation of soldiers while controlling for deployment history.

***American Leaders and the Public: Areas of Convergence and Discord in How They View Foreign Affairs*, Alvin Richman, U.S. Department of State, Office of Research; RichmanA@State.gov**

This paper compares the views of the U.S. public with those of U.S. leaders of various government and private organizations involved in foreign affairs, using the Chicago Council of Foreign Relations surveys taken last summer. Analyses of the Chicago Council surveys taken between 1994-2002 have shown that American leaders and public share similar ways of organizing their views on U.S. foreign policy, even though their preferences on certain foreign policies may differ sharply. Both leaders and public organize their foreign policy opinions in

four distinct groups, which are examined separately in this paper. A large public/leader split on GLOBAL ALTRUISTIC policies has emerged during the past several years, with American leaders now clearly more supportive of initiatives on political issues (e.g., promoting democracy) and much more supportive of providing economic assistance. In contrast, the public continues to give much higher priority than leaders to DOMESTIC ISSUES bearing on foreign policy. The public is more concerned than leaders about the adverse personal financial impact of U.S. policies, and therefore gives higher priority to restricting the flow of people and products into the U.S. Both leaders and public give high priority to the GLOBAL INTEREST objective of preventing nuclear arms proliferation and lower priority to protecting the global environment and strengthening the U.N Both U.S. leaders and public also give high priority to combating international terrorism, but differ on a number of other MILITARY SECURITY issues. Leaders are generally more willing than the public to defend U.S. allies and participate in international peacekeeping operations. However, the public is more willing to use American military forces to protect U.S. oil supplies and topple unfriendly regimes that support terrorist groups threatening the U.S. Both groups are reluctant to use military force without U.N. authorization to prevent a country from acquiring nuclear weapons.

***The Hall of Mirrors: How the US Public and Elite Misperceive Each Other on Foreign Policy*, Steven Kull, Program on International Policy Attitudes; skull@pipa.org**
For the democratic process to work effectively the public must have an accurate understanding of the positions and votes of their elected representatives and elected representatives must have an accurate understanding of their constituents' preferences. However this study reveals that in regard to US foreign policy the US public and elite have major misperceptions of each other. In particular the public overestimate leaders support for multilateral efforts while the leaders underestimated the public's support for them. Nationwide polls found that the public showed little awareness of how their own Congressional representative had voted on key foreign policy questions. Constituents tended to assume that their member voted the way they wanted them to vote: more often than not, incorrectly. Perceptions of the president were only slightly more accurate. Likewise in a survey of 450 US leaders, including members of the administration and 100 Congressional staffers, when asked about their perceptions of majority public preferences only a minority were correct. Congressional staffers were poor at estimating attitudes in their own districts. Public surveys were conducted by the Program on International Policy Attitudes and the elite survey was conducted by the Chicago Council on Foreign Relations.

***Modeling the Support for the U.S.-Led war on Terrorism Among Egyptian Citizens: Implications for International Public Diplomacy*, Michael Elasmar, Boston University; elasmar@bu.edu**
This paper utilizes the results of a public opinion survey of a probability sample of Egyptian citizens. The focus of this paper is on the variation in the level of support for the U.S.-led war on terrorism among the survey respondents (n=933). This paper uses structural equation modeling techniques for testing a theoretical model of international attitude formation. The dependent variable in this model is the level of support for the U.S.-led war on terrorism. This paper finds that both direct and indirect antecedents explain the variation in the dependent variable. These antecedents include respondents' cultural capital, level of education, interpersonal contact with America, mediated contact with America, and beliefs about the intentions of the U.S. government. Implications are drawn about the type of Egyptian respondent most likely and least likely to support the U.S.-led war on terrorism. Implications for international public diplomacy are also discussed.

NONRESPONSE

***The Effect of Survey Follow-up on Nonresponse Bias: Joint Canada/United States Survey of Health, 2002-03*, Stephen Blumberg, National Center for Health Statistics, CDC; sblumberg@cdc.gov, Karen Davis, National Center for Health Statistics, CDC; KEDavis@cdc.gov, Meena Khare, NCHS/CDC; mxk1@cdc.gov, and Michael Martinez, National Center for Health Statistics; MEMartinez@cdc.gov**
Response rates for random-digit-dial (RDD) telephone surveys continue to decline. Yet several studies suggest that nonresponse bias remains constant for response rates in the range of 40% - 70% (e.g., Curtin et al., 2000; Keeter et al., 2000). The 2002-2003 Joint Canada/United States Survey of Health (JCUSH) provides another example of how nonresponse bias remained constant despite efforts to increase response rates. JCUSH was conducted jointly by Statistics Canada and by the National Center for Health Statistics of the U.S. Centers for Disease Control and Prevention. This RDD survey was designed to collect data regarding health status and access to health care for both Canadian and U.S. residents, using comparable procedures. Toward this end, interviews for both populations were conducted from Statistics Canada's regional offices. After 5 months of data collection (November 2002 - March 2003), the Canadian response rate was 66%, relative to a U.S. response rate of 45.3%. To achieve greater response rates for the U.S. sample, additional telephone calls and refusal conversion attempts were conducted in April and June 2003. The final U.S. response rate (AAPOR RR4) was 50.2%. Sampling weights for U.S. data collected prior to March 31st were produced, and resulting estimates for key health indicators were compared with the final weighted survey estimates and with estimates from the 2002 U.S. National Health Interview Survey. (The NHIS estimates were considered the "gold standard" for analyses of bias.) Similar comparisons were conducted for interviews completed with no more than 10 dials (RR4 = 37.0%), interviews completed with no more than 15 dials (RR4 = 43.3%), and interviews completed without refusal conversion efforts (RR4 = 41.8%). Nonresponse bias was constant for nearly all key health status and health care access measures.

***Analyzing Marginal Response Rates in the CAHPS Medicare Fee-for-Service Survey*, Larry Campbell, RTI International; campbell@rti.org, G. Gordon Brown, RTI International; ggb@rti.org, Lisa Carpenter, RTI International; lac@rti.org, and Linda Dimitropoulos, RTI International; lld@rti.org**
This paper is an effort to extend recent work that evaluated the effect of increasing the level-of-effort by raising the maximum number of call attempts in a telephone follow-up of the CAHPS Medicare Fee-for-Service Survey. In that work, we used survival analysis to analyze marginal response rates among non-respondents to the mail survey who participated in the telephone follow-up to determine the appropriate number of attempts to make to each sample member. The goal was to optimize the response rate and marginal cost of a completed survey. Our key findings were: 1. With the increased level-of-effort, we did not penetrate an under-represented segment of the population with greater number of call attempts; 2. With the maximum number of calls set at twelve, we observed neither significantly diminishing returns with greater calls, or increased marginal cost per completed survey; and 3. The CAHPS measures themselves were unaffected by increased call attempts. This paper extends this work by further developing the statistical methods used to analyze the response data, including statistical tests for checking the significance of our previous and new findings. Discussion will focus on further

analysis of the marginal effects on response rate and cost for increased call attempts, further development of the survival modeling, and of the balancing of the effects of bias and precision in the estimates.

***Determinants and Outcomes of Initial Contact in the National Health Interview Survey, 2004*, James Dahlhamer, National Center for Health Statistics; fzd2@cdc.gov, Catherine Simile, Nat'l Center for Health Statistics; CSimile@cdc.gov, Barbara Stussman, National Center for Health Statistics; bjs6@cdc.gov, and Beth Taylor, National Center for Health Statistics; bft8@cdc.gov**

Previous research has demonstrated the utility of collecting detailed contact or call record data, including observations of neighborhood characteristics, features of the interactions between interviewers and householders, and information on the day and time of contact attempts, for exploring nonresponse in household surveys. Indeed, at a 2002 response rate summit sponsored by the U.S. Census Bureau, call records or contact histories were unanimously recommended by an expert panel as a source of data for tracking and better understanding survey nonresponse. With this in mind, the National Health Interview Survey (NHIS), an ongoing population based health survey conducted by the National Center for Health Statistics, has adopted a stand-alone, Blaise-based instrument (Contact History Instrument, or CHI) designed to capture critical information on each contact attempt. Utilizing nearly 200,000 NHIS contact records collected via CHI in 2004, this paper focuses primarily on the issue of making initial contact with sample households. More specifically, it assesses the probability of making contact for specific times of day or days of the week, and whether the time of day or day of the week influences the outcome of a contact (e.g., appointment set, refusal, break-off, complete interview, etc.). It explores whether interviewers are making visits to households at optimal times, and the systematic differences in interviewer calling patterns by social environmental characteristics (e.g., region of country, U.S. Census regional office, urban/rural residency, MSA status, and others). Answers to these questions will aid future work on modeling contact and cooperation in the NHIS, and inform efforts to improve the efficiency and productivity of field operations.

***NHANES Converted Refusals: Are They Different from Willing Respondents in Cardiovascular Risk Factors?*, Yinong Chong, Centers for Disease Control and Prevention; ychong@cdc.gov, Margaret Carroll, CDC; mcarroll@cdc.gov, Vicki Burt, Centers for Disease Control and Prevention; vburt@cdc.gov, and Pat Montalvan, Westat; Pat Montalvan MONTALP1@WESTAT.com**

Over the years the National Health and Nutrition Examination Survey (NHANES) has maintained response rates over 80% for household interview and over 75% for medical examination. The extensive efforts to maximize response rates are driven by the assumption that converting hard-to-get respondents not only improves precision, but also reduce nonresponse bias. However this assumption has seldom been tested in health surveys. This study compares cardiovascular risks between converted refusals and willing respondents to investigate if any differences exist. The study population included adults 20 years and older who had completed the interview and medical examination (n=9471). Administrative data from 1999-2002 were used to identify respondents who initially refused the household screening or health interview (n=2036), or refused the medical exam (n=485). Age-standardized prevalence rates of current smoking, hypertension, obesity and total cholesterol level for these two groups were compared with those of willing respondents (n=6950). Compared with willing respondents, those who initially refused medical examinations were more likely to be current smokers (35 versus 24%, $p<0.001$). No significant differences were detected in hypertension, obesity or total cholesterol level between the above two groups. The group who initially refused screening/interview did not show any significant differences from willing respondents in any of these risk factors. These findings indicate that converted refusals for the medical examination showed significant difference in one cardiovascular risk but not in others. Future investigations should focus on finding patterns and the cause(s) for the differences, in order to discern whether these differences are random occurrence or systematic, and if the differences are confounded by other factors.

POSTER SESSION

***When a Stranger Calls: The Impact of Caller ID on Telephone Surveys*, Kathy Krey, Baylor University; kathy_krey@baylor.edu, and Jodien Matos, Baylor University; Jodien_Matos@baylor.edu**

Nonresponse is a critical issue in public opinion research and can weaken the representativeness of samples and diminish the reliability of survey data. Call screening technology, in particular, presents a potential dilemma as it enables subscribers to screen calls and therefore contributes to survey nonresponse. Subscriptions to caller identification services have increased noticeably sine the 1990s with a greater number of households using the technology to manage incoming telephone calls. In this study we examine caller ID subscription patterns, respondents' screening habits, and responses to caller ID listings. In a random digit dialing telephone survey conducted with a CATI system, we asked 1286 McLennan County, Texas residents if they subscribed to caller ID, used it to screen calls, and if the caller ID listing made them more or less willing to answer the phone. Because we regularly conduct telephone surveys in this area, it is important to understand how caller ID screening affects our sample. We found that 71 percent of our sample subscribed to a caller ID service, with 66 percent of those subscribers reporting that they use it all of the time to screen calls. In addition, we discovered that respondents' reaction to the unique caller ID listing varied demographically which suggests that a subset of the population may not be represented due to call screening. Our study raises important issues regarding caller ID subscription and its impact on survey response rates.

***Who's Calling?: The Impact of Caller-ID Displays on Telephone Survey Response*, Allan McCutcheon, UNL-Gallup Research Center; amccutcheon1@unl.edu, and Jack Ludwig, Gallup; jack_ludwig@gallup.com**

Caller-ID technology raised a national level debate when in January 29, 2004, the Federal Communication Association mandated telemarketers to transmit caller-ID information when making telemarketing calls (Federal Communication Commission-FCC, 2004). This initiative followed the establishment of the Do Not Call national registry. Survey research organizations, however, are exempted from these rules (Council for Marketing & Opinion Research -CMOR, 2004). Stimulated by the debate on the caller-ID, in the spring of 2003 The Gallup Organization conducted a caller-ID randomized, pre- and post-experimental design to test the effectiveness of different caller-ID displays (names) and their impact on response, contact, and cooperation rates for telephone surveys. This research focuses on the impact of Caller ID listing on the frequency of final dialing dispositions, and on associated rates that are often used to assess the quality of data collection. The data include sampling designs employing both RDD and (client-supplied) list samples. The analysis examines the AAPOR (2004) standard response reports (run separately for "pre" and "post" periods) as the basis for investigating the impact of the

implementation of Caller ID listing on relevant dispositions and rates. We find evidence for the hypothesis that the caller-ID transmission works as a sort of "compact invitation letter," similar to that found for advance letters which underscore the legitimacy of a survey, take away suspicion, and communicate the value of the survey thereby positively influencing response rates (Dillman, 1978; Goyder, 1987; Groves & Couper, 1998).

***Strengthening Higher Education Through Gridiron Success? Public Perceptions of the Impact of National Football Championships on Academic Quality*, Robert Goidel, Louisiana State University; kgoidel@lsu.edu, and John Maxwell Hamilton, Louisiana State University; jhamilt@lsu.edu**

In this paper, we consider whether the public connects athletic success to academic quality, economic health, and quality of life, and whether the connections between athletic success and academic quality are stronger during a year in which a state university wins a national championship. Data for the study are based on two statewide, randomly selected samples of Louisiana voting age residents with the first survey conducted in February 2004 in the wake of LSU's BCS national championship and the second survey conducted in December 2004 in the wake of a 9-2 football season. Using the data, we examine: (1) What type of respondents are most likely to make connections between athletic success and academic quality; (2) Whether these connections make a difference in terms of support for state funding for higher education; and (3) Whether respondents are less likely to make these connections during a relatively less successful football season. Embedded within the February 2004 survey instrument, we also consider whether respondents primed with specific question wording identifying the national championship are more likely to make these connections than respondents given more generic language referring success in athletics.

***Improving In-Person Data Collection in a Random Household Survey in Low- and Moderate Income Census Tracts*, Ronald Bass, ORC Macro; Ronald.H.Bass@ORCMacro.com, and Jaceey Sebastian, ORC Macro; jaceey.sebastian@orcmacro.com**

ORC Macro recently conducted a study of the financial attitudes and practices of residents in low- and moderate-income census tracts in Chicago, Los Angeles and Washington, DC. Data collection in one-third of the selected census tracts consisted of random in-person surveys with the household financial decision-maker. A 1997 study, covering similar subject matter, presented a range of problems, including difficulty in achieving high response rates within a reasonable time frame and the need to continually recruit new interviewers as experienced interviewers abandoned the project. Management strategies employed to maximize the response rate and ensure accurate data include: 1) Partnering with local community groups to generate awareness of the study and recruit interviewers from within the community; 2) Recruiting highly experienced survey research supervisors to train and work closely with new interviewers; 3) Pre-survey canvassing of census tracts by the field manager and city field supervisor to verify their appropriateness for inclusion in the sample; 4) On-going continuing training provided by field manager and city field supervisor; 5) Listing/screening of eligible housing units performed by city supervisors and experienced interviewers, with call-back appointments handed off to less-experienced interviewers; 6) Pre-contacting apartment complex offices to arrange for interviewer visits; 7) Distributing respondent incentives (U.S. Postal money orders) in the field at the end of interview; 8) Providing field supervisors and field interviewers with the availability of contact with the project director and survey manager whenever it was needed. The overall response was 78%. The response rate was significantly higher in Los Angeles (92%) and Chicago (90%) than in Washington, DC (62%). We will analyze the two higher response rate strata against the lower one (DC), looking at two measures of data quality: item non-response and comparison of demographics with known population parameters.

***Seeking Quality With Secondary Survey Data*, Thomas R. Marshall, University of Texas at Arlington; tmarshall@uta.edu**

Most articles on improving surveys focus on primary survey research -- that is, polls and surveys a researcher might carry out in the future. However a considerable body of research is based on secondary survey research -- polls done at an earlier time by other researchers. When using secondary survey research, a pollster has no control over how the research was conducted, and often cannot even be sure how key decisions were made. Very little research is available on the quality standards that should be used in secondary survey research. This paper reviews recent work on secondary survey research for the quality standards that are now being used, identifies common shortcomings, and makes recommendations for quality standards that should be used.

***Identical Study Protocol, Substantially Different Outcome Rates: Determining the Reasons*, Karen Foote Retzer, Survey Research Laboratory, University of Illinois; karenr@srl.uic.edu**

Between June 2002 and May 2005, the Survey Research Laboratory at the University of Illinois will have conducted 12 cross-sectional waves for a single RDD study (to date we have completed 10 waves). For each wave (with the exception of wave 1), we followed the same protocol-we purchased sample for the same geographic area, used the same questionnaire, and utilized the same telephone/interviewing procedures. Despite the uniform field and sampling procedures, our outcome rates have differed greatly. For example, response rates (AAPOR response rate 3) have ranged between 37.4% and 52.8% for different waves; cooperation rates have ranged from 56.1% to 73.5%; and refusal rates have ranged from 17.1% to 32.4%. We also have noted that the screening completed by the sampling vendor for each wave has resulted in screening out as little as 19.4% to 30.0% of the sample ordered. Finally, the proportion of cases that are listed phone numbers fluctuates by wave. This poster will review the varying outcome rates and will try to discern the reasons behind the differences, taking into consideration issues including--but not limited to--seasonality, differences in screening by the vendor, and proportion of sample that is listed.

***"The Art of Associating": The Central Role of Peers in Civic Life*, Casey Klofstad, Harvard University; klofstad@fas.harvard.edu**

This paper assesses the role that interpersonal social relationships play in democratic systems. Specifically, this project uses a unique new data source that I collected to more comprehensively examine the impact that the people with whom we hold close interpersonal relationships with (our "peers" or our "social network") have on how we participate in civil society. Recent research on civic participation shows that peers could have an impact on how individuals participate in the political process. For example, existing works show that individuals who talk about politics with their friends and family are more likely to be politically active. However, this line of inquiry is incomplete because existing data sources do not allow us to establish a definitive causal relationship between social interactions with peers and individual behavior, nor do they allow us to establish what causal mechanisms drive this relationship. This research makes use of new survey data that I collected in a way that allows us to directly address these omissions in the literature. These data show that peers can cause individuals to become more civically active. This is case because interactions with peers can lead individuals to become more engaged with civil society, become more informed about the democratic political process, and to be subject to recruitment to participate in

civic activities. Additionally, my analysis shows that peers can be as influential, if not more so, than traditionally studied agents of politicization such as political parties and voluntary associations. This project therefore highlights the fact that in order to fully understand the nature of contemporary participatory democracy, we must continue to increase our understanding of peer influence.

***From inches to centimeters: The uncritical use of available measurements*, Pablo Paras, University of Connecticut; pp@dataopm.net, and Luis Estrada, UCSD; lmestrad@weber.ucsd.edu**

Can we safely assume that if a measurement is appropriate for established democracies it will also be reliable or valid in emerging democracies? Can we and should we capture potentially complex constructs using single-item indicators? The present paper assesses the reliability and validity of at least two important indicators that have been widely used in survey research. The first, interpersonal trust, is a key indicator of social capital theory, which is often used as a correlate of modernization and quality of democracy. The second, the left-right ideology scale, has been used in many electoral and political studies. Both are a single-item measures of potentially complex phenomenon, and most of the time, both have been used uncritically, that is, without assessing if the indicator is an appropriate measure - for the specific social system under study. By using secondary data and original survey research, we test if these measurements are adequate for the case of Mexico. Our main data sources are: (1) the 1990, 1995 and 2000 WVS; (2) the first national panel study conducted in Mexico between 2000 and 2002; and (2) two original surveys conducted in Mexico City in 2004, that provide alternative measures of the constructs under study. We use longitudinal and conceptual consistency analyses to test the reliability of indicators. We use confirmatory factor analysis to test the internal and external validity of multi-item constructs. Finally we replicate scholarly findings by testing if the proposed multi-item indicators are better predictors than their single-item predecessors. By replicating previous analyses on the subject, and - by addressing several gaps in the - literature, we find that the assessment of such concepts is not entirely adequate. Our main contribution is to call the attention of the survey industry about the relevance of using constructs without previously making a theoretical justification.

***Modeling Question-Answer Processes of Structured Interviews using Interactive Hidden Markov Models*, Adriaan Hoogendoorn, Vrije Universiteit; aw.hoogendoorn@fsw.vu.nl, cees elzinga, Vrije Universiteit; ch.elzinga@fsw.vu.nl, and Wil Dijkstra, Vrije Universiteit; w.dijkstra@fsw.vu.nl**

In this paper we will try to get a better understanding of the question-answer process in structured interviews by studying the sequences of coded utterances of the interviewer and the respondent, and by trying to fit a model to these data. Ideally, structured interviews give rise to very short and simple sequences: correctly worded question, admissible answer, acknowledgement of perceipt of that answer (Maynard & Schaeffer, 2002). However, in practice, many of such sequences deviate from this simple paradigm. Sequences may deviate for a wide variety of reasons: both respondent and interviewer may behave inappropriately by misphrasing or misinterpreting questions, becoming engaged in social conversation, suggesting certain answers, etc. Such behavior contributes to measurement error. A better understanding of the mechanisms that generate non-paradigmatic sequences could lead to improved interviewing techniques, improved question wording and improved interviewer training. In the past decades, several attempts have been made to explain the frequent occurrence of non-paradigmatic sequences. These explanations (e.g. Suchman & Jordan, 1991 and Tourangeau et. al, 1988) have been formulated as quite general and qualitative theoretical frameworks, the tenability of which is hard to test empirically. In this paper we will make an attempt to formalize these cognitive frameworks in order to test their tenability and relate model-parameters to characteristics of questions, interviewers and respondents. The models discussed are so-called Interactive Hidden Markov Models. This class of models are used to mimic the interactive processes and allow for hidden Markov chains that are affected by observable utterances of both the respondent and the interviewer. An interpretation of the hidden states of the modelled process may shed new light on understanding the interaction between respondent and the interviewer.

***Adapting to a Changing Data Collection Environment in an Ongoing Survey*, Mark Pierzchala, Mathematica Policy Research; mpierzchala@mathematica-mpr.com, Tom Barton, Mathematica Policy Research; tbarton@mathematica-mpr.com, Paul Guerino, American Institutes for Research; pguerino@air.org, and Cheryl Hills, Mathematica Policy Research; chills@mathematica-mpr.com**

In recent years, survey research organizations have faced substantial challenges in the conduct of telephone surveys, both in contacting respondents and securing their cooperation. Challenges affecting contact are caused by the proliferation of caller ID, number blocking, and the increased use of cell phones. Challenges affecting cooperation include reluctance to answer incoming phone calls, effects of the national "do not call list," and security concerns in light of September 11. Conducting a survey in this more difficult climate results in higher costs to achieve an acceptable response rate. The 2003 National Survey of Recent College Graduates (NSRCG), sponsored by the National Science Foundation (NSF), was fielded as multimode survey including paper, CATI, and Web modes. In its first weeks, the survey achieved response rates below expectations. As data collection progressed, the challenges listed above needed to be confronted. In this paper, we describe our experience in re-engineering the survey while it was underway. Steps taken in the re-engineering process included offering an incentive, adding email reminders, improving survey introductions, adjusting call scheduling parameters, using individual phone lines instead of T-Lines (to get around call blocks), conducting additional mailings, and intensifying locating efforts. These actions were taken at a rapid pace and in a compressed timeframe. The re-engineering resulted in a marked improvement in the response rate. In this paper, we describe the implementation process and the effects of selected actions both collectively and individually. Some actions, such as the incentive offer and email reminder, were clearly more effective at promoting response than others. Our experience demonstrates how an intensive, collaborative effort combining multiple actions can improve response rates in surveys underway.

***When Census Data Don't Add Up in the Field: Two case studies*, Kathleen Tiefenwerth, Battelle Centers for Public Health Research and Evaluation; tiefenwerthk@battelle.org, and Karen Tucker, Battelle Centers for Public Health Research and Evaluation; tucker@battelle.org**

In two studies of disease prevalence among young children, census data were used to develop projections of how many eligible participants should be residing in the study areas. Each of the studies used the 2000 census data to predict the number of potentially occupied households that should have been distributed throughout the study areas. Enumeration of the households and subsequent screening of the homes for eligible children, has yielded a much lower than predicted number of both occupied households, and children in the right age range for the studies. The poster will present summary descriptive statistics comparing census predictions and actual findings after visiting

homes in these communities. Statistics will include the percentage of census estimates that were confirmed by field visits, for both occupancy of dwelling units and presence of eligible children in the home. A description of validation of our field screening efforts will also be included.

***The Survey Organization-Respondent Relationship: A Taxonomy of Conceptualizations*, James Chesire, National Opinion Research Center at the University of Chicago; chesire-james@norc.org, and Jake Bartolone, National Opinion Research Center (NORC); bartolone-jake@norc.org**

The cooperation of survey respondents affects data quality directly, through their willingness to participate candidly, and indirectly, through overall response rate and conversion and locating costs. To address this, survey research organizations teach interviewers skills for gaining cooperation and avoiding refusals. The prevailing industry model of the relationship between the survey organization and the respondent implicitly assumes that these interactions are transactional in nature; there is only one purpose: to obtain information from the respondent. However, anecdotal evidence from field staff on the National Longitudinal Survey of Youth (NLSY) at the National Opinion Research Center (NORC) suggests that a successful interviewer often maintains many different types of relationships with her respondents. It is conventionally accepted that the better an interviewer understands a respondent's personal motivations for participating, the more likely they are to gain that respondent's cooperation. In order to understand the reasons why respondents choose to participate in surveys, we first identified several possible varieties of respondent motivations and interviewer approaches, through literature review (primarily Don Dillman's Social Exchange principles) and informal interviews with experienced field staff at NORC, respectively. We reviewed case management data from previous rounds of the NLSY, focusing on refusals and converted refusals because of their richness of detail, and where possible we will classify the respondent's motivation and the interviewer's approach into the appropriate codeframes. Then, for each combination of motivation and approach, we will determine the average completion rate across the next several rounds of the survey, in order to identify which approaches work best with which motivations, and if there is one default approach that works the best in the absence of information about the respondent's motivation for participation.

***A Comparison of Within Household Random Selection Methods for Random Digit Dial Surveys*, Molly Longstreth, University of Arkansas; mlongstr@uark.edu, and Todd Shields, University of Arkansas; tshield@uark.edu**

Two methods, the Rizzo, Brick and Park (2004) and the last-birthday, were used to randomly select respondents from within households telephoned in two statewide RDD studies. The Rizzo et al. method employs both probability and quasi-probability selection techniques; the last-birthday method, quasi-probability selection. Although the last-birthday method is relatively unobtrusive and simple to administer, numerous studies have shown that the last-birthday method is about 75 to 91 percent effective in selecting respondents randomly from within households (Lavrakas et al., 1993; Lavrakas et al., 2000; O'Rourke and Blair, 1983). The Rizzo, Brick and Park (2004) method for selecting respondents within households is less obtrusive for the approximately 85 percent of households with fewer than three adult members, but uses the last birthday method to select among more than two adults and the Kish method if informants don't know birthdays. In this comparison, interviewers were randomly assigned to each selection method and were seated in different parts of the telephone laboratory. About 278 surveys were completed using each method. Results of the first survey show that although surveys conducted via the last-birthday method had shorter average duration than those of the Rizzo et al. method, the differences are not statistically significant. The Rizzo et al. method used fewer calls per complete. Using AAPOR formulae, the response rates of the last birthday (23.9%) were slightly higher than those for the Rizzo et al. (21.4%) method. The two methods resulted in approximately the same ratios of men and women, 37 to 63 percent, compared with expected ratios of 45 to 55 percent. The average ages, levels of education and income of respondents in each group are comparable. Number of adults in households is similar. Responses on the substantive questions in the survey did not vary by within-household selection method.

***Sixteen Tons (of Paperwork) and What Do You Get? Self-Reported Workweek Length of Doctoral Scientists and Engineers*, Tom Hoffer, NORC; Hoffer-Tom@norc.org, and Karen Grigorian, NORC; grigorian-karen@norc.uchicago.edu**

This paper has two goals: first, to present new findings on workweek differences among a particularly interesting segment of the U.S. labor force, doctoral scientists and engineers; and, second, to provide a critical review of the methodological literature on measuring the number of hours that individuals devote to work. The data analyzed are drawn from the 2003 Survey of Doctorate Recipients, a biennial panel survey of doctoral scientists and engineers in the U.S. sponsored by the National Science Foundation. Descriptive findings show substantial differences within the doctoral labor force along the lines of employment sector, time since doctorate and (for those employed in academe) tenure status, marital status, and number of children. The methodological review argues that while self-reports to standard survey questions about hours worked per week may well be problematic for some segments of the labor force, they are likely to provide valid and reliable data for professional workers with established work routines. This criterion is met by the vast majority of doctoral-level workers and, while we are not able to directly validate them, the self-reports are likely to prove valid and reliable.

***The Impact of Follow-up Contacts on Survey Data and Response Rates*, Elizabeth Westin, Westat; elizabethwestin@westat.com, Michele Harmon, Westat; micheleharmon@westat.com, and Kerry Levin, Westat; kerrylevin@westat.com**

Standard practice in self-administered surveys includes the use of follow-up contacts to the respondent, often incorporating additional data collection modes, in order to increase the survey's overall response rate. The method of utilizing multiple and carefully timed contacts is recommended for mail surveys (Dillman, 1991), and studies have shown this method to have a favorable impact upon the administration of electronic mail surveys as well (Schaefer and Dillman, 1998). However, the gains associated with obtaining more survey responses from additional follow-up measures are not always clear. Certainly, decreasing the number of nonrespondents in the study sample will reduce the likelihood of nonresponse bias. However, as noted in a study by Lynn, Clarke, Martin, and Sturgis (2001), extended efforts to contact respondents significantly reduced nonresponse bias for some demographic variables but had no impact upon attitudinal measures. A recent study sponsored by a Federal granting agency solicited respondents from five stakeholder groups to complete an Internet survey rating their satisfaction with current grant application procedures. Respondents were first sent a pre-notification letter in the mail describing the upcoming study, followed by an email containing the initial request to complete the Internet survey. Two follow-up email reminders were sent to respondents approximately ten days apart, and a final contact to nonrespondents was made by telephone with a request to complete the Internet survey. This study examines how each follow-up effort affected the survey results and corresponding response rates throughout the field period. It also discusses how the follow-up efforts impacted the response rates by increasing the number of completed surveys and by identifying ineligible respondents in the sample frame. Additionally, the impact of follow-up contacts will be examined across the individual stakeholder groups, with particular emphasis on differences between groups of agency employees and non-agency employees.

***You Don't Have to Accept Low Survey Response Rates - How we achieved record survey cooperation rates*, Jane Traub, Commercial; jtraub@scarborough.com, Kathy Pilhuj, Commercial; kpilhuj@acarborough.com, and Daniel Mallett, Commercial; dmallett@scarborough.com**

The increasing trend of respondent non-cooperation is an important concern for survey quality. Contrary to industry trends, Scarborough Research has achieved our highest survey response rates in the 27-year company history. Our syndicated survey, with over 200,000 respondents in local markets across the US, uses a two-phase survey method: a phone interview, followed by a mailed, self-administered survey booklet and television diary. In 2001, our phone survey response rate was 31.4%. By 2004, we posted a gain of 7.5 percentage points to achieve a 38.9% phone survey response rate (a 24% increase). Our booklet return rates have increased from 49.6% in 2001 to 56.2% in 2004. Similarly, our television diary return rates have increased from 34.5% in 2001 to 38.7% in 2004. We accomplished these gains not by any single technique, but by using a combination of many different initiatives across our organization. We built on knowledge identified in the research literature and previously presented at AAPOR, and applied it across our entire survey process. In our paper we will discuss in detail the changes and innovations applied to each survey component, which enabled our success. These include: experimental-design tests of different letters and cash incentives that identified effective and efficient survey treatments, including a Refusal Letter with Cash (RLWC); an overhaul of interviewer training incorporating adult learning techniques; adding a range of interviewer incentives; changes to the number and timing of sample call-attempts; increasing the number of respondent contacts; better analysis of interviewer performance metrics; enhanced interviewer recruiting and selection methods; and technological enhancements such as Caller ID and call-in surveys.

***Developing a Prescription for Physician Surveys*, Janice Ballou, Mathematica Policy Reseach Inc; jballou@mathematica-mpr.com, Brian Roff, Mathematica Policy Research; broff@mathematica-mpr.com, Julita Milliner-Waddell, Mathematica Policy Research; jmilliner-waddell@mathmatica-mpr.com, and Frank Potter, Mathematica Policy Research; fpotter@mathematica-mpr.com**

Conducting surveys of physicians is particularly challenging. Probably more than any other population segment they are regularly contacted for all types of surveys, and they have gatekeepers that prevent contact. In the fall of 2003, working with Harvard University, Mathematica Policy Research conducted a survey of physicians to obtain baseline information about the current status of the profession. A self-administered questionnaire was mailed to 3,504 physicians in six specialty areas from an AMA listing. Based on our experience and the methodological literature the mail approach included the following best practices to maximize response and reduce cost: (1) introduction letter from a respected academic entity (Harvard) and a well-known physician (Dr. Blumenthal); (2) altruistic appeal to the contribution the physician would make to the understanding of the profession; (3) enclosed pre-incentive check for $20.00; (4) used priority mail; (5) focused questionnaire that could be completed in about 15-20 minutes (6) promise of total anonymity, (7) reminder postcards, a second full mailing, and a second post card reminder, and (8) telephone reminder calls with follow-up new mail or faxed packets. The data collection period was from December 2003 to June 2004. Even though the initial questionnaire packets were mailed during the time between Thanksgiving and the winter holidays, 70 percent of the total number of completed interviews were returned within the first few weeks of the initial mailing. This paper will conduct an analysis of data quality for four types of respondents: Wave 1 first responders (1,160), Wave 2 middle responders(377) , Wave 3 late responders (125), and Wave 4 non-responders. In addition to the analysis of the data quality the paper will also assess the return on investment (ROI) for continuing the effort to improve the response rate.

***Examining the Use and Effectiveness of a Help Desk in Providing Support for Self Administered Web Interviews*, David Roe, RTI International; droe@rti.org, Lisa Carley-Baxter, RTI International; lcbaxter@rti.org, Jeff Franklin, RTI International; jwf@rti.org, and Rodney Baxter, RTI International; rbaxter@rti.org**

Researchers are beginning to utilize the World Wide Web as a mode of data collection, either as a stand-alone instrument or as part of a mixed mode approach. While the use of the Web as a mode of data collection presents some advantages over other methods such as mail or telephone, new challenges have arisen. One such challenge is ensuring that respondents are able to overcome technical difficulties they may experience while attempting a self administered Web interview. Respondents may also misunderstand, misinterpret or be confused about the meaning of a given survey question, and not have a trained professional to turn to for clarification as they would during a telephone interview. Some research organizations have begun providing toll-free telephone numbers to web respondents where trained professionals can assist them with issues ranging from lost passwords to item related questions. This analysis presents results from two such surveys; one of students enrolled in postsecondary education institutions in the US, and another comprised of postsecondary faculty, both sponsored by the US Department of Education's National Center for Education Statistics. We expect that receiving assistance from a Help Desk Agent leads to a completed interview in a high proportion of cases who call in for assistance. We also expect that among these completed cases, a relatively small amount of time elapses between contact with the help desk and completion of an interview. The analysis will also explore and present results on why respondents call help desks, the frequency of multiple calls from respondents and whether or not these multiple contacts were for assistance with the same or different issues.

***The More you know the Less you chat: General political knowledge as a predictor of online chat room use*, Masaki Hidaka, University of Pennsylvania; hidakam@yahoo.com**

A recent PEW Internet survey found that 16% of those respondents with Internet access went online to be part of chat rooms and other online discussions. An earlier report also issued by PEW found that of those who went to chat rooms and participated in online communities, many did so to find other individuals with shared interests or to discuss their feelings about significant events such as 9/11 or personal health scares. Yet personal interest, health concerns and significant events are not the only predictors of chat room use. Using survey responses from a representative sample of Internet users (The Electronic Dialogue 2000 Project from the Annenberg School at the University of Pennsylvania), this research found that after controlling for issues of personal interest and variables such as media use, low levels of general political knowledge predicted higher levels of online chat room use. This paper will attempt to illustrate this relationship using quantitative analysis and will investigate reasons as to why lower levels of general political knowledge are associated with chat room use.

***Using Proxy Reports of Quality of Life for Nursing Home Residents*, Annette Totten, University of Minnesota; tott0007@umn.edu**
Complex outcomes like Quality of Life are of interest in studies of health and social service programs that serve the disabled, critically ill, and elderly needing long-term care. The paradox is that a significant portion of any sample selected from these populations may be unable to complete a questionnaire/interview. At this point the researcher much choose between the error introduced by obtaining data only from those able to self report or the error introduced by obtaining the desired information from other sources. This study evaluated 1471 pairs of family and staff reports on a sample of residents from 40 nursing homes in 5 states which included residents able to complete an interview, nonrespondents, and 'poor' respondents (who answered a subset of questions). Agreement between proxies and between residents and proxies varied widely across residents and items. When answering 57 identical items about a resident, the mean exact agreement between the two proxies was 42.5% (range 8% - 87.5%); agreement within one category on items with a four-level response averaged 79.2%. Exact agreement on the specific items ranged from 73% on an item about leaving the nursing home to 23% when reporting how frequently clothing was lost. Characteristics that significantly contributed to explaining this variation included residents' cognitive and functional status (a nonlinear relationship with proxy agreement highest at the highest and lowest levels of function); family satisfaction with the nursing home; staff race; and the staff proxies' self-evaluations of how well they knew the resident. A priori classification of items into categories of predicted level of agreement based on item characteristics was better at explaining higher than lower levels of agreement. Understanding influences on proxy responses is important in assessing their validity and appropriate use when the population of interest includes those unable to self-report.

***A Cumulative Count, General Approach to Item-NonResponse*, Michael Wood, Hunter College, CUNY; mwood@hunter.cuny.edu**
Understanding of item nonresponse has been hampered by conceptual inconsistency and ambiguous empirical results. "Item nonresponse" covers a wide variety of survey non-substantive responses, from "don't know" answers for attitude items to refusal (missing data) for income questions. Theories about opinion "nonattitudes" or "satisficing" seem to have limited applicability to objective (e.g. income) question nonresponse, and vice versa. This paper proposes a complementary addition to existing conceptual and empirical perspectives on nonresponse: consideration of the quality of response across items. Survey investigators are typically interested in particular items and relationships in a survey. Respondents however answer survey items in a "longitudinal" context --the survey is a comprehensive event. Item nonresponse is posited to be the result of two fundamental processes: a component associated with a particular item, and a more general component associated with a respondent and the quality of his/her overall response to a survey. Empirically this approach means that nonresponse is cumulated across items in a questionnaire. Such cumulation can be organized by type of question (e.g. objective vs. attitudinal), as well as by type of outcome (e.g. don't know, not applicable, missing data). The data produced by across-item cumulation of nonresponse is typically skewed (80% or more zero count) and can be modeled using the Poisson and negative binomial distributions (especially 'zero inflated'). The outcome of such analysis can contribute to understanding of a general component of item nonresponse, as well as improved substantive understanding of the dynamics of nonresponse to particular questions. The approach is illustrated using data from a national survey.

***Incidence and Impact of Controlled Access Situations on Unit Nonresponse*, David Cunningham, RTI International; dbc@rti.org, Laura Flicker, RTI International; lflicker@rti.org, Joseph Murphy, RTI International; jmurphy@rti.org, Jeremy Aldworth, RTI International; jaldworth@rti.org, Susan Myers, RTI International; smyers@rti.org, and Joel Kennet, Substance Abuse and Mental Health Services Administration; joel.kennet@samhsa.hhs.gov**
The term "controlled access" applies to any situation where an obstacle keeps an interviewer from reaching the door of a potential respondent. Failure to reach controlled access dwelling units may introduce bias through systematic under-representation of certain sub-groups. For example, high-income and urban households are more commonly found in controlled access situations than other sub-groups in the United States population (Blakley and Snyder, 1999). In recent years, the National Survey on Drug Use and Health (NSDUH), a federally sponsored annual survey that gathers data on substance use and abuse among the non-institutionalized household population of the United States, has seen an increase in the amount of nonresponse attributed to controlled access. Through this experience, NSDUH researchers have developed protocols for overcoming barriers to data collection. However, until 2004 the project had not collected the data necessary to analyze controlled access at the dwelling unit level nor analyzed the effect of controlled access on unit or item nonresponse. In order to develop these capabilities, the NSDUH now systematically captures housing characteristics, controlled access status, and the outcome of all controlled access situations for all 180,000 housing units in the annual NSDUH sample. This presentation will summarize the incidence of controlled access by dwelling unit type, population density, and state, and introduce a model that predicts the effects of controlled access barriers on unit nonresponse. We predict that multi-unit dwellings, areas with more dense population, and areas with higher income will exhibit higher incidence of controlled access and that controlled access will be correlated with unit nonresponse. Finally, we will discuss ideas for investigating the role that controlled access barriers have on nonresponse error and data quality. Reference: Blakley, E. and Snyder, M. (1999). Fortress America: Gated Communities in the United States. The Brookings Institution.

***Mixed Mode Data Collection Using Paper and Web Questionnaires. A Cost and Response Rate Comparison in a Survey of Students Housing Conditions*, Peter Werner, Linkping University; pewer@mai.liu.se, and Gosta Forsman, Swedish Road Administration; gosta.forsman@swipnet.se**
In this study we examine two different mixed mode approaches, each offering the respondents to choose between a paper questionnaire and a Web questionnaire. These two approaches are compared with a traditional mail approach, as well as an e-mail approach, with respect to costs and response rates. A survey requesting students at Linkping University to provide data regarding their housing conditions was conducted during spring 2004, in collaboration with the statistical office of the local government of Linkping. The principal purpose of the study was to examine if mail invitation to a Web questionnaire, thus making probability sampling possible, differ in response rate and costs from a traditional mail survey. In an experiment embedded in the survey, a random sample from the student population was randomly divided into four groups distinguished by the type of invitation and questionnaire delivery: (1) a traditional mail invitation and paper questionnaire with one mail reminder including a paper questionnaire, (2) a mail invitation to a Web questionnaire including a paper questionnaire (thus offering the respondent a choice between Web and mail) with one mail reminder not including a paper questionnaire, (3) a mail invitation to a Web questionnaire with one mail reminder including a paper questionnaire (thus offering a choice in the reminder), and (4) for reference an e-mail invitation to the Web questionnaire and two e-mail reminders (all students at Linkping University have Internet access and individual e-mail addresses). Cost estimates and response rates per survey group will be presented in the paper. The findings will be compared and contrasted to similar studies reported in the survey literature.

***Assessing the Accuracy of Respondent Reports of the Location of their Home Relative to Geographic Boundaries and Other Characteristics*, John Baldridge, Bureau of Business and Economic Research, Univ. of Montana; john.baldridge@business.umt.edu, James Sylvester, BBER, The University of Montana; jim.sylvester@business.umt.edu, and William Borrie, Coll. of Forestry and Conservation, Univ. of MT; bill.borrie@umontana.edu**

Logging and forest fires are vital policy issues. Government agencies increasingly use survey research, especially random digit dial telephone surveys, to gather information about these issues. Respondent proximity to forests is an important variable in this research. Many researchers do not ask respondents for their address due to the perceived threat associated with this question. We assess the accuracy of responses to questions that ask respondents the location of their home relative to geographic boundaries without asking their address. Our analysis assesses the accuracy of respondent reports on forest cover in the area surrounding their home. We also validate respondent reports of their Zip code. The data that we assess were gathered in a random digit dial telephone survey (N = 1,164) of adult residents of Ravalli County, Montana. The survey examined respondents' trust of and satisfaction with the management of the U.S.D.A. Forest Service's Bitterroot National Forest and was conducted in May and June of 2004. Responses are validated by using reverse telephone directories to code respondent home addresses and then by assigning the address a geo-code. GIS software is then used to compare the respondents' reports of the location of their home relative to the boundary of the Bitterroot National Forest, the boundary of the nearest incorporated town, and the respondent's actual Zip code. Existing satellite imagery that has been enhanced by University of Montana's College of Forestry and Conservation will then be combined with GIS software to determine whether the respondent's home actually resides in a forested area. The survey has been completed and address geo-codes have been assigned. The remaining work includes coding the amount of forest cover around the respondent's home, analyzing the accuracy of respondent reports, examining the limits of the validation methods, and developing the poster presentation.

***Accuracy of reports of behavioral health service use among public assistance HMO members: Results from a record check study*, Scott Beach, University of Pittsburgh; scottb@pitt.edu, Janet Schlarb, University of Pittsburgh; jschlarb@pitt.edu, Donald Musa, University of Pittsburgh; dmuc@pitt.edu, and James Schuster, Community Care Behavioral Health; schusterjm@ccbh.com**

Since the advent of managed health care, patient satisfaction surveys have become an important component of quality improvement efforts by managed care organizations. Surveys of patient satisfaction with behavioral health treatment (psychiatric, substance abuse, counseling) involve special challenges. A recently developed instrument for the evaluation of behavioral health treatment, the Experience of Care and Health Outcomes (ECHOTM) survey, includes an initial question asking whether or not any treatment was received in the last year. This gives the member the option of not revealing and evaluating treatment that did in fact occur. While ensuring that patients can maintain privacy, this also raises the possibility of biased survey estimates. We have been using the ECHOTM instrument annually since 2001 to assess patient satisfaction with behavioral health treatment among members of Community Care Behavioral Health (CCBH), a managed behavioral health organization that coordinates behavioral care for public assistance enrollees in Allegheny (Pittsburgh), Berks, York, Adams, and Chester counties in Pennsylvania. This paper uses CCBH administrative data on behavioral health service utilization to validate patient self-reports of service receipt from the survey. Data are presented from a total of 8,795 respondents between 2001 and 2003. Analyses focus on predictors of reports of no treatment in the last year where administrative records show that, in fact, such treatment was received (i.e., false negatives). Predictors include mode of survey completion (self-administered mail vs. telephone; a mail with phone follow-up protocol was used), frequency of service utilization in the year prior to the survey, time since most recent service use, diagnosis, type of services used (inpatient, outpatient, both), along with various demographic characteristics. Adults reporting about their own treatment are also compared to parents of children receiving treatment. Demographic correlates of likely "memory-based" errors versus more deliberate misreporting are also examined.

***What's the benefit of a mail supplement to a telephone survey for gathering data from "hard to reach" populations?*, Kerry Levin, Westat; kerrylevin@westat.com, cynthia helba, Westat; cynthiahelba@westat.com, barbara forsyth, Westat; barbaraforsyth@westat.com, and Karen Masken, IRS; Masken Karen C [Karen.C.Masken@irs.gov]**

When conducting telephone surveys, multi-mode methods for contacting sampled individuals are the industry standard in order to maximize response (e.g., Dillman, 1978, 2000). With a list frame, incorporating modes other than telephone also enhances the likelihood that respondents without telephone numbers are included in the data collection. Of course, multi-mode designs are considerably more expensive than telephone-only studies. This study of low income, mobile taxpayers allowed us to explore factors that may affect the cost-effectiveness of supplementing a telephone survey with a mail survey among this unique population. We sent mail surveys to two different populations who could not be reached by our telephone interviewers--those for whom a telephone number could not be found and those with a telephone number but who were non-locatable (e.g., non-working telephone number). Our analyses examine demographic and attitudinal differences among two groups of taxpayers--those who responded by telephone and those who responded by mail. Among the mail respondents, we also examine differences between those without telephone numbers and those coded as non-locatable. Based upon these results, we identify factors to be considered when weighing the costs and benefits of a multi-mode design, especially when potential respondents are low income and mobile.

***Spanish-Speaking Households: Reaching Them Faster*, Lorayn Olson, Abt Associates Inc.; lorayn_olson@abtassoc.com, Martin Frankel, Abt Associates and Baruch College, CUNY; Martin_Frankel@abtassoc.com, Larry Osborn, Abt Associates Inc.; Larry_Osborn@abtassoc.com, and Sergei Rodkin, Abt Associates Inc.; Sergei_Rodkin@abtassoc.com**

On a random-digit-dialed telephone survey there is little information available about each sampled telephone number that might be used to tailor the strategy for completing each case. This paper presents an approach that makes use of information gained during a pre-dialing name/address match in order to enhance the efficiency of the data collection for cases requiring interview administration in Spanish. The National Immunization Survey (NIS) measures vaccination coverage among children aged 19-35 months. In order to interview 34,000 parents of age-eligible children, the NIS samples over 3 million telephone numbers annually. A name/address match for the selected telephone numbers is used to gain addresses for an advance letter mailing. An experiment designed to test the effectiveness of a new approach for cases more likely to require a bilingual interviewer was conducted. Surnames having at least one case where an NIS interview was completed in Spanish were identified. Included in the experiment were cases having one of the 6,000 surnames that had the highest probability of leading to a Spanish-language interview. These cases with selected surnames have been divided into two groups: an experimental group, in which the 11,500 cases are being delivered to bilingual interviewers from the first call; and a control group in which 11,500 cases remain part of the overall sample being delivered to the next available interviewer, routed to bilingual interviewers only if the

need is identified during a conversation with a household member. The impact of this alternative approach with respect to the number of dials required to complete an interview (or otherwise finalize a case) as well as the components of the overall response rate will be presented.

***Cognitive Interviewing en Espanol -- Extending the Realm of Cross-Cultural Pretesting*, Martha Hunt, National Cancer Institute; huntm@mail.nih.gov, and Gordon Willis, National Cancer Institute/NIH; willisg@mail.nih.gov**
Cognitive interviewing techniques are increasingly being extended to cross-cultural applications, where surveys are administered in multiple languages, and/or across sub-cultural groups. Many of these applications involve pretesting of survey questions intended for Hispanic respondents who speak Spanish as a first language. Such endeavors present a set of challenges to the questionnaire designers, who must concern themselves not only with the usual issues surrounding sources of response error, but must as well take into account several additional sources of error, including (a) the translation (and possibly mis-translation) of the instrument; (b)natural sub-group differences in language use or in survey response tendencies (e.g., between Mexicans, Cubans, and Puerto Ricans); (c) uncontrolled variation in cognitive interviewer behavior across staffs speaking different languages or representing different cultures; and (d) reconciliation of interview findings that may reflect any of the above factors, when attempting to develop a coherent sense of the testing results and their implications for question re-design. For the current study, we examined these issues through cognitive testing of a cancer-risk-factors questionnaire, involving three interviewing rounds in Spanish, across three Hispanic subgroups. As a comparison, two interviewing rounds were conducted on non-Hispanics in English. In all cases, the investigators used a testing protocol (i.e., a set of probe questions) that was standardized across groups, in order to minimize extraneous error. Further, a coding system was used to process results, again in an effort to standardize the approach and render the results more comparable. Our findings, along with implications for cross-cultural cognitive interviewing, will be discussed.

***Modeling Components of Response Propensity in Centralized Telephone Surveys*, Barbara C. O'Hare, Arbitron Inc.; barbara.o'hare@arbitron.com, Sheila Cross, Arbitron Inc; Sheila.Cross@arbitron.com, Sonja Ziniel, University of Michigan; sziniel@isr.umich.edu, and Robert M. Groves, University of Michigan; bgroves@isr.umich.edu**
Increases in noncontacts of sample households and reluctance of respondents to participate in a survey have led over the past two decades to declining response rates (de Leeuw and de Heer, 2002). The common reaction of survey organizations to these two phenomena was to undertake additional callbacks and increase refusal conversion efforts. Unfortunately, all these measures increased the costs of RDD surveys dramatically. Cost constraints make it desirable to customize the survey design in order to be able to minimize the increase of additional costs while efforts to reduce survey nonresponse are undertaken. Customization requires collecting information on sample performance effects of key design features, and adjusting the survey design during the survey process. These include what interviewer to assign to a case, how many callbacks to make, what refusal conversion protocol to use, and what incentive regimen to apply. Real-time propensity models for survey participation seem to be a suitable way to allow the development of a customized design and thus maximize nonresponse-reducing measures within a fixed budget (Groves and Heeringa, 2004: Lavrakas et al., 2004; Burks et al., 2004.) This paper identifies possible predictors for response propensities at different stages of the research process in a mixed-mode diary survey about radio listening by Arbitron. Prior data available on the sample numbers, interviewer attributes, calling history, as well as household characteristics are used to predict cooperation in the telephone survey and participation by sending back the diary. Significant predictors that discriminate the sample can then be used as leverages for maximizing the usefulness of response rate enhancing methods within a given budget and without introducing nonresponse error.

***Survey Outcomes and the Characteristics of Telephone Number Exchanges*, John Kennedy, Indiana University Center for Survey Research; kennedyj@indiana.edu, and Nancy Bannister, Indiana University Center for Survey Research; banniste@indiana.edu**
This presentation demonstrates how the information provided with list-assisted random digit dial telephone samples can be used to help understand survey nonresponse. List-assisted samples generally include additional information about the characteristics of the households in the area code/exchange. Often, this information is used to stratify samples to target area code/exchanges that have selected characteristics. For example, oversampling exchanges with high proportions of households with incomes over $100,000 can improve estimates of philanthropic giving (Rooney 2003a, 2003b, 2004). The information provided by list-assisted samples can also be used to understand survey nonresponse. Groves and Couper (1998) include environmental attributes as factors that help understand survey nonresponse. The characteristics of the households in the telephone exchanges can be used as indirect measures of these attributes. For example, exchanges characterized as having lower income households might indicate conditions that make potential respondents less willing to participate in telephone surveys. This presentation analyzes differences in survey outcomes using AAPOR codes for 3 surveys - a statewide RDD and 2 national RDD surveys. One national survey (the Timesharing Experiments for the Social Sciences) uses presurvey announcement letters and offers an incentive. The 2 national RDD surveys allow comparisons of their outcomes across exchange types but using different survey procedures. We analyze survey outcomes (completed interviews, refusals, etc.) by various types of exchanges. For example, we found that the response rates across 3 categories of exchanges - low, middle, and high proportions of African Americans - differed but not substantially. However, the reasons for nonresponse differed substantially across the 3 groups of exchanges. The outcomes from this research can be used to better understand ecological correlates of survey nonresponse; to estimate more accurately the number telephone numbers needed to accomplish a specific sample; and to explore the potential for targeting incentives based on exchange characteristics.

***Partial Interviews and Data Quality in a Large Telephone Screening Survey*, Kristen Olson, University of Michigan; olsok@umich.edu, Mary Cay Murray, Abt Associates; Mary_Cay_Murray@abtassoc.com, Jessica Cardoni, Abt Associates; Jessica_Cardoni@abtassoc.com, Michael Battaglia, Abt Associates Inc.; Mike_Battaglia@abtassoc.com, and Jim Lepkowski, University of Michigan; jimlep@umich.edu**
Partial interviews are an unexplored source of nonresponse and measurement error in telephone surveys. AAPOR standards permit partial interviews to be included in the numerator of response rate calculations; however, data quality can be seriously compromised when important questions come after the point of break-off. Additionally, once persuaded to complete the interview, partial interview respondents may put less effort into giving accurate answers or retrieving records necessary to assist recall. A recent investigation into break-offs in a large telephone screening study showed that approximately 25 percent of all completed interviews in a large national screening survey at some point had a partial interview. Additionally, partial interviews were in telephone exchanges with different

demographic characteristics and have different call histories than those cases that completed the interview without being a partial interview (Olson, et al., 2004). However, the impact of including these partial interview cases on survey estimates remains unknown. This paper takes the first look at the effect of partial interviews on survey estimates in a large, national RDD study that screens for a special population, the National Immunization Survey. This paper addresses the following questions: (1) Do completed interviews that had a partial interview put more or less effort into completing a questionnaire (e.g., level of item missing data; retrieval of medical records for use while completing the questionnaire)?; (2) Are completed interviews converted from partial interviews more or less likely to provide access to medical records?; (3) Do converted complete interviews differ from never partial complete interviews on important items in the questionnaire? On data found in medical records? On the difference between questionnaire and medical records? (4) What is the impact of converting these cases on statistics calculated from this survey?

***Presence of Others: Korean General Social Survey (KGSS) and General Social Survey (GSS)*, Youngshil Park, Survey Research Center, SKKU; youngsp12@skku.edu**

Presence of others in the survey has an effect on the response errors due to the limitation of privacy or confidentiality. Although much has been known about the presence of others in the West, little is known the effect of presence of others in the East. We examined the presence of others in Korean General Social Survey (2003) and General Social Survey (2002). Because Korean General Social Survey adopted General Social Survey in US, both surveys include same questions. Our preliminary analysis shows that only 45% of respondents were interviewed alone in Korea, but 70.1% of respondents in US. This difference appears to relate to the household structure reflecting sociocultural differences. Although people whose culture emphasizes collectivism or interdependence are more likely to be influenced by the presence of others, the effects of presence of others on attitudes about gender roles, marriage, and trusting others were very similar in both countries.

***Making sense of the third-person effect: A personality approach*, Yanjun Zhao, Southern Illinois University at Carbondale; yanjun@siu.edu**

The purpose of the present research was to investigate the relationship between the third-person effect and personality traits. Specifically, the personality traits studied were trust, sociability, optimism, and self-esteem. A self-administered survey was conducted with 218 participants as a convenience sample in Carbondale, IL in April 2003. The 31-item questionnaire measured personality traits, the third-person effect and demographics. Items measuring personality traits were taken from the Eysenck Personality Questionnaire and the Couch Trust Inventory. The results identified a signigicant third-person effect, but failed to support a relationship between the third-person effect and the personality traits as examined in this study. This study identified gender as a new predictor for the third-person effect. Females are more likely to exhibit third-person effect than are males. Although the personality variables examined in this study were not significant predictors, other personality traits still have the potential to predict the third-person effect.

***Testing a shorter interview for refusal conversion in a telephone survey*, Vasudha Narayanan, Westat; vasudhanarayanan@westat.com, Sherman Edwards, Westat; ShermEdwards@westat.com, Vickie Mays, UCLA; maysv@nicco.sscnet.ucla.edu, Stephanie Fry, Westat; stephaniefry@westat.com, and Susan Cochran, UCLA; cochran@nicco.sscnet.ucla.edu**

A shorter questionnaire/interview should result in higher response rates. However, literature on this topic is not definitive. The length-response rate relationship varies by the sample studied, content, saliency and mode. Also there is little in the literature on using a shorter interview for refusal conversion. Westat is conducting an experiment to test the effectiveness of a shorter interview as a refusal conversion method. The design is being tested in the California Quality of Life Survey, conducted for UCLA and funded by the National Institutes of Health. The sample was selected from respondents to the 2003 California Heath Interview Survey who agreed to be recontacted. All respondents receive $25 after completing the 45-minute interview on various health topics. This study has two categories of refusals. (1) The respondent refuses to participate in the interview after being read the study disclosure statement. (2) The respondent hangs up in the first few minutes, refuses before the study disclosure is read, or breaks off during the interview. By IRB ruling, only the latter category is eligible for nonresponse follow-up and thus is included in the experiment. Refusers are divided in two groups, both receiving a letter requesting them to talk to the interviewer. Group 1 receives the standard refusal conversion script and gets the full 45-minute interview. Group 2 is asked to complete a 20-minute interview. The authors will determine if there is a significant difference in the cooperation rate between the two groups and conduct a cost-benefit analysis.

***What we think others think: A motivated reasoning model of public opinion perception*, Lilach Nir, U of Pennsylvania; LNir@asc.upenn.edu**

Theorists posit that a public--unlike a mass of individuals--forms its opinions through a communicative action that fosters a shared universe of discourse, awareness of multiple viewpoints, and recognition of the legitimacy of opposition in a polity. Modern differentiated societies rely on mass-mediated news to communicate these multiple viewpoints on a scope and scale that reaches the citizenry. Whether individuals pursue information on others' political preferences, however, is another matter. While some are motivated to seek as much information possible, others either seek information that supports their own preference, or shun information altogether. This differential pattern of awareness has implications for people's assessment of collective preferences. In this paper, I develop and test a motivated reasoning (MR) model that explicates the relationship between accuracy goals (to reach correct conclusion) and directional goals (to reach preferred conclusion). Predictions regarding perceptions of collective preferences were tested using a large-scale probability-sample survey and experimental group data. Results show that MR (1) is a significant predictor of the ability to generate counter-attitudinal messages, i.e., reasoning a political opponent's point of view, (2) affects perceptions of public opinion--some people systematically overestimate support for their own opinion, while others systematically underestimate support. These predictions held for estimates of national-level opinion as well as estimates of the modal leaning in discussion groups. Implications of these findings for a theory of public opinion (spiral of silence) are discussed. The paper concludes with a normative reflection on the attainability of considered public opinion.

***Socioeconomic status, race, and ethnicity on information gain from public information campaigns*, Dianne Rucinski, University of Illinois-Chicago; drucin@uic.edu**
Research on the knowledge gap has explored individual and social factors that may influence the creation, maintenance, and demise of SES-based differences in information gained as the result of information campaigns. The finding that the information rich tend to get richer while the information poor either stay information poor or gain information at a slower rate has been replicated over time and on a number of topics, leading to a formalization of the research under the knowledge gap hypothesis (Tichenor, Donohue and Olien, 1970). The knowledge gap hypothesis holds that as information flows into a social system, those of higher socioeconomic position are more likely to acquire information at a faster rate than those of lower socioeconomic position, leading to knowledge gaps between higher and lower socioeconomic position populations, compounding existing disparities and inequities (Tichenor, Donohue, and Olien, 1970). While research on the knowledge gap phenomena has explored factors that can mediate the influence of socio-economic status on information gain, few studies have considered the influence of race and ethnicity on information gain in response to information campaigns. Using data from the National Survey of American Families (NSAF), we examine the impact of socioeconomic status, race, ethnicity, and other factors on social program awareness. The analysis considers awareness of two social programs targeted to those with low- incomes: the States Children's Health Insurance Program (SCHIP), and the Earned Income Tax Credit (EITC). Preliminary analyses indicate that race and Latino ethnicity are important additions to the traditional knowledge gap model of income and education, and mediate the influences of income, education, and awareness of community services. Implications for public communication campaigns, especially those involving eliminating racial and health disparities, will be discussed.

***Voter Turnout and Likely Voter Models*, Joanne Miller, Univeristy of Minnesota; jmiller@polisci.umn.edu, Lawrence Jacobs, Univeristy of Minnesota; ljacobs@polisci.umn.edu, Samuel Best, University of Connecticut; sam.best@uconn.edu, and Chase Harrison, University of Connecticut; chase.harrison@uconn.edu**
Why do Americans commit the time and effort to cast a ballot that is unlikely to affect the outcome of an election? The apparent "irrationality" of voting is one of the enduring questions in American politics. Three explanations have emerged to explain this puzzle. One suggests that the perceived importance of an issue motivates voters to exert the effort to vote. A second points to institutional arrangements (such as same day registration) that lower the costs of voting. A third points to early socialization that gets voters into the habit of participating. The 2004 election offers an opportunity to examine these explanations. First, turnout rose to its highest level in decades. Do the existing accounts explain this surge? Second, these accounts have received little comparative testing. A rigorous examination of the relative credence of each account would be quite valuable. Third, turnout in 2004 offers an important opportunity to examine the reliability of likely voter models. How well did they perform in anticipating the rise in turnout? This paper uses a unique panel design to analyze likely voter models and to reconsider the causes of voter turnout. In June and October, representative samples of adults in Iowa (which has a registration deadline) and Minnesota (which has same-day registration) was contacted via telephone and asked questions to gauge likelihood of voting, attitudes towards the candidates, and issue positions. Both samples for Minnesota and Iowa were re-contacted after the election. We test the accuracy (early and late in the campaign) of 3 likely voter screens - past voting behavior, interest in the election, and voter registration, and report the effectiveness of the screens at capturing new/young voters. This paper contributes to the ongoing debate among survey researchers about likely voter models and addresses enduring questions about the causes of voter turnout.

***Are Web Options Making a Difference?*, Geri Mooney, Mathemtaica Polify Research; gmooney@mathematica-mpr.com, Barbara Rogers, MPR; brogers@mathematica-mpr.com, Melissa Wood, MPR; mwood@mathematica-mpr.com, and Deborah Trunzo, Methematica Policy Research; dtrunzo@mathematica-mpr.com**
In an effort to maintain or encourage higher response rates, an increasing number of surveys are including a Web option. Up until 2002, N-SSATS was a multi-mode mail survey with telephone follow-up. A Web option was introduced in 2002, and to date, there have been three survey administrations with a Web option (2002, 2003, and 2004). In that time, the percentage of sample members who have chosen to respond by Web has increased nearly 75 percent, from around 17 percent to nearly 30 percent. Despite these gains in the percentage of Web respondents, how successful have these Web efforts really been overall? Using data from N-SSATS, we will look at such questions as: (1) Is including a Web option increasing response rates or are the Web respondents being plucked from other the existing modes; (2) To what extent are Web respondents enamored of this option, that is, once a Web respondent, do you tend to remain a Web respondent? and (3) how does the accuracy of the data provided by Web respondents compare to the accuracy provided by respondents in the other modes?

Saturday 12:00 pm-2:15 pm - AAPOR LUNCH PLENARY

2004 EXIT POLLS

***2004 Exit Polls*, Warren Mitofsky, Mitofsky International; mitofsky@mindspring.com, Kathy Frankovic, CBS News; kaf@cbsnews.com, and Fritz Scheuren, NORC; scheuren@aol.com**
The 2004 exit polls by Edison Media Research and Mitofsky International were done for the nation's news media and paid for by NEP (ABC, Associated Press, CBS, CNN, FOX and NBC). While there were no mistakes in any projections in either the presidential primaries or the general election, and the analytic data were and still are widely reported by the news media and others, there was controversy. Premature interpretation on election day of the early results by some news organizations, and leaked results reported on the Internet in early afternoon, led many to assume John Kerry was headed for victory. It also led to accusations of election fraud by a handful of academics, particularly for Ohio and Florida. A subsequent report by Edison/Mitofsky attributed the primary source of error in the exit polls to differential response rates by Bush and Kerry voters. Despite the thorough work done, however, considerable controversy remains.

COGNITION

***Using cognitive interviews to improve the quality of questions on working time in the Spanish Labour Force Survey*, José Luis Padilla, University of Granada; jpadilla@ugr.es, Miguel Angel Martinez, National Statistical Institute; mimartin@ine.es, Teresa Moratilla, National Statistic Institute; moratilla@ine.es, and Andres Gonzalez, University of Granada; andreito@ugr.es**

Statistics on working time are essential for revealing characteristics of the labour market. Organisation of working time may be measured in different ways, two of the most common being 'actual hours worked' and 'usual hours worked'. The Spanish Labour Force Survey is the principal source of information on working hours in Spain. The LFS asks respondents to provide information on usual and actual hours worked, as well as overtime during the survey reference week. In view of changes to be introduced in the Spanish LFS in 2005, various studies have been carried out to improve the quality of working time measurement. The present paper considers results of using cognitive interviews to improve the formulation of questions concerning actual working time, usual working time and overtime. A total of 57 cognitive interviews were conducted. Participants were selected on the basis of socio-demographic and educational characteristics of particular interest for the study: immigrants, young people, employees, housewives and people aged over 65. Interview protocol was designed with a view to obtaining information on the understanding of the concepts, the information retrieval and the formulation of answers to questions on the three variables. Results indicate that interviewees did not distinguish between the concepts of "actual hours worked" and "usual hours worked" during the survey reference week. It is also significant that immigrants related "usual hours worked" to the number of hours "settled" or "agreed" with the employer. Moreover, the concept of overtime was interpreted as paid working hours exceeding the usual number of hours. In addition to discussing the implications of these results in relation to the characteristics of study participants, the article proposes procedures for analysing cognitive interviews when the target questions are open questions.

***Evaluating Argument Repertoire: Opinion Quality, Political Sophistication, and Ambivalence*, Nam-Jin Lee, University of Wisconsin-Madison; namjinlee@wisc.edu**

The argument repertoire, which takes into account both the number and the balance of considerations not only supporting one's viewpoints but also favoring the opposite viewpoints, has been widely used to gauge the effect of various forms of public deliberation on opinion quality. However, a fundamental question of whether (and if it so, in what sense) the possession of large and balanced argument repertoires is a clear indication of holding a higher-quality opinion has rarely been empirically addressed. Moreover, since the two components of considerations arguably generated by different psychological and social mechanisms are mixed together in this measure, it is often difficult to differentiate it from such related concepts as political sophistication or ambivalence. This paper attempts to address these questions by carefully evaluating argument repertoire especially in terms of its conceptual and empirical roles in public deliberation process. For this purpose, this paper constructs multiple measures of argument repertoire using the 1992 and 2000 ANES open-ended items (e.g., likes/dislikes for presidential candidates and the good/bad things emerging from the Gulf War) and tests various models examining the underlying mechanisms generating argument repertoire and its consequences in terms of various democratic outcomes. The findings suggest that generation of one's "own considerations" and "opposing considerations" goes through different processes, leading to quite different outcomes. Results further show that these two components of one's argument repertoire tend to behave much like political sophistication and ambivalence, respectively, often operating as countervailing forces in the opinion formation process. Thus, simply employing a summary measure either by adding up or multiplying those two components masks an important mechanism of public opinion formation. This paper concludes by discussing the implication of these findings and other specific routes to further conceptualize and theorize cognitive outcomes of deliberation.

***Establishment Surveys from the Perspective of Socially Distributed Cognition*, Boris Lorenc, Department of Statistics, Stockholm University; boris.lorenc@stat.su.se**

From its introduction into the survey field in the mid-1980's, the cognitive view contributed considerably to understanding of the response process. Nowadays, the approach is an indispensable component of any serious surveying endeavour. Yet, during the preceding two decades there emerged some difficulties in cognitive science as a discipline. These unresolved issues, like the concepts of representations and of 'processing information', led to creation of alternative approaches, beginning actually as early as the 1970's. These came in various flavours, one of which is socially distributed cognition (Hutchins, 1995). It is to this specific approach that the attention in this paper is focused, for two reasons: - it has a potential to be a successful (and even necessary) complement to the usual study of cognitive processes related to responding, - it is particularly applicable to establishment surveying, which thus far has stayed behind cognitive developments in individual responding. A vehicle for demonstration was a Statistics Sweden census of pupils taking part in primary education in Sweden. The data collection process, knowledge of which the author gained through observation and interviews, turned out to be a socially distributed endeavour including several people, their skills and tools (paper, computers), and persevering through time by establishing social practices and by use of record data. An analysis of the "path" that pupils' data - eventually reported to Statistics Sweden - need to traverse was performed in the spirit of socially distributed cognition, indicating variables where measurement errors would seem more prone to appear. A mathematical tool for the analysis of the paths is also sketched, and generalisations are given using the examples published in previous works. Reference Hutchins, E. (1995) Cognition in the Wild. Bradford: MIT Press.

***When do respondent misconceptions lead to survey response error?*, Anna Suessbrick, New School for Social Research; anna suessbrick@hotmail.com, Michael Schober, New School for Social Research; schober@newschool.edu, and Frederick Conrad, University of Michigan; fconrad@isr.umich.edu**

In a survey about people's use of tobacco and their opinions toward tobacco use, respondents have been shown to interpret terms like "smoking" and "cigarette" in surprisingly variable ways (Suessbrick, Schober, & Conrad, 2000). We carried out a detailed analysis of 104 telephone survey interviews about tobacco use in which we tracked when and how interviewers clarified survey terms for respondents. We also had access to details of respondents' conceptualizations about the terms in the questions as well as whether their responses changed when they were re-interviewed with the request to answer according to standard definitions. For the 86% of answers that were reliable (did not change on re-interview), respondents' conceptualizations differed from the survey designers' more than half the time on components of the definition that were not essential for their answers. Across the survey, nearly 12% of answers were unreliable on re-interview, even when respondents had been provided clarification during the interview. Closer analyses of the transcripts demonstrate that respondents' conceptual alignment with the survey designers was less critical to their responding accurately than the relevance of the concept (or concept dimension) to the respondent's personal circumstances (e.g., a periodic smoker, who doesn't inhale, correctly interprets "smoking" to

include any puffs). In fact, very minimal conceptual alignment could yield reliable answers as long as the area of conceptual overlap was relevant to the respondent's circumstances. Overall, clarification alone was not enough to guarantee reliable answers; clarification needed to be helpful to the respondent in resolving her particular misconception. Results suggest that clarification during a survey, though valuable under some circumstances, is not always needed or effective. On the other hand, preventing clarification, as advocated under the most strictly standardized approaches to interviewing, is insufficient to address the high and unpredictable rates of conceptual variability among respondents.

***Who is being heard? Item non-response bias and response quality bias in open ended comments.*, Mark Andrews, ORC Macro; mark.w.andrews@orcmacro.com**

A great deal of research has been conducted in order to describe and measure both unit and item non-response bias on quantitative surveys. Relatively less work, however, has been done to adapt this work to open-ended responses. This lack of attention is unfortunate because the topic is relevant for a number of reasons. First, while some purists treat verbatim comments as purely qualitative data, most researchers go through great lengths to code and tabulate responses for analysis. If a large bias is detected, this method of quantifying qualitative data will be brought to question. Second, this topic is important because there are reasons to believe that item non-response bias is greater for open ended comments. Open-ended comments have traditionally low response rates and there are also obvious education, language, or other demographic issues that may have an effect on the likelihood and quality of response. Finally, open ended responses have an added dimension-response quality. Item non-response, therefore, is not the only source of bias in verbatim comments. This paper will examine the issue of item non-response bias and response quality bias in a large (13000+ respondents) government employee survey. This survey provides a substantial quantity of demographic data including income, race/ethnicity, gender, age, education, and whether the employee is in a supervisory position. To measure item non-response bias, we will develop a logistic regression model to predict response likelihood to the one open-ended question in the survey. For the second stage, we will use length of response as a proxy for response quality. Again we will compare means and develop a regression model to predict response length using available demographics. To the extent possible, findings will be corroborated with other survey data to determine if the results are unique to the institution or are more universal in nature.

CONSUMER EXPENDITURES/FINANCES

***Relationships Among Expenditure Reporting Rates, Household Characteristics and Interview Processs Variables in the U.S. Consumer Expenditure Interview Survey*, John Eltinge, U.S. Bureau of Labor Statistics; Eltinge.John@bls.gov, and Moon Jung Cho, Bureau of Labor Statistics; cho.moon@bls.gov**

In the U.S. Consumer Expenditure (CE) Interview Survey, consumer units (roughly equivalent to households) are asked to provide item-level monthly reports of expenditures. For this survey, reporting rates (defined as the proportion of expenditure categories in which a consumer unit reports a purchase in a given month) are potentially useful indicators of data quality. Cho, Eltinge and Steinberg (2004) reported preliminary results on relationships between reporting rates and some characteristics of interviewers and the timing of the interview. The current paper extends this previous work in four ways. First, we present detailed results from logistic regression modeling of reporting rates with predictor variables associated with household characteristics (e.g., family size and composition, as well as prior-interview expenditure reports) as well as variables associated with the interview process (e.g., the number of respondents and the use of records by the respondent during the interview). Second, we supplement standard development and significance-testing results with diagnostics for the predictive power of the resulting models. These diagnostics are especially important for evaluation of the potential practical benefits of modification of field procedures suggested by the modeling results. Third, we extend these results to some specific groups of expenditure categories, and explore differences among the results for these groups. Finally, we compare and contrast these empirical results with previous literature on response and reporting rates in single-wave and panel surveys.

***Learning to Say No: Conditioned Underreporting in an Expenditure Survey*, Nhien To, BLS; nhien@bls.gov, and Jennifer Shields, Bureau of Labor Statistics; shields_j@bls.gov**

The Consumer Expenditure Interview Survey (CEQ) is a comprehensive federal expenditure panel survey. Respondents are asked to provide detailed information about their expenditures over the course of five quarters. For each reported expenditure, the respondent is asked a series of detailed follow-up questions; the more expenditures a respondent reports, the longer the interview becomes. Previous research has shown respondents may become conditioned to provide negative responses to reduce interview length (Jensen & Edelbrock, 1999; Lehnen and Reiss, 1978). In addition, survey methodologists suggest respondents realize answering a question a certain way can lead to additional questions (Fowler, 1983; Sudman and Bradburn, 1982). A study looking specifically at CEQ data found evidence in some commodities of a decline in mean expenditures across interview waves (Silberstein & Jacobs, 1989). Our study hypothesizes that respondents may become 'conditioned to underreport,' that is, they learn to say 'no' to questions about purchases after discovering that responding 'yes' increases interview length. To find evidence that interview length may condition underreporting, we propose to examine rates at which respondents report no expenditures. Focusing on these rates, as opposed to expenditure values, allows inspection of reporting patterns across waves without consideration of the amount of purchase reported. This study will use respondents who started the CEQ over a one year period, approximately 9,000 cases. Using respondents who reported a purchase in a previous interview, the probability that they report the same type of purchase in subsequent interviews will be estimated. If the respondent is conditioned to underreport across interviews, the probability of reporting a purchase in subsequent waves will be lower if they reported the same type of purchase in a preceding quarter. It is also hypothesized that demographic characteristics might impact the likelihood that respondents are conditioned to underreport. Our analysis will also examine this possibility.

***Consumer Sentiment and Consumer Spending: The Gambler's Fallacy*, Nat Ehrlich, Office for Survey Research, Mich. State University; nathaniel.ehrlich@ssc.msu.edu**

Economic theory states that when consumers feel that their personal financial situation is poor they react by spending less, and when their personal financial situation is good they spend more. The reason for this is that economists believe that all economic trends tend to continue until some major situational event [e.g. job loss] reverses the direction of the trend. By contrast, the "Gambler's Fallacy" is the belief that everything in life is the product of random fluctuation; for example, if a coin comes up "heads" three times in a row, there is an increased likelihood that the next coin toss will come up "tails". This attitude predicts that consumers who feel that their personal financial situation is poor right now will be likely to spend more, in anticipation of an improvement. Michigan State University's Office for Survey Research has

conducted a quarterly survey of Michigan adult residents from 1994 to the present, amassing over 35,000 responses. First order results - examining the change from quarter n to quarter n+1 and relating whether the change has been positive or negative to a measure of consumer spending in quarter n+2 - indicate that the Gambler's Fallacy model is a better predictor of behavior than economic theory. In addition, this paper will present analyses of the length of trends in both consumer sentiment and the relation of absolute value of consumer sentiment to same-quarter and next-quarter spending.

***Bilingual Approaches and Influences on Measurement and Data Collection: Evidence from the Survey of Consumer Finances*, Man-chi Mandy Sha, NORC; sha-mandy@norc.org, and Leslie Athey, NORC; athev-leslie@norc.org**
Interviewer effects on non-response and other measurement errors may be traced back to various design factors such as interviewer selection and training. However, the significance of bilingual approaches and influences has only recently generated more interest among survey researchers. With the increasing need for, and attention to, collecting and measuring survey data in a language other than English, this issue merits the attention of survey researchers. The 2004 survey of the Survey of Consumer Finances (SCF), which fields interviews with US households every three years on assets and finances, offered an interesting opportunity to provide evidence on bilingual approaches and influences on measurement and data collection. For a survey that is not traditionally known for its Spanish component, the 2004 survey saw an exponential growth in its numbers of bilingual interviewers and interviews administered in Spanish. We will discuss the merit and practice of developing a comprehensive protocol on bilingual interviewer selection and training developed by the Culture Methodology Program at NORC using data from a post-interviewer training survey administered to 23 Spanish certified bilingual interviewers. To measure its influences, we will look at interview length and production by language, and measurement differences between interviews completed in English and Spanish by bilingual interviewers. Finally, we will look at reasons for attrition specific to bilingual interviewers.

IMPORTANCE OF RESPONSE RATES

***How different are hard-to-reach respondents? Is it worth the effort to track them down?*, Mary Outwater, University of Oklahoma; outwater@ou.edu, and Jeanette A. Norris, University of Oklahoma; jeanette.a.norris-1@ou.edu**
A great deal of effort is put into increasing response rates in telephone surveys based on the premise that higher response rates reduce non-response bias and increase generalizability. Making numerous call attempts to reach hard-to-reach respondents is a common way survey organizations use to increase response rates. Logically it would seem that reaching these elusive populations for interviews is crucial, since they may have different behaviors and/or opinions. The general notion is that if we exclude such populations, we may get a biased aggregate estimate of the populations' opinions and attitudes. However, do the opinions and attitudes of hard-to-reach respondents really differ from the rest of the population? Would excluding these populations bias the overall distribution of opinions? This is an area not yet fully investigated but be explored in this research. Survey data from both a national study as well as a local (state of Oklahoma) study will be used to investigate the differences between respondents who are reached quickly and easily and those who must be called numerous times before a survey can be completed. The purpose of this research is to determine how much effort should be made in contacting respondents. If there is a great deal of difference (in terms of opinions, behaviors, and/or demographics) between those who are readily accessible and those who are not, it is very important to pursue this portion of the sample and increase the validity of the data. On the other hand, if there is little to no difference, the labor and resources required to reach this group might be put to better use in another manner.

***The relation between likelihood of response and data quality in the American Time Use Survey*, Scott Fricker, BLS; fricker_s@bls.gov**
Federal statistical agencies spend a good deal of effort and money in an attempt to improve response rates. Recent papers by Keeter et al (2000) and Curtin et al (2001) suggest that the higher response rates achieved through these attempts may not lead to gains in nonresponse bias reduction. These findings are in opposition to the generally held notion that nonresponse reduction procedures help to reduce nonresponse bias, and seem to suggest that such efforts are inefficient and perhaps unnecessary. The primary concern with missing data from nonresponse is the risk that they may cause bias in the survey estimates. However, information on nonresponse bias is rarely available, and nonresponse rates are at best only indicators of the potential for bias and of the quality of the data collection procedures. More attention could usefully be given to nonresponse bias, in particular, to the potential bias resulting from the poor data quality of respondents who are reluctant or otherwise less likely to respond to a survey request. The present study examined the link between response propensity and measurement error using data from the Current Population Survey (CPS) and the American Time Use Survey (ATUS). The aims of this research were: (1) to explore the usefulness of utilizing information available from the CPS and ATUS sample frames to predict response propensity (both contactability and probability of cooperation given contact); (2) to examine the interaction between response propensity predicted from CPS and nonresponse reduction techniques on final case disposition from ATUS; and (3) to look at the effects of nonresponse reduction efforts and predicted response propensity on response data quality. The analyses were based on propensity models derived from relevant theories of survey response, and latent class models of measurement error using a variety of quantitative and qualitative data quality indicators.

***Are Lower Response Rates Hazardous for Your Health? Do Higher Response Rates Translate Into Better Estimates of Health Insurance Coverage and Access to Care?*, Michael Davern, University of Minnesota; daver004@umn.edu, Kathleen Thiede Call, University of Minnesota; callx001@umn.edu, Meg Brown Good, University of Minnesota; mbgood@umn.edu, and Jeanette Ziegenfuss, University of Minnesota; zieg0100@umn.edu**
Response rates for random digit dial surveys have been falling over recent years. Recent Pew studies (Pew Research Center 2004) have found that national surveys with response rates as low as 27 percent can be as representative as surveys with 51 percent response rates on opinion, civic engagement and attitude items. These studies have pointed to a non-response mechanism that meets the criteria of "missing at random" as opposed to "missing completely at random" (Little and Rubin 1987). We examine whether this holds for health insurance and health care access variables from statewide surveys. Low response rates may lead to biased estimates of state health insurance coverage and access. We examine two recent surveys conducted by the University of Minnesota for the states of Oklahoma (n=5,847, AAPOR response rate #4=45%) and Minnesota (n=13,512, AAPOR response rate #4=56%). Using these data we estimate the probability of being uninsured, having different types of insurance coverage, and lacking access to care by whether the household refused to participate during a previous call, and whether the household took 5 or more days to be completed. Although certain demographic characteristics varied significantly between the two groups such as age (showing the data were not "missing completely at random"), there are no statistically significant

differences in multivariate models predicting key health access and health insurance coverage estimates controlling for the demographic differences (i.e., our data meet the criteria for "missing at random."). Not including the initial refusals and surveys completed after 5 days would result in response rates that are half of the actual rates but would not affect the quality of our estimates after imposing weighting controls for demographic variables. Thus we should consider developing additional summary measures of survey quality that are related to the estimates generated from the survey.

***High Response Rate or Better Data Quality? Examining the Trade-offs for an Establishment Survey*, Lauren Harris-Kojetin, Institute for the Future of Aging Services/AAHSA; lharris-kojetin@aahsa.org, and Kristen Kiefer, The National Council on the Aging (NCOA); kristen_kiefer@hotmail.com**

Survey researchers desire both a good response rate and quality data within the limits of available resources. We report on data collection results from a mixed-mode representative sample survey of 800 continuing care retirement communities (CCRCs) for seniors, where the targeted respondent was the CCRC director. The CCRCs belong to the American Association of Homes and Services for the Aging (AAHSA)-a national membership organization. To achieve a better response rate than in previous establishment surveys of AAHSA members (averaging 20% for mail surveys), we drew on Dillman's recommendation to be flexible. We provided sample members multiple modes to respond to a 45-item survey about resident participation in physical activity. Our data collection design included: awareness messages about the upcoming survey using AAHSA's established channels for communicating with members; an advance letter; U.S.-mailed questionnaire; U.S.-mailed reminder post card; e-mail reminders; phone call reminders; and, the option to complete the survey by web. By adding the web option, we increased survey completions by 47% (adding 128 web completions to 270 mail completions, for an overall response rate of 50%). However, web respondents had four times higher average item non-response than mail respondents (12% to 3%, $p < .01$). A similar inverse relationship between unit- and item-level non-response occurred in a randomized experiment with a subset of the sample ($n = 157$) for which we had no e-mail addresses (for whom we only offered the mail option). The reminder postcard treatment group (no advance letter) had a higher response rate than the advance letter treatment group (no reminder post card) (35% to 23%, $p = .10$). Yet, the reminder postcard treatment group had higher average item non-response than the advance letter treatment group (4% to 3%, $p < .01$). We will discuss this inverse relationship and its implications for establishment survey design.

DESIGNING QUESTIONNAIRES

***Fostering Storytelling: A Qualitative Approach to Enhance Questionnaire Development*, William Mickelson, Ph.D., Chamberlain Research Consultants; wtm@crcwis.com, and Jon Harrington, MS, Chamberlain Research Consultants; jharrington@crcwis.com**

Recently, it has become popular to use open-ended interview questions prior to survey development in a pseudo-qualitative research pre-study into public attitude and opinion. The goal of such an activity is to uncover key issues or concerns and then incorporate them into a choice-based survey questionnaire. Particularly in business and marketing, key decision points often drive the type of open-ended questions that are posed. For example, when a bank is considering what types of products and services to offer, it might pose the survey question, "What are the most important factors when choosing a bank?" The research goal behind such a question is to gain insight into potential business differentiators. In addressing the bank's open-ended questions, respondents will often answer from a predictable array of responses, like low fees, convenience, or flexible hours. Unfortunately, respondents are not always creative enough to spontaneously give sufficiently deep answers that truly address the research goals. If the public was allowed to tell stories about experiences, services, or beliefs, what would be learned? In the bank example, if respondents were encouraged to describe a time when they wanted to scream at their bankers, you would receive completely different responses. By using a "storytelling" research technique prior to survey development, researchers can assess emotion and potential attributes or hot buttons that are truly important to the public through meaningful language, with the potential to uncover issues that otherwise may have been missed or misunderstood by the researcher. This presentation synthesizes findings and experiences gleaned from putting the "storytelling" technique into practice in multiple research projects with various clients over the past 18 months. Salient examples to illustrate issues and improvements in questionnaire development will be given. In addition, recommendations for best-practice implementation of "storytelling" techniques will be offered, as well as pitfalls to be avoided.

***A Picture is Worth a Thousand Words: Results from Cognitive Testing of Graphic Show Cards*, Anne Hartman, National Cancer Institute/ DCCPS/ Risk Factor Monitoring & Methods Br.; Anne Hartman@nih.gov, Jennifer Crafts, Westat; jennifercrafts@WESTAT.com, Suzanne McNutt, Westat; susiemcnutt@westat.com, Gordon Willis, National Cancer Institute/NIH; willisg@mail.nih.gov, Tracey Summerall, Westat; traceysummerall@WESTAT.com, and Amy Yaroch, National Cancer Institute/ NIH; YarochA@mail.nih.gov**

Previous testing suggested item wording alone is insufficient to convey to survey respondents which types of hats provide complete protection from the sun. Respondents often inappropriately include commonly worn hats such as baseball caps and sun visors in their responses. The research question was whether showing pictures of the types of caps and hats to consider would lead to more accurate reporting, make it easier for respondents to understand and respond to the questions, and require less response time. In preparation for fielding the 2005 National Health Interview Survey (NHIS) Cancer Control Topical Module, three respondent show cards were developed that contained 3, 4, and 5 pictures each of a range of "hats" commonly used outdoors to protect the face, ears and neck from the sun. A lead-in question about use of baseball "caps" and sun visors was also developed along with a 2-picture show card to determine whether asking about cap use first would lead respondents to more accurately report hat use. Nine respondents over 18 years old with a mix of demographic characteristics were recruited for cognitive testing. Each of the respondents was given different versions of graphics and questions, varying the order. Most preferred the five picture version of the hat card because it made them think of all types of sun-protective hats; triggered memory, thus reducing respondent burden; and allowed more accurate reporting. Regarding caps, respondents understood the two distinct types when asked the question and preferred no picture card. The hat pictures were to be viewed generally (not focusing on details) so that they would represent a range of appropriate hats. However, respondents tended to interpret the hat card specifically (rather than generally) when preceded by a cap picture card. Implications of selecting pictorial versus text only approaches to question design will be discussed.

***Using cognitive procedures and traditional pilot studies to improve the quality of questions concerning earnings in the Spanish Labour Force Survey.*, José Luis Padilla, University of Granada; jpadilla@ugr.es, Miguel Angel Martinez, National Statistical Institute; mimartin@ine.es, Teresa Moratilla, National Statistical Institute; moratilla@ine.es, Andres Gonzalez, University of Granada; andreito@ugr.es, and Isabel Esteban, National Statistic Institute; iesteban.con@ine.es**

Data concerning earnings are important for a variety of users: political leaders, private companies, trades unions, etc. In Spain, the most important source of information on the labour market is the Labour Force Survey (LFS). The LFS is a three-monthly panel survey which collects data from all members of the houshold using a sample of approximately 70,000 households. On all occasions the survey is conducted by means of a face-to-face interview. This paper considers various studies designed to analyse the viability of introducing questions concerning earnings, in which different question format and interview methods were studied. The principal sources of error analysed in these studies are as follows: the effect of order when enquiring about gross earnings and net earnings; the closed or open format of the question; the precision of information provided by proxy and the impact of presentation letters prior to the interview. The paper presents the results of 40 cognitive interviews and two focus groups. Participants were selected on the basis of socio-demographic and educational characteristics. In addition, the different proposals were examined by means of a traditional pilot study in which 270 subjects took part in a survey with a quasi-experimental design. Participants in the cognitive procedures agreed on the need for a previous presentation letter. Analysis of the interviews reveals that interviewees respond with greater precision to questions on net salary, and show a preference for questions situating their salary in bands. Results also indicate the lack of precision of information provided by proxy. In the pilot study, question format was not shown to make a difference. Again, questions regarding net salary achieved a higher answer rate. Finally, the article proposes criteria for analysing the convergence between results obtained in traditional pilot studies and cognitive procedures during the pre-test phase of questionnaires.

***The Importance of Properly Measuring Importance*, Amy Gershkoff, Princeton University; agershko@princeton.edu**

What is issue salience and how can we measure it accurately in survey research? These questions have been hotly debated by psychologists, political scientists, and experts in survey methodology, but we have yet to reach a theoretical or methodological consensus. In this paper, I offer a new way to measure issue salience, using open-ended questions from the National Election Study, and I demonstrate why I believe this method is more accurate than previous methods. Moreover, the way we measure issue salience has implications for the conclusions we draw about the role issues play in the citizens' voting decisions. Using previous methods of issue salience measurement, numerous scholars have concluded that issue voting, defined as issue positions having an effect on vote choice that is separable from party identification, does not occur in American elections. Using my method of measuring issue salience, however, I find that a sizeable percentage of the electorate does engage in issue voting. In the paper, I analyze the historical role that issue voting has played in several previous presidential elections, and I find that in recent times, the propensity of voters to become "single-issue voters" has dramatically increased. For example, using my methodology to analyze the 2000 presidential election, I find that a sizeable percentage of registered Democrats voted for Bush because of his position on abortion. However, I show that using traditional methods of measuring issue salience, we would find abortion having no impact separable from party identification. I replicate this result across a tremendous number of issues and several elections. My results herein have implications not only for scholars of voting behavior, but also for survey researchers more generally who want to have an improved tool for assessing how consumers weight various factors in decision-making.

VOTER INTENT AND CHOICE

***To Vote or Not to Vote?: A Comparison of Vote Intention Measures*, Randall Thomas, Harris Interactive; rthomas@harrisinteractive.com, Mitchell Sanders, Harris Interactive; msanders@harrisinteractive.com, Renee Smith, Harris Interactive; rsmith@harrisinteractive.com, and Susan Behnke, Harris Interactive; sbehnke@harrisinteractive.com**

Through a series of web-based studies conducted in 2000 and 2004, we examined the efficacy of a variety of vote intention measures. Specifically, in 5 primary elections in 2000 (NH, CA, OH, GA, NY), in the 2000 general election, and in the 2004 general election we examined numerous variants of vote likelihood. In all studies we recontacted respondents within a few days of the election to determine whether or not they voted. Some vote intention measures appeared to be superior to others and we discuss the implications for determining appropriate likely voter models in pre-election polls.

***Voter Intent, Voting Technology and Measurement Error*, Frederick Conrad, University of Michigan; fconrad@isr.umich.edu, Emilia Peytcheva, University of Michigan; emiliap@isr.umich.edu, Michael Traugott, University of Michigan; mtrau@umich.edu, MIchael Hanmer, Georgetown University; mjh72@georgetown.edu, Paul Herrnson, University of Maryland; pherrnson@capc.umd.edu, Ben Bederson, University of Maryland; bederson@cs.umd.edu, and Richard Niemi, University of Rochester; niemi@mail.rochester.edu**

In principle, elections directly measures the public's political preferences. However, ambiguity about voter intent was at the heart of the contentious Florida recount in the 2000 presidential election. New computer-based voting technology promises to eliminate the ambiguity of partially punched chads but may introduce usability problems that obscure voters' intent in other ways. In the proposed talk we report a usability study of six electronic voting machines that examines voters' problems when they use this technology - technology that will be used throughout the US as the Help America Vote Act of 2002 is implemented. Forty-two community members, most selected for their limited computer experience, voted in a mock election on each machine. Voters first indicated their intentions by circling candidates and ballot questions in a paper voter booklet and were instructed to vote accordingly. Then they (attempted to) cast those votes on each of the machines, advancing through the machines in different random orders. Their interactions were videotaped and they completed a satisfaction questionnaire after voting on each machine. Thus our data concern voting time, satisfaction (based on the questionnaire) and frequency of usability problems (based on the videos) for each machine. We interpret differences between the preferences in the voter booklet and the votes actually cast as reflecting particularly serious usability problems. Our discussion distinguishes between problems that distort voter intent (e.g., voting for someone other than intended) and those that occur even though the voter ultimately casts the intended vote (e.g., difficulty changing votes because the process is counterintuitive). We also address the usability problems introduced by printed records of what appears on the touch screen, the method of verifying electronic votes advocated by many. We conclude by discussing how problems might affect voter turnout and how designers can overcome some of these problems.

***Measures of difficulty in election polling*, Joseph Shipman, SurveyUSA; jshipman@surveyusa.com, and Jay Leve, SurveyUSA; jleve@surveyusa.com**

Some elections are harder to poll than others. Therefore, it is not fair across disparate elections to calculate one pollster's average "error" and compare it to another's, since one pollster may have polled disproportionately "easy" contests and the other may have polled disproportionately "hard" contests. Data from 104 polling organizations in 120 statewide November 2004 election contests are analyzed. The 15 most active statewide pollsters are ranked. Two measures of pollster performance that take "degree of difficulty" into account are developed. One measure borrows from golf, and assigns a "par" score for each different type of contest (President, State Office, Ballot Measure). Pollsters are judged to have scored below par or above par. The most active 15 pollsters are ranked on a Leader Board. A second measure borrows from Olympic Figure Skating, and assigns an ordinal rank to the accuracy of each poll that a pollster conducted, relative to all of the other polls conducted in the same contest. The average of each pollster's ordinal rankings is then compared to determine, using this approach, which statewide pollsters were Medalists in 2004.

***Moment by Moment: Analyzing the Presidential Debates*, Anthony Salvanto, Ph.D., CBS News; salvantoa@cbsnews.com**

The 2004 Presidential debates had a dramatic impact on the campaign: the belief that John Kerry won them brought him back to even in the polls, where he'd been trailing. But the fullest understanding of how the debates shaped views of the candidates, and of how views on who "won" them were formed, comes from analyzing voters' real-time reactions to what they saw during those face-offs. During the debates, CBS News (partnering with Knowledge Networks) tracked moment-by-moment reactions of a nationally-representative panel of uncommitted voters while they watched the debates in their homes, recording reactions on a "thermometer" scale via Web-enabled TV's, in real time. They also completed a survey before and immediately after the debate. This paper shows how the panel responded to stimuli from the candidates and analyzes how those reactions impacted views of who won or lost the debate, and affected broader favorability ratings of the two men. It also shows whether important stimuli that emerged in the debates foreshadowed the reasons that voters gave for their choices on Election Day. The data show how Kerry's policy plans won him consistent favor and drove the view that he had won relative to Bush, but rarely hit the highest levels on the meter. Bush made fewer statements that beat Kerry's, but his mentions of faith and morality prompted the highest meter spikes seen from either man. (Paper presentation will include video clips with the meter ratings alongside.) Data on an individual-respondent level are used to explore methodological issues in detail and how they affected the results here -- e.g., whether certain types of viewers are pre-disposed to marking extreme ratings; how uncommitted voters leaning toward one candidate going into the debate reacted differently than the purely undecided - and their implications for broader research of this type.

CROSS-NATIONAL RESEARCH II

***The Changing Patterns of Religious Practice and Belief in Western Europe: A Cross-National Cohort Analysis*, Allan McCutcheon, UNL-Gallup Research Center; amccutcheon1@unl.edu**

One of the several elements of secularization includes the declining role of organized religion in modern society. In particular, a number of survey researchers have focused on recent declines in self-reported church attendance and religious identification as possible indicators of the declining role of organized religion in the U.S. and Europe (e.g., Ester et al. 1994, Bruce 2002). Other researchers challenge this interpretation arguing that many respondents who claim no religious identification, and who report low, or no, church attendance continue to affirm questions on religious beliefs (e.g., belief in God, life after death)-the so-called "unchurched believers" (Hout and Greeley 1987, Hout and Fischer 2002), or "believers without belonging" hypothesis (Davie 1999, 2000). This research focuses on a log-linear cohort analysis of eleven Western European nations, and examines the "believers without belonging" hypothesis. Data for the eleven nations are taken from representative national samples in the 1981, 1990, and 1999 European Values Surveys (EVS) and the 2002 European Social Survey (ESS). The analyses indicate only weak support for the "believers without belonging" hypothesis, and then, largely among the older cohorts. Younger (post-WWII) cohorts express substantially lower levels of religious identification, church attendance, and traditional religious beliefs. In comparison to the Protestant (Lutheran) nations and mixed religion nations of Europe, most of the Catholic nations of Europe (excepting France) appear to have delayed this movement away from identification, attendance and belief. Among recent birth cohorts, however, the Catholic nations appear to moving away from traditional practice and belief at a pace not too different from the other nations.

***Equivalence of the Schwartz Value System across 20 European countries - Results from the ESS 2002*, Peter Ph. Mohler, ZUMA - Centre for Survey Research and Methodology; mohler@zuma-mannheim.de, and Katrin Wohn, ZUMA - Centre for Survey Research and Methodology; wohn@zuma-mannheim.de**

Numerous experimental studies across a wide range of cultures indicated cross-cultural equivalence of the Schwartz Value System. However, in most cases respondents were convenience samples of teachers and students. The European Social Survey (ESS) from 2002 provides an excellent opportunity to test this equivalence asssumption for a general population and using strict probability sampling. A detailed analysis revieled that the hypothesis of a general equivalence of the Schwartz Value System across all 20 countries could not be supported. This paper will firstly explain the analytical approach (MDS - identical to Schwartz's original publications), then identify the non-equivalent clusters of countries, and finally discusss possible reasons for the non-equivalence. Among them are the reduction of the scale for the purpose of the ESS, general problems with question wording of that scale, and, of course, the underlying theoretical approach.

***Cross-national Measurement of the Religion Effects and Its Application on Public Attitude Toward Science and Technology in European New Member Countries*, Dan Liao, University of Nebraska, Lincoln; ashley_bnu@hotmail.com, David Palmer, Gallup Research Center; dpalmer1@bigred.unl.edu, Lingrui Jiang, UNL-Gallup Research Center; lily_wind2000@hotmail.com, Xiaoming Liu, Gallup Research Center; samliu@bigred.unl.edu, and Allan McCutcheon, Gallup Research Center; Allan_McCutcheon@gallup.com**

Debate between contradicting religious beliefs and scientific theories are not limited to philosophical settings. These debates are often at the heart of policy decisions regarding science and technology research. A nation's religious atmosphere may influence not only the amount of money and effort the country invests in science and technology research, but also the types of research that are pursued. In May 2004, ten candidate countries, Bulgaria, Cyprus, the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Romania, Slovakia, Slovenia, and Turkey, joined the European Union (EU). Because eight of these ten countries were previously under communist rule, their religious atmospheres are different from veteran EU members. These differences need to be understood and anticipated because they may

translate into policy conflicts between veteran and new member countries. The Cross-national measurement for the effects of religion on public opinion was attracted increased attention from social scientists over the past decade. The current research focuses the impact of religious belief on public attitude towards science and technology. The analysis examines the influences of individuals' religious belief and religiosity, as well as these variables' interactions with age and gender on response rates and responses toward science and technology within each new member country of the EU and the comparisons of the effects across countries are next performed. The data are from a series of surveys modeled on the Standard Eurobarometer in the candidate countries of the European Union, conducted by the Gallup Organization-Hungary in September 2002.

***Attitude towards Science and Technology In European New Member Countries: A Cross-National Study*, Dan Liao, University of Nebraska, Lincoln; ashley_bnu@hotmail.com, Lingrui Jiang, UNL-Gallup Research Center; lily_wind2000@hotmail.com, David Palmer, Gallup Research Center; dpalmer1@bigred.unl.edu, Xiaoming Liu, Gallup Research Center; samliu@bigred.unl.edu, and Allan McCutcheon, Gallup Research Center; Allan_McCutcheon@gallup.com**

The public attitude towards science and technology has a prominent impact on the level of governmental support for research, the number of young people devoted to the science and technology career, and the application and improvement of technology. In this study, confirmatory factor analysis models are conducted to measure three dimensions of the public attitude towards science and technology (knowledge of fundamental science and technology facts and concepts, optimism regarding science and technology, and attitude toward the sciences and related areas as professions) and demographic influences on these dimensions. Due to the fact that a considerable amount of data is not missing at random, full information maximum likelihood estimators with missing data were utilized. . As an efficient technique to test the cross-national construct comparability of the three dimensions across the 10 countries, multiple-group confirmatory factor analyses are performed. This implies testing the equality of factor loadings and intercepts across groups. This equality is required to ensure that factors have the same interpretation in all groups, which is necessary when comparing any aspect of the factor distribution across groups. Finally, comparable groups of countries are compared through three dimensions and the effects of the differences on several economic indices of new member countries are discussed. The data are from a series of surveys modeled on the Standard Eurobarometer in the candidate countries of the European Union, conducted by the Gallup Organization-Hungary in September 2002.

MEDIA RESEARCH

***Using Public Opinion Polling to Evaluate Mass Media Health Promotion Campaigns: The Good, The Bad and The Not-so-Good*, Corinne Hodgson, Corinne S. Hodgson & Associates Inc.; corinne@cshodgson.com**

Beginning in 1999, the Heart and Stroke Foundation of Ontario has frequently used public opinion polling technology to evaluate the impact and reach of health promotion campaigns utilizing mass media (television and print). Three evaluation studies conducted for the Foundation will be discussed and compared: (1) The Good -- a controlled study of the impact of advertising on public knowledge of the warning signs of stroke; (2) The Bad -- a pre- and post-test of knowledge, attitudes and behaviors concerning high blood pressure; and (3) The Not-So-Good -- longitudinal tracking of attitudes about smoking. The ability of public opinion polling to effectively evaluate the impact of a health promotion program is in large part determined by the nature of the campaign. Campaigns that focus upon knowledge (the stroke awareness campaign) and for which comparable control groups can be constructed are relatively easy to implement and interpret. Campaigns that focus upon attitudes (the high blood pressure and smoking campaigns) or for which control groups can be difficult to establish can be much more complicated to analyze and interpret. Specific examples will show the sorts of problems that can arise in this type of research. It is important that polling reflect not what the health promoter would like to see as the end result of his/her campaign (often, a change in behavior) but rather the true focus of the advertising. Sometimes, the pollster must resist the desire of health promoters to "stretch" the scope of questioning. Polling can be an effective evaluation tool but only when the research question is focused and measurable and health promoters understand both its strengths and limitations.

***Who learns from the news? The moderating role of cognitive processing styles in learning from the news*, Clarissa David, University of Pennsylvania; cdavid@asc.upenn.edu**

This study investigates the role of general cognitive styles, namely need for cognition and need to evaluate, in predicting news media use habits and moderating learning effects from news exposure. Research is conducted on data from two cross-sectional Internet-based probability sample surveys. Preliminary results indicate that cognitive styles do not have strong main effects on political knowledge after controlling for demographic and motivational variables. However, supporting evidence is found for the hypothesis that those who are high on need for cognition and high on certain types of news media use gain more political information at comparable levels of exposure than those who are low on need for cognition. Moreover, tests also suggest a negative interaction between exposure to network television news and need to evaluate such that evaluators tend to gain less political information from equal levels of network TV use. While results show statistically significant interaction effects, little contribution is made by way of increasing variance explained on political knowledge. Findings suggest that general cognitive processing styles may influence depth or type of processing, and consequently the amount of learning, when exposed to news. Implications on information processing approaches to understanding how people learn from the news and the potentials of utilizing survey-based measures of information processing styles are discussed.

***News discrepancy Perception and News Credibility judgment: The Role of the Self as a Comparison Anchor in Judgmental process of News Credibility*, Hyunseo Hwang, University of Wisconsin-Madison; hyunseohwang@wisc.edu, Gun Hyuk Lee, Department of Communication, Changwon National Uni; ghlee21@hotmail.com, and Sung Gwan Park, Department of Communication, Seoul National Univer; park@snu.ac.kr**

Research on news credibility and its antecedents consistently shows perception or judgment of news credibility is both subjective and relativistic. Even though this subjectivity and relativity element of news credibility judgment are key theoretical components in research on media bias perception, they have not been conceptually developed and empirically tested in the research. For example, some scholar measures media bias perception by asking respondents to evaluate whether a given article was neutral, or biased in favor of one side or the other, while others measure news bias perception as asking respondents to evaluate news slant compared to their own views. In addition, different but similar concepts such bias, slant, trust, and credibility have been used without any clear conceptual distinction. To address these measurement and conceptualization problems in the research on media bias perception, this study distinguished media bias perception into two different components - news discrepancy perception and news credibility judgment and then tested the role of the self as a

comparison anchor in audience judgmental process of news credibility. Specifically, we constructed two different news discrepancy perceptions - news discrepancy from neutral point and news discrepancy from one's own view and examined the relationships of each discrepancy perception to both its antecedents (e.g., issue involvement, political ideology, and strength of political ideology) and news credibility judgment. A Web-based survey with the issue of revision of the Status of Forces Agreement (SOFA) spurred by U.S military vehicle accident in South Korea are used to test several hypothesized relationships among the main variables. Findings reveal that neither antecedent variables of the model nor news credibility were related to perceived news discrepancy from neutral point. In contrast, perceived news discrepancy from one's own views had strong relationship to issue involvement, political ideology, strength of political ideology, as well as news credibility.

***News Discrepancy and Information Search: The Effects of News Slants on Audiences' Information Search Patterns Over the Internet*, Hyunseo Hwang, University of Wisconsin-Madison; hyunseohwang@wisc.edu, Kwangjun Heo, University of Wisconsin-Madison; kheo@wisc.edu, and Sun-Young Lee, University of Wisconsin-Madison; sunyounglee@students.wisc.edu**
This study examined how people in feeling of alienation from mainstream media message consciously engage in information search through the Internet, by introducing the concept of "media dissociation"- the discrepancy between mainstream media portrayals and an individual's own view on an issue. Our analysis focused particularly on individuals' information seeking behavior over the Internet, in doing so we proposed that the more individuals dissociate their views from media portrayals, the more they are engaged in the issue-related information search. Using an experimental design embedded within a Web survey, we exposed Korean college students to mainstream news stories that were manipulated in slants on the deployment of South Korean troops in Iraq. The manipulation focused on whether the mainstream news story was supporting for or opposing to the issue. Following the exposure to the mainstream news stories, accessible 10 online news articles were offered in a Web-style format page that presented hyperlinks with the headlines and leads of the articles. Research participants were instructed to follow their own personal interest in choosing and reading the articles. The headlines and leads of the articles were adjusted to show the tone of the article clearly so that readers recognized the slant of the article easily. The respondents' hyperlinks usage was monitored and recorded into a log file. The results showed that the respondents in dissonant news story condition read the news articles more in both supporting for and opposing to their own views than those in consonant condition. In addition, the respondents in dissonant condition were attracted to news articles that were consistent with their views significantly more than the respondents in consonant condition. The further implication of the study was discussed.

***Does Television Change Us? An Analysis of Three Experiments*, Bethany Albertson, University of Chicago; blalbert@uchicago.edu**
In the last half century, television has become a primary source of news and information for the American public. As such, it has the potential to inform viewers and influence their attitudes and behavior. A growing literature has attempted to study the effects of television, but research has been hampered by methodological challenges. Laboratory experiments face problems of external validity that impede generalizability. Observational studies take advantage of real world conditions, but introduce numerous confounding variables that make it difficult to isolate the effects of television. This paper reports on three field experiments that stake a methodological middle ground, combining the random assignment of experimental work with everyday settings. Using instrumental variable analysis, we analyze the effects of watching three informational broadcasts on viewers' knowledge, attitudes and behavior. We find that the effects of watching this kind of broadcast are more limited than previous studies have suggested.

HOW AAPOR CAN HELP MEET THE NEEDS OF MID-CAREER MEMBERS

Sandra Berry

STANDARDS

***ISO Standards for Market, Opinion, and Social Research: A Preview*, Tom W. Smith, NORC, University of Chicago; smitht@norc.uchicago.edu**
The International Organization on Standardization (ISO) has drafted standards for the survey research industry which will be adopted later this year (2005). This paper reviews the major features of those standards.

***Professional Standards: New Challenges, Changing Ethics?*, Diane Bowers, CASRO; dbowers@casro.org**
As the survey research industry continues its evolution--responding to advances in technology, client demands for more actionable, targeted, even respondent-specific data, and regulatory (and public) roadblocks, researchers face more "grey" than "black" lines on standards questions. Our balancing act between traditional research principles and practices and research (information services) progress is becoming more difficult to maintain. Our arguments and explanations of why . . . why we must protect confidentiality and privacy; why we should be free to contact anyone we want to; why (and how) we're "different" from direct (tele) marketing; etc. are not as compelling or convincing as they used to be. This session will present recent ethical and professional "dilemmas" that challenge CASRO's (and AAPOR's) standards experts . . . and ask you what you would do.

***Reporting Standards for Internet Surveys and Polls*, Richard Tychansky, University of Toronto; richardt@eratos.erin.utoronto.ca**
Internet surveys and polls today lack accountability. Researchers need guidance in making the results of their data collection efforts scientifically and sociably acceptable. Although, the Internet has provided researchers and pollsters alike with a means of conducting surveys and polls at a minimal cost, which is diametrically opposed to methods such as RDD based telephone interviewing, standards for the reporting on the data collection process are non-existent. Presented is a standard set of measurable quantities, which can be used in evaluation and comparison of Internet based methodological approaches all on one common ground. Three case studies are evaluated in terms of their response rates, contact rates and sampling strategies, they are: a national survey of legal scholars and professionals, a national survey of university students and a long running political Internet poll. Each study contributes to the development of reporting standards for Internet based surveys and polls. The measurable quantities derived from each study allow researchers, funding agencies and the public alike to audit any claim of representativeness of a population, or the accuracy and precision of the response data. Full disclosure is the key

to understanding the success or failure of a study, as is a requirement for Internet based researchers to ensure the privacy and confidentiality of data stored and transmitted from their systems. A standardized reporting format is presented for researchers to use in calculating and reporting contact rates, cooperation rates, and response rates for Internet surveys and polls. These three standardized quantities provide researchers and peer reviewers alike with the necessary metadata that is associated with a study in order to effectively measure, report and judge the Internet data collection process.

WAR IN IRAQ

***Support for the War in Iraq: American Casualties as a Survey Item*, Mark West, U. North Carolina at Asheville; west@unca.edu, and Donald Diefenbach, U. North Carolina at Asheville; ddiefenbach@unca.edu**

Support for the war in Iraq has proven difficult to quantify, with respondents often reporting substantial ambivalence if not confusion concerning their opinions re the further prosecution of the war. In an effort to understand the characteristics of those who support further prosecution of the war, and to understand better the nature of their support, a regional random-digit dialed telephone survey applying a willingness-to-pay question to Iraqi war costs was conducted in the first two weeks of October in the Western counties of North Carolina. The survey used a next-birthday method for within-household respondent selection. A total of 493 complete interviews were conducted, with an AAPOR cooperation rate 1 of .48 and a contact rate 2 of .43. Seventy-eight percent of respondents said that they would support no more deaths of U.S. soldiers in Iraq, even after the explanatory introduction to the survey question. Fifteen percent of respondents said they would support any number of deaths in order to fulfill the U.S. mission in Iraq, while 7 % said they would support some intermediate number of deaths, with a modal value of about 100 deaths. A nonlinear regression model was constructed using respondent characteristics to predict the number of U.S. deaths the respondents would find tolerable. Age, income, education, gender, race, newspaper readership, television viewership, and party affiliation were all found to be significant predictors of the number of deaths the respondent would tolerate; the model R2 was .21, and the model was significant at the .01 level. Respondents who were older, watched less television, who were male, had higher income and educational levels, and were white were significantly more likely to support higher levels of U.S. casualties than were other respondents.

***Exemplification, War and Iraq: Examining the Effects of Exposure to Violence in Iraq on Attitudes on the Iraq War*, Russ Tisinger, Annenberg School for Communication; rtisinger@asc.upenn.edu, and Kelli Lammie, Annenberg School for Communication; klammie@asc.upenn.edu**

Research has consistently shown that the news media's tendency to use exemplars in relaying news stories has a demonstrable effect on audiences' perceptions of incidence rates. Viewers ignore the more reliable base-rate information, which reports actual rates of occurrence, and concentrate on the more vivid, detailed and concrete pictures and information presented in exemplars. When considering occurrence rates, viewers' judgments often reflect the information presented in exemplars rather than the base-rate information. Although exemplification has been shown to occur within experimental settings, it has not been as thoroughly explored using observational data. Furthermore, the implications of exemplification for viewers' policy positions are largely unknown. This paper addresses these issues by examining the effects of the murder of four civilian contractors in Iraq. In April 2004, when United States civilian contractors were violently killed in Fallujah, Iraq, news media organizations reported graphic descriptions of the killings and published pictures of the event. In this paper, secondary data analysis is performed on a Pew Center for People and the Press survey in which respondents were asked if they saw the pictures, their perceptions about the war, and their support for various policy issues regarding the war. The study suggests that not only does the news media's tendency to rely heavily on exemplars have a large effect on audiences' perceptions of incidence rates but it also has implications for the way viewers' express their policy positions.

***Foreign Policy and America's Place in the World:Generational Divisions on U.S. Priorities and Iraq*, Nicole Speulda, Pew Research Center; speuldan@people-press.org**

For the first time since the Vietnam era, foreign affairs and national security issues loomed largest in the presidential election. For most of the 1990s, fewer than 10% of Americans rated foreign policy as the most important problem facing the nation. In 2004, 41% cited defense, terrorism, or foreign policy as the most important national problem, compared with 26% who mentioned economic issues. This paper examines findings from a foreign policy survey by the Pew Research Center in July of 2004. Missing from the analysis of foreign affairs is substantial insight into the generational differences when it comes to key facets of foreign policy issues. As in the Vietnam era, polls consistently show young people (under 30 years of age) more supportive of the war in Iraq than older people. While a 58% majority of Americans across most demographic groups consistently rate as a top priority the goal of getting other nations to assume more of the costs of keeping world order, the goal of persuading other nations to share international burdens is much more important to older Americans than it is to younger people. This paper highlights generational opinion differences in terms of foreign policy priorities and the way older and younger people view America's place in the world. These questions remain: Why are older Americans less supportive of the war in Iraq (as in Vietnam) than those younger, both before it started and since? Are older Americans more or less likely to stick to their partisan identification on this issue? This paper also offers a regression analysis to predict war support or opposition drawing from a series of Pew polls from the last two years. Are the coefficients that shape opinion different among older people than they are among younger people?

SATISFICING

***Effect of respondent motivation and task difficulty on nondifferentiation in ratings: A test of satisficing theory predictions*, Sowmya Anand, Ohio State University; anand.13@osu.edu, Jon Krosnick, Stanford University; krosnick@stanford.edu, Kenneth Mulligan, University of Georgia; kmulliga@uga.edu, Wendy Smith, Research by Design; wsmith@researchbydesign.com, Melanie Green, University of Pennsylvania; mcgreen@psych.upenn.edu, and George Bizer, Eastern Illinois University; cfgyb@eiu.edu**

Survey respondents are routinely asked to answer batteries of items employing a single rating scale. Satisficing theory raises the possibility that respondents might shortcut the cognitive processes required to provide optimal responses and respond identically or nearly identically to all items in the battery so as to provide answers that appear to be reasonable while investing minimal effort, a behavior termed "nondifferentiation." Nondifferentiation is posited to be most likely to occur among respondents who have limited cognitive skills and who have minimal motivation to answer carefully, and when the response task is cognitively demanding. Some existing evidence suggests that

respondents lower in cognitive ability are most likely to nondifferentiate, but little prior work has explored the effects of respondent motivation and task difficulty on nondifferentiation. We sought to test thoroughly whether satisficing implicated factors (low cognitive ability, low levels of respondent motivation, and high task difficulty) in fact increase nondifferentiation. To do so, we analyzed data from several sources, including telephone surveys conducted for the National Election Studies and by the Center for Survey Research at The Ohio State University, and Internet surveys conducted by Harris Interactive, Knowledge Networks, Greenfield Online, GoZing, and similar companies. This research is unique in that it also tested the mediational hypotheses posited by satisficing theory. Specifically, we tested whether cognitive ability, need for cognition, and need to evaluate affect respondent effort and the difficulty of the response task, which in turn affect the extent of nondifferentiation. Separate structural equation models estimated using each dataset provided converging evidence supporting these predictions. Thus, it seems that nondifferentiation in particular, and perhaps other satisficing-related compromising response strategies, may be reduced by increasing participant motivation and decreasing task difficulty.

***Comparing Major Survey Firms in Terms of Survey Satisficing: Telephone and Internet Data Collection*, Jon Krosnick, Stanford University; krosnick@stanford.edu, Norman Nie, Stanford University; nhnie@stanford.edu, and Douglas Rivers, Stanford University and Polimetrix, Inc.; doug@polimetrix.com**

At AAPOR some years ago, Chang and Krosnick reported a national survey experiment in which the same questionnaire was administered via RDD telephone interviewing and by Internet survey firms. Those investigators found that the Internet respondents provided self-reports containing less random and systematic measurement error than did the telephone respondents. We have conducted a new experiment in which a single questionnaire was administered by eight different, major national survey firms to national samples. One survey was done by telephone by SRBI, and the others were done by Internet survey firms. The questionnaire included an array of embedded experiments designed to assess the extent of survey satisficing, including measures of non-differentiation, response order effects, don't know filter effects, acquiescence response bias, and more. Comparisons across the firms will be reported showing that the extent of survey satisficing varied depending on the mode of data collection. Additional analyses will show that these effects of mode and firm are especially prominent among subsets of the populations likely to be most sensitive to such data collection factors.

***Effect of Interview Pace and Items Position on Satisficing*, Michael Lemay, Joint Program in Survey Methodology; mlemay@survey.umd.edu, Scott Fricker, Joint Program in Survey Methodology; sfricker@survey.umd.edu, Mirta Galesic, Joint Program in Survey Methodology; mgalesic@survey.umd.edu, Roger Tourangeau, Joint Program in Survey Methodology; rtourangeau@survey.umd.edu, and Ting Yan, Joint Program in Survey Methodology; tyan@survey.umd.edu**

Satisficing (Krosnick & Alwin, 1987) is one of the models explaining measurement error due to respondents in surveys. While a few studies have been conducted on the circumstances and characteristics that are conducive to satisficing, none of them have focused on the interview pace or the position of the items within the questionnaire. We varied the position (beginning vs. end of questionnaire) of a series of items prone to response effects and the pace of the interview (fast pace, fast pace with pauses between response options, and slow pace) in a 2x3 design. Optimum control of the reading speed will be assured by the use of a text-to-speech synthesizer. ACASI was used to administer the questionnaire in a lab setting. The hypotheses were that satisficing will increase with speed and for items at the end of the questionnaire. Respondents should be more likely to show primacy effect, nondifferentiation and acquiescence in the faster speed condition. They should be more prone to these same effects for the items toward the end of the questionnaire. Several covariates will also be included such as working memory tests and personality items to examine the characteristics of persons prone to response order effects.

GENES AND BIOTECH

***Knowledge, Communication and Opinion Formation: Comparing Attitudes Regarding Genetically Modified Foods*, Gregg Murray, University of Texas at Brownsville; gmurray@utb.edu, Susan Banducci, Texas Tech University; susan.banducci@ttu.edu, and Jeff Karp, Texas Tech University; j.karp@ttu.edu**

The contentious debate about genetically modified (GM) food is often framed in terms of perceptions of their benefits versus risks. In addition to perceptions of risks and benefits, a number of other factors have been found to be associated with support for biotechnology, particularly issue engagement and confidence in the pertinent actors. Both the salience of how the issue is framed and these factors imply a significant role for issue knowledge in attitudes toward GM products. In order to examine these questions, we rely on Eurobarometer data collected in 2002 that surveyed Europeans about their attitudes toward biotechnology. These survey results suggest that Europeans generally lack support for GM food, though there is substantial variation in support for these products within the European Union. Support for GM food is not simply a matter of scientific knowledge but rather is influenced by a complex relationship between instrumental concerns about the relative costs and benefits, ideology and values and trust in groups associated with biotechnology issues.

***Biotechnology and Trusts: What Drives Citizens to Support Biotechnology?*, Zuoming Wang, Cornell University; zw34@cornell.edu, Dietram Scheufele, University of Wisconsin, Madison; scheufele@wisc.edu, and Dominique Brossard, University of Wisconsin-Madison; dbrossard@wisc.edu**

The introduction of biotechnology has generated varying public responses from the public. Researchers found the public support of biotechnology is related to the knowledge about science and the trust toward relevant governing bodies. The extant literature tends to view different types of trust as a whole and does not specify how the different types of trust influence the public opinion on biotechnology. Moreover, trust in government and trust in science itself are relatively less explored than the institutional trust. This paper explores the relationship among ideology, media science news use, and three different types of trust (trust in government, trust in science, and trust in regulative institute), as well as their direct and indirect impact on the support of biotechnology. It answers two key questions. First, what is the role of the media science news use in promoting public support for biotechnology? Second, which type of trust is the most important predictor for the public support for biotechnology? Data for this study came from a national telephone survey (N=781, RDD sampling, response rate: 55 percent). Initial path analysis results showed public are slightly leaning toward the support of biotechnology. Media science news use has both a direct and an indirect positive effect (through trust in science) on support of biotechnology. Trust in regulative institute only has an indirect positive impact (via trust in science) on the support of biotechnology, while trust on government is negatively related to the support of biotechnology. Among different types of trust, the trust in science is the most important predictor for the public support of biotechnology.

***How Clean is the Air You Breathe?: Predicting Beliefs about Regional Air Quality*, Donald Diefenbach, U. North Carolina at Asheville; ddiefenbach@unca.edu, and Mark West, U. North Carolina at Asheville; west@unca.edu**
Western North Carolina is home to the Great Smokies and extensive mountain forestland. The region is also in the pathway of industrial air pollution from the Tennessee Valley and the Ohio River Valley. The present research explores variables that predict regional residents' beliefs about the quality of their air. Do respondents believe the manifest visual impression, which is a clean and undeveloped landscape, or are they aware of the realities of significant air quality problems in the community? 439 respondents were surveyed and asked to provide an estimation of the number of air quality days rated as code yellow or worse in the last year. This open-ended count estimation produced a skewed distribution, as estimates are bounded by zero but unbounded above. A form of non-linear regression modeling known as Poisson Regression was employed, which accounts for this non-normal count estimation data. All variables included in the analysis were demonstrated to be significant predictors of beliefs about regional air quality, including basic demographic variables and both television and newspaper exposure. The media exposure variables allow modeling of beliefs about regional air quality using information processing and cultivation theories. This paper discusses non-linear modeling and its broader application as a valuable tool for enhancing survey research methodology.

TERRORISM

***Psychological and Political Responses to the Madrid Terrorist Attacks*, Kenneth Rasinski, NORC, at the University of Chicago; rasinski-ken@norc.org, Tom W. Smith, NORC, University of Chicago; smitht@norc.uchicago.edu, and Juan Diez-Nicolas, Analisis Sociologicos Economicos Y Politicos; Juan Diez-Nicolas [100613.2721@compuserve.com]**
The terrorist attacks in Madrid, just prior to the national elections, seemed to have turned public sentiment against the incumbents to give a surprise victory to a party that supported withdrawal of troops from Iraq. NORC received a grant from the National Science Foundation to study public responses to the attack. Spanish data were collected by Analisis Sociologicos Economicos Y Politicos (ASEP). Key findings are summarized below. Comparisons to the NORC post 9/11 survey will be made. Exposure. Typically, respondents heard about the attacks through media reports. A higher percentage of residents of Madrid compared to the rest of Spain said that they heard the news through phone calls or personal message. Most respondents, especially those in Madrid, were not able to carry on their activities as usual after the attacks. Behavior. We asked the Spanish sample whether they gave money, clothing or other items to charity, donated blood, did extra volunteer work, or avoided going to government buildings or the downtown of a large city. While the rate of activities was generally low, proportionally more residents in Madrid compared to those outside of the city reported engaging in these behaviors. Emotional Response. A higher percentage of Madrid residents compared to those in the rest of Spain said that they were more upset than others. The two predominant emotional responses were crying and feeling nervous and tense. More emotional responses were found in the Madrid sample compared to the national sample. Political and Personal Response. A higher percentage of residents of Madrid than those outside the city worried about how the attacks might affect the political situation in their country. Similarly, a higher percentage of those in Madrid expressed anger about the attacks, wondered about their own safety and expressed a great deal of confusion.

***Willing to Believe: Explaining the Belief that Saddam Hussein aided with the September 11th Attacks*, Russ Tisinger, Annenberg School for Communication; rtisinger@asc.upenn.edu**
Before, during, and after the major combat operations in Iraq of Gulf War II, public opinion polls showed that a large proportion of Americans - a majority in some cases - believed that Saddam Hussein was involved in the September 11th attacks. This paper uses June 2004 data collected by the National Annenberg Election Survey (NAES) to test how well John Zaller's model of persuasion predicts this belief. Zaller's model holds that the probability of holding a certain belief depends largely upon the probability of being exposed to a certain message and the acceptance of the message. Preliminary analyses suggest that the model does indeed largely explain belief that Saddam Hussein aided the September 11th terrorists. An already-completed analysis of NAES data gathered in June 2003 provided evidence that 1) approval for President Bush was highly correlated with the likelihood of believing in Saddam Hussein's involvement in the September 11th attacks and 2) frequency of newspaper-readership interacted significantly with approval for President Bush to predict likelihood of believing in Saddam's involvement. In other words, the more Bush detractors read the newspaper, the less likely they were to believe in Saddam's involvement. One year later, NAES asked again about Saddam's involvement. This paper will use both the June 2003 data and the June 2004 data to engage in a more precise test of Zaller's model with regard to beliefs about Saddam's involvement in the September 11th attacks.

***Public Prudence, the Policy Salience of Terrorism and Presidential Approval following Terrorist Incidents*, William Josiger, Georgetown University; josigerw@georgetown.edu, and George Shambaugh, Georgetown University; shambaug@georgetown.edu**
The public is generally assumed to "rally around the flag" in the event of national emergencies, often granting the Executive expanded support to respond as needed to promote national security. Since September 11, 2001 the Executive Branch has presented terrorism to the public as a matter vital to national security. It has linked much of its policy agenda to the persistent threat of terrorism and, more recently, it focused much of its reelection campaign strategy on the same. Public support for the president and U.S. policy in the aftermath of terrorist events has been high, but it is not as uniform as these reelection strategies might suggest. Our analyses indicate that the importance that the public assigns to terrorism as a policy issue and the rallying effect that terrorist incidents have on public support for the president and his policies vary as a function of the costs of the incidents in human and material terms, the types of people targeted and the weapons used by the terrorists. Our findings indicate that the public is cost and risk averse and it tends to assess risk in personal rather than nationalist terms. This variation lends support to the work of Bruce Jentleson and others suggesting that, even in the context of a terrorist incident, the public exercises some prudence when offering its policy support to the president.

***What is terrorism? Challenges to Improving the Construct Validity of a Complex Concept*, Stephanie Willson, National Center for Health Statistics; zex8@cdc.gov, and Kristen Miller, National Center for Health Statistics; ksmiller@cdc.gov**
With the events of September 11, 2001, public perception regarding the threat of terrorism has shaped national priorities and, arguably, has altered the nation's political course. Undoubtedly, monitoring the nation's attitudes about terrorism and the threat of another terrorist attack will be a primary goal for opinion poll researchers for many years to come. However, although salient in the mind of the American public, the concept of terrorism is also elusive, containing multiple dimensions of meaning that shift depending upon specific events, as well as social and political contexts. As a result, asking about terrorism is difficult. Survey questions on the topic are likely to prove vague and be

subject to many interpretations, thus weakening construct validity. Based on 25 in-depth interviews conducted in 2004 by the National Center for Health Statistics, this paper will present the various definitional components of terrorism, and illustrate how this interpretive complexity lends difficulty to survey research on terrorism by increasing the potential for response error. In this discussion, the paper will also illustrate how the meaning of terrorism has not materialized as a unified, collective definition-one emerging from the 9/11 attacks against the US-but rather, as a series of interpretations dependent upon the reality of individuals' lives and the context of their communities. For example, analyses reveal several dimensions to the concept of terrorism. One dimension relates to perpetrators of violence (organized group vs. loner), while the other focuses on the targets (random individuals vs. specific people). When survey questions failed to define either of these groups, response error occurred as respondents drew upon their own experience and decided for themselves what was meant by the term "terrorism." Other dimensions of the concept will be discussed, as will suggestions for improving question design and construct validity

ORDER AND CONTEXT

***Response Order Effects in Online Surveys*, Randall Thomas, Harris Interactive; rthomas@harrisinteractive.com, Susan Behnke, Harris Interactive; sbehnke@harrisinteractive.com, and Alyssa Johnson, Harris Interactive; amjohnson@harrisinteractive.com**
Response order effects have been found in phone surveys and mail surveys (Krosnick, 1999). Through a series of online studies, we examined vertical response order effects by randomly assigning respondents to be presented with attitudinal scales either showing the positive end of the scale at the top or at the bottom. We found a strong response order effects for most attitude scales and will detail the implications for inferences derived from such data and the potential impact on validity of measurement.

***Causes of Context Effects: How Questionnaire Layout Induces Measurement Error*, Andy Peytchev, University of Michigan; andrey@isr.umich.edu, and Roger Tourangeau, Joint Program in Survey Methodology; rtourangeau@survey.umd.edu**
Placing related questions together can alter the associations between answers to them. We conducted two experiments that varied whether related questions were presented together on a single screen in a web survey. The first experiment replicated that inter-item associations and scale reliability were highest when the questions were presented together. However, a structural equation model revealed that these higher associations reflected correlated measurement error and decreased rather than increased construct validity. We carried out a second experiment to test three possible mechanisms for the heightened correlations, but reduced validity. First, the questions may be perceived as being multiple measures of the same construct, inducing more similar interpretations of the items. Secondly, when no actions are needed to get to the next question, the same material may be retrieved from working memory in answering all the questions. Thirdly, respondents may be minimizing effort by clicking response options in the same columns and paying less attention to the individual questions when they are presented in a grid. Our second experiment used a factorial design in a web survey with 2,694 respondents. Respondents answered 4 questions on diet and 4 on exercise, where the layout (together in a grid, together on a screen but listed separately, in separate screens), the accompanying instructions (related, independent, no instructions), and the order of the questions (by topic, intermixed) were varied randomly. We also expected these manipulations to interact with the location of the experiment in the questionnaire and randomly assigned its placement. Respondents' Body Mass Index was calculated in order to estimate and compare measurement error properties and validity of the diet and exercise constructs. The findings will allow us to understand the mechanisms generating differences in responses to questions on the same topic and guide survey design decisions that affect measurement error, nonresponse, and cost.

***Context, Race, Gender and Policy Considerations*, David Wilson, Gallup/Michigan State University; david_wilson@gallup.com**
To what extent do beneficiaries of public policies and programs gain from being mentioned in the comparative context in public opinion surveys? For instance, previous research on Affirmative Action (AA) opinions has shown that attitudes can vary across the frame, or context, by which the policy is presented. In addition, research has also shown that the public supports AA programs for women over racial minorities. These differences in response across content and context have led to muddled conclusions about the consistency of AA policy attitudes. Using data collected by the Gallup Organization in 2003, we reexamine these questions using a split-ballot design experiment, randomly ordering questions assessing perceptions of "equal employment opportunities" (EEO) and support for "affirmative action" for women and racial minorities. The results show that racial minorities have increased sympathy and advocacy -- lower perceptions of EEO and higher support for AA -- when they are considered in a comparative context to women. Accordingly, women have the highest levels of EEO sympathy and AA support, however this advocacy declines when gender is mentioned in a comparative context to race. We further consider how gender and race interplay with these contextual effects. The results of the analysis have implications for questionnaire design, attitudinal measurement, and the interpretation of trends in EEO and AA.

DETROIT AREA SURVEY

***Detroit Area Study*, Elizabeth Martin, US Census Bureau; emartin@census.gov, Diane Colasanto, Consultant and Mick Couper, University of Michigan; mcouper@umich.edu**

BLOGS AND BLOGGERS 2004: EXTENDING THE REACH, INFLUENCE, UNDERSTANDING AND MISPERCEPTIONS OF PRE-ELECTION AND EXIT POLLS

***How did Blogs and the Internet Influence Polling in the 2004 Election?*, Nancy Belden, Belden Russonello & Stewart; nancybelden@brspoll.com, Mark Blumenthal, Mystery Pollster, Steve Soto, The Left Coaster, Joel Bloom, Oregon Survey Research Lab, Tom Bevan, RealClear Politics, Tom McIntyre, RealClear Politics**
The advent of blogs in recent years has changed not only the tome frame for the dissemination of polling data during an election cycle but has also greatly increased the amount of information about pre-election and exit polling to the interested public. During the 2004 elections, millions of Americans downloaded detailed pre-election polling data as well as incomplete exit poll data by accessing new web sites. Political "blogs" (short for web logs) were alive with debate and commentary on the meaning, accuracy and methodology of political polling during the pre-election season. What was the impact on campaign coverage? What are the implications of blogs and the dissemination of polling data via the web for the practice of survey research?

IDEOLOGY AND ISSUES

***Is ideological self-identification in Internet speech silenced by opinion opposing the speaker?*, David Fan, Univeristy of Minnesota; dfan@cbs.umn.edu**

There is a substantial literature on the impact of persuasive information on public opinion. For example, studies have shown that news messages can push opinion in directions presented in news stories. In contrast, there has been much less study of the impact of opinion on public self-expression. Noelle-Neumann's Spiral of Silence suggests that unfavorable opinion depresses public utterances. A resulting hypothesis is that speech with self-identifications as liberal or conservative should diminish to the extent that opinion favors the other side. The advent of the Internet permits this hypothesis to be tested directly because public speech can be captured without filtering by gatekeepers. In particular, Usenet bulletin board postings can be retrieved from the Google Groups archive back to the 1980s so long time trends can be constructed of public self-expression. The hypothesis will be supported if Usenet postings by self-identified liberals is smaller when polled opinion favors conservatives and vice versa. Data for the public opinion predictor will be retrieved from the POLL database of the Roper Center. This database has over 1,729 poll questions asking whether a person is politically liberal or conservative. Therefore, it is possible to construct a finely grained time trend of the predictor. Furthermore, a search of Google Groups using the search command (("I am" OR "I'm") conservative liberal) yielded approximately 365,000 postings. Therefore, it is possible to construct a detailed time trend of the dependent variable of public utterances including overt self-identifications as being conservative or liberal. The content analysis will be performed using the InfoTrend computer method used in previous studies for examining the impact of the press on public opinion, and opinion on conservatism and liberalism will be used in the ideodynamic model to predict the time trend of the frequency of ideological self-identifications.

***Ideology, Issue Preferences, and Political Choice: The Paradox of "Conflicted Conservatives"*, Christopher Ellis, UNC-Chapel Hill; crellis@email.unc.edu**

Macro-level studies of public opinion have noticed a seeming paradox in American public opinion: the American electorate is operationally liberal, but symbolically and ideologically conservative. The mass public, in other words, holds liberal preferences on a wide range of spending, social, and cultural issues, but identifies primarily as politically "conservative." At the micro level, this "paradox" means that a large segment of the population possesses conflicted political preferences, identifying as ideologically conservative, but holding predominantly liberal positions on issues. These "conflicted conservatives," who comprise around one-fifth of the American public and identify in roughly equal numbers with the Democratic and Republican parties, present an interesting and important puzzle for scholars of public opinion. This paper builds on the ideological self-identification and elite framing literatures to develop a micro-level theory for the preponderance of "conflicted conservatives." In both political and non-political contexts, the word "conservative" is both more multidimensional and more popular than the word "liberal." Because of the way arguments are framed by party elites, religious and cultural leaders, and the mass media, the dominant messages that citizens receive from political elites are themselves conflicted in a way that leads certain segments of the mass public to identify as "conservative" despite holding predominantly liberal issue preferences. The idea of "conflicted conservatives" suggests that the meanings of ideological terms-and the use of ideology in structuring mass decision making-are more multifaceted than is often assumed. The impact of ideology on political choices is not simply conditioned by education and sophistication; rather, ideological terms mean substantively different things to different groups of people. These findings have implications for the study of both mass opinion and American electoral dynamics.

***The Future of the First Amendment: A National Survey of High School Students*, Kenneth Dautrich, University of Connecticut; k.dautrich@uconn.edu, John Bare, Blank Family Foundation; JBare@ambfo.com, and David Yalof, Univ. of Connecticut; david.yalof@uconn.edu**

In recent years, many First Amendment liberties have come under pressure. The University of Connecticut has conducted annual national surveys on the state of the First Amendment every year since 1997. Yalof and Dautrich use these surveys to document how the general public views free expression in "The First Amendment and the Media in the Court of Public Opinion" (Cambridge University Press, 2002). The pressures of questionable journalism techniques, overzealous reporters, and events such as September 11 and the Patriot Act have strained free expression rights. The Knight Foundation commissioned the University of Connecticut to conduct a national survey of high school students in 2004, in part for the purpose of comparing student opinions about the First Amendment to the attitudes of adults. The report on the 'Future of the First Amendment" will be issued in February 2005. In this AAPOR paper, the principal investigators for the Knight-funded survey explore specifically student orientations toward the First Amendment and compare such opinions to their adult counterparts. The sampling design and data collection methodology also enhance our ability to examine possible sources of how opinions form regarding First Amendment rights. Not only is the national sample of students quite large (over 100,000), but school building data from the 350 schools from which these students were sampled is utilized in the analysis. In addition, teachers and principals from these 350 schools were also surveyed, enabling us to trace the effects of educational actors and circumstances on opinion formation.

SURVEYS OF CHILDREN AND YOUTH

***Children and Juveniles as Respondents - Results from Field Experiments on Measurement Error*, Marek Fuchs, University of Kassel; marek.fuchs@uni-kassel.de**

The theory of the question answer process as well as rules for designing good questionnaires are well developed for surveys of the adult population. Based of a wide array of experiments on question order effects, response order effects and on other aspect of the questionnaire design a detailed understanding of the cognitive processes applied by respondents when answering a survey question has been developed. In recent years, however, many surveys have addressed samples of children and juveniles. In the early 90th Jaqueline Scott and Edith de Leeuw have pointed out that standardized surveys of these populations should take into account a wider understanding of the respondents' cognitive and social capacities. So far, little empirical work has been done on the question answer process of underaged respondents. This paper aims to present results from several field experiments on whether or not the well known findings for the adult population can be reproduced for children and juveniles. Furthermore, we will analyze the effects of age and education as indicators for cognitive capacity on the size and the direction of the response effects. Thus we will contribute to a more general discussion whether or not the general assumptions regarding the cognitive processes in the respondents mind do also apply to children and juveniles. We will present results from

a large scale self-administrated survey conducted in 2004 with children and juveniles aged 10 to 21 (n=4,500). Several experiments on question order, response order and scale effects were incorporated in each survey. Respondents where randomly assigned to several split groups in order to produce response effects. The results indicate considerable differences among respondents of different ages and educational levels.

***Evaluating Sources of Variation In Survey Estimates of Children with Disabilities Using Federal Education Data*, Stacey Bielick, ESSI/AIR; sbielick@air.org, Jennifer Park, National Center for Education Statistics; jennifer.park@ed.gov, and Elvira Hausken, National Center for Education Statistics; elvira.hausken@ed.gov**

Estimates of children with disabilities can be difficult to measure in national surveys. For example, the estimated number of children with disabilities as reported by parents from the National Household Education Surveys Program (NHES) has been consistently higher in each data collection than data collected by the Office of Special Education Programs (OSEP) through state administrative records. However, reliable estimates of school enrollment, receipt of special education and related services, and academic performance of children with disabilities are critical for effective school budget, staffing, and program planning at national, state and local levels. The goal of this research is to determine if different measures of children's disabilities-specifically, parent-reported general and detailed questions about disability status, therapy and services-yield different estimates and to measure the manner the aggregate estimates diverge from counts from extant administrative records. In the paper, we compare survey estimates of children with disabilities in kindergarten, first, and third grade as measured by parent reports in two federally sponsored surveys, the Early Childhood Longitudinal Survey Kindergarten Class of 1998-99 (ECLS-K) and the NHES of 1999, 2001, and 2003-a cross-sectional national RDD survey of children and adults. Sources of variation in estimation between these surveys are examined by drawing from multiple measures of disability in the ECLS-K and NHES. NHES data are also compared with age-specific counts of children with disabilities as compiled by OSEP. We use standard errors adjusted for complex survey data and Student's t tests to make comparisons of estimates. This paper furthers understanding of survey measurement issues in the estimation of special populations. The analysis broadens our understanding of the ways in which survey measurement can affect population estimates, and offers a method of improving data collection and analysis by drawing from relationships observed in comparable datasets.

***Factors Contributing to Assessment Burden in Preschoolers*, Cassandra Rowand, Mathematica Policy Research; crowand@mathematica-mpr.com, Susan Sprachman, Mathematica Policy Research; ssprachman@mathematica-mpr.com, Holly Rhodes, Reserach Triangle Institute; hrhodes@rti.org, and Ina Wallace, Research Triangle Institute; wallace@rti.org**

The emphasis placed on school readiness by the No Child Left Behind Act has brought child assessment research into preschool classrooms. Assessments are a regular tool for examining topics as diverse as the impact of maternal employment on child wellbeing and the impact of implementing preschool curricula. Assessment batteries administered to four-year-olds can range in length from the 15-minute Head Start National Reporting Survey to the 70-minute Preschool Curriculum Evaluation Research (PCER) battery. This presentation examines the extent to which assessment length and assessor experience impacts children's performance. This topic is particularly relevant because high quality surveys collecting child assessment data must strike a balance between the amount of data gathered and the acceptable burden placed on young children. This topic also impacts the development of trainings for those administering assessment batteries to children. Two contractors for the US Department of Education's Institute of Education Sciences administered lengthy PCER assessments to four-year olds in 313 classrooms across 20 sites. Data were collected in the fall and spring of the pre-kindergarten year. We examined both waves of data with respect to assessment length, whether the assessment was completed in one or two sittings, assessor experience, child behavior, and how these factors related to assessment score. The resulting data address five research questions: 1) Do children completing the assessment battery in less than one hour perform better than children who take more than 70 minutes? 2) Do children who are administered the assessment in two sittings perform better than those who complete it in one sitting? 3) Do assessors rate children's behavior as more cooperative if the assessment is done in two sittings rather than one? 4) Do children assessed by experienced assessors perform better than those who have less experience? 5) Do experienced assessors rate children's behaviors differently than inexperienced assessors?

***Evaluating Differences in Health Care Quality Ratings in Children Based on the Level of Proxy Respondent Involvement with Child's Care*, Timothy Beebe, Mayo Clinic; beebe.timothy@mayo.edu, Walter Suarez, Midwest Center for HIPAA Education; walter.suarez@sga.us.com, and Kathleen Thiede Call, University of Minnesota; callx001@umn.edu**

Use of proxies can be a valuable and necessary approach to assess the quality of health care provided to children. Some suggest that the accuracy of parent proxy response is higher for younger than older children. Although this issue has received little empirical attention different standards exist for including children 13-17 in consumer surveys via proxy response. The National Committee for Quality Assurance excludes 13-17 year olds in the CAHPS 3.0H survey sampling protocol for health plan accreditation. This practice is based on the assumption that parents cannot accurately reflect the health care experiences of their older children because they rarely accompany older children to the clinic, and when they do, they may not accompany the child into the examination room. By contrast, the CAHPS Research Consortium and the National CAHPS Benchmarking Database allow parents to respond for their teens up to age 17. This study evaluates differences in proxy responses to selected quality of care items based on how often responding parents of older children (aged 13-17) accompany them to the clinic and/or physical examination room. We use data from a statewide survey (mail with telephone follow-up) of Minnesota's public health care program (e.g., Medicaid) enrollees (AAPOR response rate #4=54%). The results indicate that the level of parental accompaniment to the clinic and/or examination room falls precipitously with age resulting in concomitant differences in health care ratings. Moreover, the findings suggest that health care ratings are more affected by whether the parent accompanies the child into the examination room than just going with them to the clinic. Specifically, we found that parents who almost always accompany children into the examination room gave significantly higher ratings of overall health care, how well the provider listens, and how well the provider explains things than those accompanying the child less often.

INTERNET AND WEB SURVEYS II

***The Market Value Survey: ensuring quality on a government web based survey*, John Flatley, Office for National Statistics, UK; john.flatley@ons.gov.uk, and Dave Ruston, Office for National Statistics; john.flatley@ons.gov.uk**
The English House Condition Survey (EHCS) is a nationally representative continuous survey which is run by the Office for National Statistics (UK) on behalf of the Office for the Deputy Prime Minister (ODPM). The EHCS is a multi-stage survey and the survey unit is the dwelling. Following a face-to-face interview with the current occupants and a physical inspection of the property there is a follow-up desk based survey. This is called the Market Value Survey (MVS) and includes the collection of the dwelling's market value and other supplementary questions relating to housing demand and the local area. The MVS is conducted via a web-based questionnaire completed by professionally trained property valuers. Pivotal to a successful MVS is the extrapolation and display of background data (from the earlier interviewer and building surveyor administered parts of the survey) such as property descriptions and digital photographs. This paper explores how quality was ensured when moving these data between the different parts of the survey. There will also be a focus on the practical considerations of official web data collection such as item non-response, on-line data validation and data confidentiality.

***Using Portal Technology to Improve Quality at the U.S. Census Bureau-Presentation and Demonstration of Portal Technology*, Cheryl Landman, U.S. Census Bureau; cheryl.r.landman@census.gov, and Deborah Stempowski, US Census Bureau; deborah.m.stempowski@census.gov**
The U.S. Census Bureau is the preeminent collector and provider of timely, relevant and quality data about the people and economy of the United States. With its goal "To provide the best mix of timeliness, relevancy, quality and cost for the data we collect and services we provide", it was natural for the Census Bureau to turn to Information Technology to achieve this goal. During 2000, we began experimenting with portal technology to improve the quality of our data and business processes. As part of the Quality program, we needed a vehicle to communicate and share knowledge about quality throughout the organization. We established a portal that would provide users with the quality standards and best practices while also providing a repository for documents that others could use in developing their quality components. At the same time, we saw an opportunity to link quality with another major process improvement activity - the Project Management Initiative. We designed the Quality and the Project Management Portals using Oracle WebDB. We modified these portals over the last two years based on customer feedback. The portals provide a way to manage data quality and share best practices, standards, and guidelines. As we progressed with the use of portals, we began designing portals to help analysts improve their work processes. These two survey portals give analysts the ability to easily review data from disparate sources, compare data from multiple surveys on one screen, annotate analytical reports electronically and share notes across reports, as well integrate search results from multiple engines. The final release of these portals is now available for analysts to use while analyzing their survey data. This presentation will demonstrate portal technology, provide background about how we developed the sites and highlight how we used customer feedback and lessons learned through this process.

***A Comparison of an Online Card Sorting Task to a Rating Task*, Randall Thomas, Harris Interactive; rthomas@harrisinteractive.com, Leonard Bayer, Harris Interactive; lrbayer@harrisinteractive.com, Alyssa Johnson, Harris Interactive; amjohnson@harrisinteractive.com, and Susan Behnke, Harris Interactive; sbehnke@harrisinteractive.com**
Card sorting tasks have been commonly used to measure issues of importance in market research. We sought to compare results that would be obtained when using a card sorting task with results using a more traditional rating scale in an online survey. Respondents were asked to indicate how often they consumed a series of 20 different foods or beverages. This served as criteria to determine the validity of the sorting task or rating scale. Respondents were then either asked to sort the 20 different foods into 1 of 5 categories which were labeled from 'Very much like' to 'Very much dislike' or were asked to indicate how much they liked each of the 20 different foods using a grid format. We found some distributional differences due to task and, though the card sorting task took longer, we found little difference in the average correlation with the key criteria.

COGNITIVE INTERVIEWING

***Cognitive Interview Usability Testing of the Redesigned Sudden Unexplained Infant Death Reporting Form*, Danna L. Moore, Academic; moored@wsu.edu**
The need for consistent coding of cause and manner of infant death, nationwide, has lead to a redesign of the Sudden Unexplained Infant Death Reporting Form (SUIDIRF). This study asked for input from medical examiner/experts and also explored the outcomes of a testing process for evaluating the usability of the visually redesigned SUIDIRF. This study is unique in that the usability testing paired bereaved parents/caregivers with death scene investigators in an interview setting to recreate the death scene investigation (DSI) interview. We were primarily interested in questionnaire improvement as a way to standardize investigator performance in DSI. Investigators used the revised form "cold" and had not previously talked with the person they were interviewing. Investigators varied in characteristics such as type of investigator (police, Medical examiners, or coroners) and investigation experience. Cognitive interviews were assessed using behavioral coding, encoding capture on the form, investigator debriefs, and parent feedback interviews to determine the adequacy and thoroughness of the new form for investigation interviews. The behavior coding of investigator parent interactions were conducted from videotapes which also captured investigator speak alouds, and respondent clarifications. Data captured on forms was also reviewed. Investigator interview style and ability impacted interview outcomes. The power of cognitive testing proved to be the ready identification of questions that caused difficulties for respondents and form inadequacies signaled by investigators probing differently than form questions. The power of expert review was feedback on acceptance of the form and the identification of specific data urgency (i.e. before autopsy, before death certification) as a way to shorten the DSI interview and segment data collection. Thoroughness of investigation was impacted by investigator interview ability, investigator demeanor with respondents, and respondent forthcoming. This study shows differences between standard survey interviewing and investigation interviewing.

***Cognitive Interviewing: the case of verbal reports*, Erik Hjermstad, University of Nebraska-Lincoln; hjerm@bigred.unl.edu, and Monica Sanchez, University of Nebraska-Lincoln; msanchez@bigred.unl.edu**

The constant desire to improve survey quality begins with sound questions and quality testing of those questions; cognitive interviewing is arguably one of the best methods for pretesting questionnaires and improving survey quality. However, there does not appear to be any uniformly accepted standard regarding the verbal reports that respondents provide to probe questions. The current research stems from two schools of thought concerning the content of verbal reports. One idea is that verbal reports should focus solely on the thoughts and experiences surrounding questions and the researcher determines whether or not problems exist. Proponents of this idea believe that the best information is found when respondents report on their own specific thoughts and experiences because they tend to have more expertise surrounding their own experiences. The alternative notion is that respondents are aware of question flaws themselves and should comment on them if they feel compelled to do so. Those in favor of this idea tend to believe that if problems with question wording are solved, the source of those details is not important. The current research will examine recent literature on cognitive interviewing, cognitive psychology, communication theory and other survey pretesting methods to provide a detailed synopsis on the debate surrounding the content of verbal reports.

***A Cognitive Tool to Predict How Well Respondents Perform Survey-Related Tasks*, Patricia A. Gwartney, University of Oregon; pattygg@oregon.uoregon.edu**

Recently, the author, with Anthony Leiserowitz and Vikas Gumbhir, has explored affective imagery as a cognitive tool for understanding the quality of respondents' telephone survey behavior. Affective imagery refers to the meanings respondents give to stimuli elicited in open-ended word association questions. One paper shows that respondents' survey behavior varies systematically with their responses to the stimulus "global warming," e.g., "skeptics" more often refused at first but then gave significantly longer interviews, with significantly longer answers to open-ended question, than respondents answering "don't know." Another paper varied the stimulus, asking for respondents' first thought and image to the keywords "survey," "U.S. Census," and "political poll." Forty percent of respondents could not supply a visual image, consistently across the three stimuli. No matter what their specific answer, "visualizers" produced less item nonresponse, more words per open-ended question, shorter interviews, and ascribed higher importance to survey research generally. In multivariate analyses, visualizing ability proved highly significant, controlling for standard predictors, with twice the net effect of education in explaining item nonresponse. This paper attempts to validate our visualizer findings in parallel dataset (RDD, n=810, Oregonians, 2003). The stimulus item is "organic food." A systematically varied word association question prompts no sensate reply. Just 10% of respondents offered concrete visual images unprompted (alfalfa sprouts, shriveled up vegetables, ugly bug-invested apples, visions of big greenhouses). About 15% more offered vague visual images (fruits, vegetables). Unexpectedly, 2% offered taste-related answers (delicious, better tasting, tasteless, yuck). The forthcoming multivariate analysis will determine whether unprompted sensate answers prove as powerful in explaining respondents' survey behaviors as did prompted visual answers previously. Such results would suggest that affective imagery yields robust indicators of cognitive sophistication, a concept proven key to understanding the quality of survey respondents' behavior.

***Designing Items to Measure Healthy Marriage: Lessons Learned from Cognitive Interviews*, Lina Guzman, Child Trends; lguzman@childtrends.org, Kristin Moore, Child Trends; kmoore@childtrends.org, Greg Matthews, Child Trends; gmatthews@childtrends.org, and Zakia Redd, Child Trends; zredd@childtrends.org**

This paper will discuss the lessons learned from a series of cognitive interviews conducted with a diverse sample of couples in the DC area to test items of healthy intimate relationships for inclusion in a federal evaluation of healthy marriage education programs. A central focus of the cognitive interviews was to determine whether existing items, which have been largely administered to white middle-class populations, were applicable to low-income couples and other diverse populations. Recent qualitative research with low-income couples suggests that existing measures of relationship quality may not adequately tap into issues that affect low-income couples. A second goal was to identify items that worked well (i.e., easily understood, similarly interpreted, etc) across racial and ethnic groups, and across immigrant and family status, gender, and income groups. Preliminary findings indicate that the domains identified by the research team and prior research match well with couples' perceptions about what defines a healthy marriage. Feedback from participants suggests that, while most of the items chosen or developed worked well with diverse populations, some presented problems. Items that were negatively worded or that used contractions or abbreviations were particularly problematic. In addition, the interviews highlighted the need to ensure that the language and terminology used in instruments administered to populations with a range of educational backgrounds is as clear and simple as possible. For example, though care was taken to word items at a fifth-grade level, some terms that were anticipated to be part of a common vernacular caused respondents problems. Interviews with immigrants helped to identify items that were less easily understood for respondents raised in other countries, though they worked well with American-raised respondents. Last, the interviews indicate that the appropriateness of items to measure violence in intimate relationship may vary by gender.

***Does Cognitive Interviewing Really Improve Survey Quality?*, Stanley Freedman, Energy Information Administration; stanley.freedman@eia.doe.gov, Robert Rutchik, Energy Information Administration; robert.rutchik@eia.doe.gov, and Kara Norman, Energy Information Administration; kara.norman@eia.doe.gov**

This paper evaluates the effectiveness of cognitive interview techniques for improving data quality and survey processing at the Energy Information Administration (EIA). Specifically, it focuses on whether pre-survey design visits and cognitive interviews on draft survey forms resulted in measurable improvements in both data quality and survey processing. EIA recently completed a major redesign of one of its electricity surveys. The survey was difficult for some types of respondents to complete, required a great deal of follow-up with these respondents after data submission, and did not always provide high quality data. The redesign involved a review of data requirements, pre-survey design site visits with respondents to identify business record keeping practices, a detailed review of the correspondence between business records and the data EIA required, and cognitive interviews on draft survey questionnaires. This year long redesign effort resulted in a new survey instrument and instructions for a subset of respondents that had been having difficulty reporting to EIA. EIA believed that the new survey would provide more accurate information, be less burdensome to respondents, and would require less processing time and follow-up at EIA. To verify this, EIA is conducting an evaluation of the data collected and survey processing for the redesigned survey. This paper describes the problems with the original survey, the cognitive interviewing procedures used to improve the survey, and the findings from the evaluation study. Both qualitative and quantitative results are presented.

INCENTIVES III

***Unintended Consequences of Incentive Induced Response Rate Differences*, Duston Pope, Market Strategies, Inc.; duston pope@marketstrategies.com, Scott D. Crawford, Survey Sciences Group, LLC; scott@surveysciences.com, Eric O. Johnson, Henry Ford Health Science Center; EJOHNSO3@hfhs.org, and Sean Esteban McCabe, University of Michigan; plius@umich.edu**

Studies are finding that many previously held hypotheses about the effectiveness of varying incentive approaches remain intact when implemented in Web surveys. Given low per respondent costs in Web surveys, researchers take advantage of the ability to increase sample sizes, and frequently use sweepstakes based incentives, where respondents are entered into a drawing to win a fixed prize or prizes. The effectiveness of this kind of incentive has been mixed. However, a recent experiment of first contact mode and incentives types in a Web survey of 2,500 undergraduate college students conducted by these authors demonstrated that: 1. First respondent contact send via US Mail improves response rates over the use of email for first contact. 2. A sweepstakes drawing for $500 improves response rates over the use of no incentive. 3. A pre-paid $2 bill improves response rates over the use of the sweepstakes drawing for $500. 4. Use of a pre-paid $2 bill AND a drawing for $500 improves response rates over the use of either single incentive alone. In this study, the resulting response rate (AAPOR RR2) spread between treatment groups was 28%. The common assumption would be that with the increased response rates comes higher data quality. In looking at key measures in this study, we have found that there may in fact be significant data quality differences between the modes, including a finding of significantly different measures of smoking prevalence between experimental conditions. Singer (2002, Chapter 11 in Survey Nonresponse, eds. Groves, et al.) described that while incentives have their well documented intended consequences, they may as well have many unintended consequences on data quality. In this presentation we will discuss what we have found with regards to the impact these incentives have had on missing data, key variable response distributions, and sample composition.

***The Use of Monetary Incentives in the Survey of Income and Program Participation*, Denise Lewis, U.S. Census Bureau; denise.c.lewis@census.gov, and Kathleen Creighton, U.S. Census Bureau; kathleen.p.creighton@census.gov**

The U.S. Census Bureau began using monetary incentives in the Survey of Income and Program Participation (SIPP) in an attempt to combat the steadily increasing nonresponse that occurred over the life of the sample. Becuase the SIPP provides national estimates of sources, amounts, and determinants of income for households, families, and individuals, it is the primary source of data for evaluating the effects of welfare reform. Therefore, it is imperative to achieve a high response rate so that policy makers have an accurate picture of the economic situation of all Americans. The literature concludes that the positive effects of incentives in mail surveys also hold in surveys conducted by interview in person or by telephone as well as for fresh respondents, panel respondents, and nonrespondents (Singer et al. 1999). Multiple experiments were embedded in the 1996 and 2001 Panels. These experiments tested the differentail effects of: $0, $10, or $20 unconditional incentives using both paper vouchers and plastic debit cards; $20 'booster' incentives for poverty household; and $0, $20, or $40 conditional incentives for previous wave nonrespondents. In this presentation, we provide a historical context of the 1996 Panel experiments; present the results (i.e., response rates, hard refusal rates, receipt of multiple incentives, and respondent reaction to receiving an incentive) of the 2001 Panel experiments which represent the last opportunity to benchmark response rates using a control group; and, conclude with the preliminary results of the 2004 Panel in which the use of incentives became standard practice.

***The Impact of Providing Incentives to Initial Refusal Cases on Sample Composition and Data Quality in a Telephone Survey*, Douglas Currivan, RTI International; dcurrivan@rti.org**

Although considerable research has examined the impact of various incentive plans on survey response, few studies have evaluated the impact of offering incentives to initial refusers on the final composition of the sample or the quality of survey data. Providing incentives to initial refusers can potentially impact both the composition of survey respondents and the quality of responses they provide. Offering an incentive payment based on initial refusal may increase the participation of sample members who would otherwise be difficult to include in the survey. As often intended, this outcome would produce a more representative set of respondents. Research also suggests that incentives may persuade sample members who have little interest in the survey topic to participate. Respondents with little interest in the survey topic might be less motivated to provide complete and accurate data than other sample members, even though their participation may reduce nonresponse error. The combined outcome of such incentive plans could therefore be increased participation from some sample members, but lower quality data from these respondents. To investigate the impact of offering incentives to initial refusers, this study evaluates an incentive plan currently being implemented in a statewide telephone survey, the New York Adult Tobacco Survey. In this survey, respondents are not originally offered any incentive to complete the interview, but are then offered $20 if they initially refuse to participate. Comparisons of key characteristics among sample members who received the incentive versus those who did not will indicate whether this incentive influenced the composition of respondents. Comparisons of response patterns to key survey items among paid and unpaid respondents - including analysis of missing data, "no opinion" answers, and other indicators of low respondent motivation - will indicate whether the incentive affected data quality. The implications of these analyses for similar incentive schemes will be discussed.

***Incentives: Do You Get What You Pay For?*, Ronald Broach, National Opinion Research Center; broach-ronald@norc.uchicago.edu, and Karen Grigorian, NORC; grigorian-karen@norc.uchicago.edu**

In the current environment of declining response rates, many researchers debate whether to include an incentive, what the incentive should be and how it should be given to the respondent, and how effective an incentive will be in increasing response rate. It is much rarer to see an evaluation of the effects of incentives on the data collected. The researchers conducting the 2003 Survey of Doctorate Recipients implemented a controlled experiment employing a $30 prepaid incentive, a $50 postpaid incentive, a non-monetary incentive, and a control group, which received a gaining cooperation letter only. In this presentation we offer a quantitative and qualitative evaluation of the incentive experiment results. As a quantitative measure, we will provide response rate information. More importantly, we will provide a qualitative comparison of cases from each treatment group (measured by the amount of data editing and cleaning required). Quantitative comparison also includes an evaluation of the amount of item non-response as well as the quality of data at selected key data items. The incentive experiment was implemented late in data collection, and so all members of the experiment and control groups were non-respondents at the time of the treatment implementation. All evaluated cases had already received at least one mail questionnaire, access to the Web-based version of the questionnaire, and telephone contact from interviewing staff. Finally, we will provide a cost comparison of the incentive groups and an overall evaluation of these findings.

VOTING: FOREIGN AND DOMESTIC

***Economic Perceptions and Voting Behaviors in the Second Swedish Election to the EU-PARLIAMENT*, Lingrui Jiang, UNL-Gallup Research Center; lily_wind2000@hotmail.com, Yu Feng, UNL-Gallup Research Center; gemfeng@hotmail.com, and Allan McCutcheon, UNL-Gallup Research Center; amccutcheon1@unl.edu**

Conventional wisdom asserts that economic conditions are closely linked to election outcomes. A growing body of literature shows that over-time changes in government approval, and election outcomes are, at least in part, structured by the public's perception of the government's economic performance. Researchers have found evidence of effects of the evaluation of economic performance on vote choice at the individual level (see e.g., Clarke et al. 2004). In this paper, the Swedish population's perceptions and evaluations of the European Union's economic performance following Sweden's 1995 entry into the EU are modeled using latent class models and data from the EU-PARLIAMENT ELECTION STUDY of 1999. The distribution of Swedish adults in each latent class clearly indicates widespread ambivalence regarding the EU's fin-de-siecle economic performance; nearly 2 of 3 respondents report mixed evaluations of the economic performance, while about 1 in 5 are negative, and the remaining 15% are more generally positive. The latent variable measurement model for the evaluation EU economic performance is included in a model of factors influencing voting intention and behavior. This study examines the hypothesis that economic perceptions strongly affect both vote intention and party choice in the second Swedish EU-parliamentary election. Multinomial logistic regression models are used to analyze the influence of economic evaluation perception, as well as other variables (e.g., age, gender, education, household income), on predicting voting behavior. References Clarke, H. D., D. Sanders, M. C. Stewart, and P. Whitely (2004) Political Choice in Britain. Oxford: Oxfor University Press.

***Response Privacy and Elapsed Time Since Election Day as Determinants for Vote Overreporting*, Volker Stocké, University of Mannheim, SFB 504; vstocke@rumms.uni-mannheim.de**

Survey respondents systematically overreport their participation in political elections and this bias was found to affect the results from research about the determinants of electoral participation. Researchers have argued that memory failure and social desirability bias cause vote overreporting. Much research has been done about the demographic characteristics and political attitudes which are associated with overreporting. Only few studies analyzed determinants which survey researchers are able to influence. In our study we utilized data from a field experiment with a locally defined random sample in order to test firstly whether an ensured privacy of the response situation reduces overreporting. This was done by randomly assigning the respondents to either an interviewer-administered (CAPI) or self-administered (CASI) mode of data collection. It was expected that the increased privacy under self-administration reduces the subjects' impression management motive and thus causes self-reported and official turnout to be significantly more congruent. Available research tested for a time span up to 8 months whether a longer time elapsed between the election day and the survey interview leads to more vote overreporting. This was confirmed empirically. However, some research designs require that respondents report their behavior in elections which were held several years ago. It is thus an interesting, but empirically unanswered question whether the probability of vote overreporting still increases when the differences in elapsed time are in the magnitude of years rather than months. Respondents were thus asked about their electoral participation in three federal elections in Germany, which were held 3.7, 7.6 and 12.3 years before the survey interviews. Although we found a tendency for stronger vote overreporting for elections more distant in time, this effect did not prove to be statistically significant. The empirical analysis however confirmed that an ensured response privacy significantly reduces the differences in self-reported and official turnout.

***A Comparison of Presidential Candidate Vote Intention Measures in U.S. Elections*, Randall Thomas, Harris Interactive; rthomas@harrisinteractive.com, David Krane, Harris Interactive; dkrane@harrisinteractive.com, Mitchell Sanders, Harris Interactive; msanders@harrisinteractive.com, and Susan Behnke, Harris Interactive; sbehnke@harrisinteractive.com**

Through a series of 5 pre-election web-based studies, we randomly assigned respondents to either a bipolar or unipolar presidential candidate vote intention scale (with the item stem being 'How likely are you to vote for George W. Bush?' and 'How likely are you to vote for John Kerry?'). We compared the likelihood measures with a categorical candidate preference measure ('Who would you vote for?') to determine concurrent validity. We then recontacted these respondents following the election to inquire who they voted for to determine the predictive validity of these measures. We compare our results with previous research that indicated that unipolar behavior intention measures (i.e. "Not at all likely" to "Extremely likely") have somewhat higher validity than bipolar behavior intention measures.

***A failing Tool or Changing Society? The Validity of the "Strength-of-Personality Scale" and the Role of Opinion Leaders in the 2002 German Federal Election*, Thomas Petersen, Institut für Demoskopie Allensbach; tpetersen@ifd-allensbach.de**

The theory of "opinion leaders" is one of the oldest and most influential in empirical election research and media effects research. Formulated by Paul F. Larzarsfeld in 1944, the theory posits that opinion leaders in all social classes, i.e. individuals whose personality traits make them especially influential in their own social circles, play a key role in the process of public opinion formation. This theory went hand in hand with the theory of the two-step flow of communication, which maintains that news and opinions first flow from the mass media to these opinion leaders and are then passed on by these persons to broader segments of society. In the early 1980s, the Allensbach Institute succeeded in developing a strength-of-personality scale that enabled us to identify those personalities that displayed the traits associated with opinion leadership. Survey findings repeatedly demonstrated, for example, that people who scored high on the strength-of-personality scale led the remaining population in the opinion formation process. In the 2002 German federal elections, however, survey findings revealed a completely different pattern: those people identified via the strength-of-personality scale as "opinion leaders" followed the rest of the population in the opinion formation process. There were only two plausible explanations for this finding: either the strength-of-personality scale was no longer able to identify opinion leaders or there had been some fundamental change in how the population forms its opinions that had prevented opinion leaders from playing their customary role. An analysis of the findings of election research conducted by the Allensbach Institute, which we also compared with the findings of media content analyses, would seem to offer support for the latter explanation, thus suggesting that the theory of the two-step flow of communication may be severely limited by modern media structures today.

SELF-REPORTS AND HEALTH

***Assessing the Quality of Height and Weight Data Collected During Telephone Interviews*, Keith Smith, Abt Associates, Inc.; Keith_Smith@abtassoc.com, and Marjorie Morrissey, Abt Associates, Inc.; Marjorie_Morrissey@abtassoc.com**

The collection of accurate height and weight data is critically important for health studies. The combination of height and weight yields a person's body mass index (BMI), which is a measure used to determine if a person meets the threshold for clinical obesity. This paper analyzes height and weight data collected from two studies sponsored by the Centers for Disease Control and Prevention (CDC) that examine the prevalence and incidence of Chronic Fatigue Syndrome (CFS). CFS is prolonged fatigue that cannot be explained by another medical or psychiatric condition. Height and weight are important measures in CFS research because obesity is one of several medical conditions that excludes a diagnosis of CFS. The analysis in this paper focuses on the self-reported height and weight data collected during telephone interviews and compares those data to the actual height and weight collected during physical examinations. Our paper examines the differences in these two sets of data and compares the results among various demographic groupings (sex, age, race, ethnicity). It also discusses the implications for health studies that collect this data via telephone interviews and examines methods and strategies for collecting more accurate data.

***Designing Self-Administered Questionnaires for Health Center Personnel to Evaluate a Quality Improvement Initiative*, Jessica Graber, NORC at the University of Chicago; graber-jessica@norc.org, and Anne Kirchhoff, University of Chicago; akirchho@medicine.bsd.uchicago.edu**

NORC, the MidWest Clinicians' Network, and the University of Chicago are conducting an innovative, multi-method evaluation of the Health Disparities Collaborative (HDC) initiative at more than 150 community health centers (CHCs) in the Midwest and West Central regions of the United States. This research is the first large-scale evaluation of the HDC, a quality improvement initiative that aims to enhance the treatment of chronic illness through evidence-based practices, clinical decision support systems, and patient self-management. Critical to our evaluation are targeted, self-administered questionnaires completed by health center leadership, health care providers, and non-provider clinic staff on the functionality, success, sustainability and potential expansion of the HDC initiative. Our research provides an original contribution to the literature, as there are no large-scale survey instruments specific to center employees on quality improvement collaborative efforts in health centers. Recognizing the challenge of obtaining cooperation from time-constrained health center staff, we aimed to develop a focused, relevant questionnaire that would facilitate participant response and provide insight on the complex nature of these organizations and their improvement efforts. Position-specific instruments for CHC personnel were developed using results from the qualitative component of our research, with further input from health center personnel on the content and wording of questions. In this paper we present how qualitative work informed our complex questionnaire design process and its utility in interpreting resulting survey data. For example, as found in qualitative interviews, personnel and funding issues emerged in the questionnaire data as important themes. In addition, we will describe the challenges of developing measures to sufficiently assess the complex CHC environment and present early survey results. Preliminary analyses show satisfactory spread for responses, an indication of well-developed instruments. Lastly, we present response rates across the multiple respondent types to demonstrate the effectiveness of conducting a mail survey with CHC personnel.

***Reliability Of Self-Reported Conditions In Health Surveys: Findings From The Army Chemical Corps Health Survey*, John Boyle, Schulman, Ronca & Bucuvalas, Inc.; j.boyle@srbi.com**

Epidemiological research relies on population estimates of medical conditions from health surveys. Despite a number of studies over the past fifty years, no consensus has emerged about the validity and reliability of survey estimates of chronic disease. The Army Chemical Corps Health Study was designed as a long-term health assessment of a sample of approximately 4,000 veterans who served in the Chemical Corps during the Vietnam War era. Health and fertility outcomes for cases and controls were assessed in a 45-minute telephone interview. Since the validity of self-reported medical conditions could not be determined from medical records for the entire sample, a study was undertaken to assess the reliability of the survey reports of medical conditions as a potential silver standard for comparative disease prevalence in the two populations. A sample of 495 cases was drawn from 2500 interviews that had been completed by the end of November 1999. This sample was stratified by case/control status (Vietnam/not Vietnam) and by consent for medical records (no medical conditions requiring consent/medical conditions with consent/medical conditions without consent). The retest survey that was initiated approximately eight months after the initial interview was limited to questions about the lifetime prevalence of eleven health conditions from the baseline survey and the date of diagnosis if a condition was reported. Ultimately, 449 out of the sample of 495 subjects (91%) were re-interviewed in the retest survey. This paper reports the reliability of chronic condition reports in health surveys based on repeated measures of the same conditions by the same respondents less than one year apart. In addition to the overall reliability of reporting in health surveys, the paper explores differences in reliability by condition and type of respondent that might explain the sources of variability in reporting of medical conditions.

***Aging and Errors of Measurement in Surveys -- The Case of Self-Reports of Health*, Duane Alwin, Pennsylvania State University; dfa2@psu.edu, Ryan McCammon, University of Michigan; mccammon@isr.umich.edu, Willard Rodgers, University of Michigan; wrodgers@isr.umich.edu, and Halimah Hassan, University of Michigan; halimah@isr.umich.edu**

Because of the cognitive demands of many survey questions and given well-documented laboratory and survey results regarding declines associated with aging in some domains of cognitive functioning, it has been suggested that there may be identifiable differences in errors of measurement associated with the respondent age in sample surveys of the general population. To date very few age-related differences in measurement error have been reported in the literature. Using data from the Health and Retirement Study (HRS) - a large national longitudinal survey of middle-aged and older adults conducted by the University of Michigan's Survey Research Center (N = 21, 410) - this paper reports the results of a project whose goal was to examine age differences in reporting reliability for several self-report measures of health. The key focus of the present paper involves an analysis of one of the most common self-report measures of health, a five-category rating of health - "Would you say your health is excellent, very good, good, fair, or poor?" Analyses of age-differences in reliability of measurement in the HRS data reveals systematic and significant linear declines in reliability with age. These patterns in reliability are shown to derive from a combination of two processes: (1) a systematic age-related decline in true variance in self-rated health, and (2) the systematic increase in measurement error variance with age. These patterns are discussed within the context of the existing literature on age-related errors of measurement and the implications of these results are considered for the study of aging using self-report measures of health.

LAYING THE GROUNDWORK FOR INCLUDING CELL PHONE NUMBERS IN TELEPHONE SURVEY SAMPLES

***Final Results from the 2004 CPS Supplement on Telephone Usage*, Clyde Tucker, Bureau of Labor Statistics; tucker_c@bls.gov, Michael Brick, Westat; mikebrick@westat.com, and Brian J. Meekins, Bureau of Labor Statistics; meekins.brian@bls.gov**
Some recent changes in the U.S. telephone system, including the size of the universe of numbers and how they are assigned, number portability and the rapid growth of cellular phone usage, raises concern about coverage error and productivity. Recent trends have undermined confidence in RDD designs. In February 2004, a supplement to the Current Population Survey on telephone service was conducted. Additional data collection was undertaken in November to obtain survey results from a representative sample of the population inadvertently missed in February. This paper incorporates these respondents , those who said they did not have access to a telephone, into estimates of the percentage of households with landline only, no landline but wireless service, both landline and wireless service, and no telephone service. Standard errors (using BRR) are calculated that take into account the complex sample design. Analyses will include detailed demographic profiles of the four telephone service subpopulations, measurement of the effect of telephone service type on labor force estimates, and estimation of the respondents' propensity to respond to subsequent waves in the panel. In addition, the trend line that measures the growth in cell-only household using the Consumer Expenditure Survey will be extended. All of this information is of critical importance to developing efficient telephone survey methodology in these changing times.

***Surveying Cell Phone Households - Results and Lessons?*, Y. Angela Yuan, Westat; angelayuan@westat.com, Bruce Allen, Westat; BruceAllen@Westat.com, Michael Brick, Westat; mikebrick@westat.com, Sarah Dipko, Westat; SarahDipko@westat.com, Stanley Presser, JPSM; spresser@survey.umd.edu, Clyde Tucker, Bureau of Labor Statistics; tucker_c@bls.gov, Daifeng Han, JPSM; daifenghan@westat.com, Laura Burns, JPSM; lburns@survey.umd.edu, and Mirta Galesic, Joint Program in Survey Methodology; mgalesic@survey.umd.edu**
As a result of the growing cellular-only population, the validity of estimates from traditional RDD surveys is being questioned. Previous research shows that over 5 percent of households in the U.S. are cellular-only, a number that appears to be growing rapidly. People who live in households without wired phones are missed in traditional RDD surveys. In addition, persons in households with both wired and wireless phones may be harder to reach in traditional RDD surveys than are those with only wired phones. Thus it is important to evaluate the possibility of sampling households from a frame including cellular numbers. The 2004 JPSM Practicum Survey of the University of Maryland examined the feasibility of sampling households from a frame of cellular phone numbers and interviewing people on their cell phones. The JPSM study sampled 8500 phone numbers from the cell frame and 4688 phone numbers from the traditional RDD frame. Data were collected by Westat between July and September 2004. Interviews averaged about 10 minutes and asked a household respondent about telephone ownership and usage, attitudes towards cell phones, social behaviors, and demographics. A text message notification experiment and an incentive experiment were embedded in the cell phone survey. This paper describes the design and operational issues surrounding sampling and data collection in this dual-frame study. Major findings that will be presented include the response rates from the two samples, patterns of accessibility and cooperation by type of sample, and the effectiveness of the text message and incentive experiments for the cell sample.

***An Investigation of Response Differences between Cell Phone and Landline Interviews*, Sarah Dipko, Westat; SarahDipko@westat.com, Pat Dean Brick, Westat; PatDeanBrick@WESTAT.com, Michael Brick, Westat; mikebrick@westat.com, and Stanley Presser, JPSM; spresser@survey.umd.edu**
The recent proliferation of cell phones has generated much speculation regarding their potential effect on surveys and particularly RDD surveys. To collect information on cell phone usage and cell phone users, the 2004 JPSM Practicum of the University of Maryland conducted a survey that included a substantial cell phone sample. One open-ended question was included to investigate response differences between those interviewed on cell phones and those interviewed on landlines ("What do you like about having a cell phone?"). This was asked of all respondents in both samples who reported having a cell phone. All responses were recorded verbatim and a standard probe for additional information was delivered. We plan to analyze this open-ended question by coding the content of responses and doing counts of the number of characters, words, and phrases. We wish to investigate whether there are any differences in open-ended responses by telephone device used. Cell phone respondents may be concerned about the minutes used (and related cost) for the interview, which may result in brevity of response for cell phone interviews as compared to landline interviews. For the cell phone sample, the data set includes their location at the time of interview, which allows examination of contextual effects, that is, effects related to being interviewed while in the car, grocery store, etc. We will also investigate possible interviewer effects by telephone device by conducting a hierarchical analysis with the interviewer at one level and respondent at the second level. To avoid confounding the analysis (i.e., individual-level analysis from a household-level survey), only household level variables will be used. For the hierarchical modeling, we expect this not to be an issue because we have interviewer characteristics at the first level and respondent household characteristics at the second level, with probably no interactions across levels

MAIL SURVEYS AND ADVANCE CONTACT

***Does envelope size affect respondent participation?*, Kerry Levin, Westat; kerrylevin@westat.com, Martha Kudela, Wesat; marthakudela@westat.com, and Karen Masken, IRS; Masken Karen C [Karen.C.Masken@irs.gov]**
Nonresponse continues to be a problem in household surveys (Singer, Hoewyk, and Maher, 2000). Researchers (e.g. Dillman, 1978; 2000) have tested a number of different methods for increasing survey response such as the use of incentives, varying the design of mailing materials, implementing different mailing procedures, and varying the types of materials sent in prenotification mailings. The current study is a telephone survey of low-income respondents. About halfway through data collection, we attempted to increase response by mailing a letter encouraging participation to 979 respondents who had not yet completed an interview. The letter explained that we had been unable to complete an interview with them and the importance of their participation. The letter was signed by a high-ranking official within the survey sponsor's organization and it appeared on letterhead especially designed for the project. To examine whether the size of the

envelope would impact survey response, our experiment involved randomly assigning half the respondents to receive the letter in a #10 window envelope and the other half to receive the letter in a larger (9x12) envelope. The mailing envelopes, regardless of size, displayed the same logo and other design elements as the letterhead. We hypothesized that respondents would be more likely to open and therefore read the letter sent in the larger size envelope. If they read the letter, we expected higher response than if they did not read the letter. To determine the impact of envelope size on survey participation, we will compare the response rates of the two groups. Preliminary results show there is no difference between the two conditions (17.5 percent of respondents who received the letter in the larger envelope completed an interview and 16.5 percent of those who got the smaller envelope completed an interview).

***Mail survey non-contact identification: Who did we not get through to?*, Benjamin Healey, Massey University; b.j.healey@massey.ac.nz, and Philip Gendall, Massey University; p.j.gendall@massey.ac.nz**

Mail survey sampling frames such as electoral rolls inevitably contain out-of-date address information for some potential respondents. In addition to the financial cost to researchers of undelivered survey packages, the misaddressing that results from out-of-date data can lead to inaccuracies in reported response rates and bias in the population estimates inferred from survey results. To date, little research has been undertaken to investigate the extent of mail survey non-contact, or whether certain sub-populations are more likely to be prone to it than others. Furthermore, it is difficult for researchers to accurately assess non-contact in surveys because 'Return to Sender' receipts do not fully represent the number of invitations that did not reach their intended recipients. Our study examined misaddressed survey return rates by sending deliberately misaddressed survey invitations to a sample of 1400 New Zealand households. The effectiveness of an envelope message encouraging recipients to return misaddressed items to the sender, aimed at improving overall 'Return to Sender' receipts, was tested as part of the study. Additionally, household composition was examined to identify demographic differentiators of misaddressed survey returning behaviour. Only 53% of the intentionally misaddressed envelopes without a 'please return' message were returned, compared to 67% of those with the envelope message. In other words, a large proportion of misaddressed survey mailings are likely to not be returned, but this proportion can be reduced by a simple instruction on the outer envelope. Analysis of demographic characteristics suggests that households composed of individuals more likely to be 'independent itinerants' (e.g., younger people, renters, and those living in cities) are least likely to return a misaddressed survey envelope. The implications for survey researchers for mail survey cost management, response rate calculation, potential bias identification, and the development of sampling methodologies to counter survey error introduced by non-contacts are discussed.

***Testing an Advance Contact Targeted-Awareness Campaign to Raise Response Rates*, Ana Patricia Melgar, Nielsen Media Research; Ana.Melgar@nielsenmedia.com, Paul J Lavrakas, Nielsen Media research; paul.lavrakas@nielsenmedia.com, Agnieszka Flizik, Nielsen Media Research; Agnieszka.Flizik@nielsenmedia.com, Rosemary Holden, Nielsen Media Research; Rosemary.Holden@nielsenmedia.com, and Maria Anatro, Nielsen Media Research; Maria.Anatro@nielsenmedia.com**

Following from Leverage-Saliency theory, the more a sampled respondent is aware of a survey organization, and the more favorably disposed is the respondent to the survey organization, the higher should be the propensity for the respondent to participate in a survey conducted by that organization. Conversely, the less aware a respondent is of the survey organization, and the more negatively disposed is the respondent to the survey organization, the lower should be the propensity for the respondent to participate when sampled by that organization. One way to try to make respondents aware of and positively disposed towards a survey organization is to do a mass media advertising campaign. Another way is to use a targeted-awareness campaign with only those households that will be sampled for a survey. In 2004 and early 2005, Nielsen Media Research is conducting an experimental test in the Atlanta metro area of the effect on response propensity of sending a series of three mailings to sampled addresses prior to being contacted to participate in two surveys. The goal of the advance mailings is to (1) raise the household's awareness of Nielsen and (2) make them positively disposed towards Nielsen. Included in each of the three mailings is a cover letter, an informational piece, and a small gift (each item being different in each mailing). The Atlanta area-probability sample was drawn in early 2004. Approximately 850 addresses were randomly assigned to receive the three mailings, with approximately 2,550 addresses randomly assigned to the control group. Dependent variables in this experiment will be (1) mail and in-person response rates to the first survey and (2) in-person response rate to the second survey. Other dependent variables will include the effort and cost in the control group versus the treatment group of recruiting the households for each survey.

***Is it Worth the Effort: RDD Telephone Surveys and Advance Survey Notification*, Leslyn Hall, ORC Macro; leslyn.m.hall@orcmacro.com, Randall ZuWallack, ORC Macro; Randall.S.ZuWallack@ORCmacro.com, and Kirsten Ivie, ORC Macro; Kirsten.B.Ivie@orcmacro.com**

This paper investigates seeks to investigate the potentially unintended consequences of mailing an advance letter to RDD telephone survey respondents. While it is generally accepted that advance notification for any survey has positive effects on response rates, what is less clear are the effects on survey estimates. Using the Behavioral Risk Factor Surveillance System RDD Telephone Surveys for a number of states, and the US Department of Housing and Urban Development's Fair Market Rent (FMR) Regional RDD Telephone Surveys, ORC Macro conducted experiments whereby the company mailed advance letters to randomly selected households of the RDD telephone sample members where addresses could be ascertained, and then did not mail advance letters to the other listed and unlisted households. Initial analyses of the FMR survey data indicated that perhaps the increased costs and efforts to obtain these higher response rates did not significantly affect the statistic of interest -- the contract rental value for two bedroom apartments; what was less clear is if the increased response rate affected the bias of the estimate. This paper seeks to thoroughly examine how the advance notification affects response and cooperation rates for these two different RDD telephone surveys by evaluating whether or not the apparent receipt of advance notification introduce any significant bias to the variables of interest; does it introduce a new source of error into estimates by increasing the potential representation of the list portion of the RDD telephone-sampling frame. To the extent that differences between the two survey experiments are found, this paper then goes on to consider how the surveys' advertised content effects participation and results.

MODES AND METHODS

***Web survey response rates compared to other modes - A meta-analysis*, Katja Lozar-Manfreda, University of Ljubljana, Slovenia; katja.lozar-manfreda@fdv.uni-lj.si, Michael Bosnjak, University of Mannheim; bosnjak@tnt.psychologie.uni-mannheim.de, Iris Haas, University of Mannheim, Dept. of Psychology II; irishaas@gmx.de, and Vasja Vehovar, University of Ljubljana, Slovenia; vasja.vehovar@uni-lj.si**

Anecdotal literature reviews suggest that in general, Web survey response rates are considerably lower. However, such unsystematic evidence might be highly misleading. As an alternative, meta-analytic procedures synthesizing controlled experimental mode comparisons could provide accurate answers, but to the best of our knowledge, these studies have not been conducted so far. To overcome this gap, we conducted a meta-analysis of 36 published and unpublished experimental comparisons between Web and other survey modes. We took the natural log of the odds-ratio as our effect size measure, weighted for the inverse effect size variance. Subsequently, a homogeneity analysis and a meta-regression were performed. To predict the response rate variance on the effect size measure, the following study descriptors were included: type of alternative mode, year of study, type of target population, sponsorship, and implementation procedures. While the weighted mean effect size indicated a slight advantage for other than Web modes, the 95% confidence interval around this mean effect size estimate included zero, suggesting that there are no systematic response rate differences. A homogeneity test revealed that the 36 effect sizes are homogeneous, indicating that the study descriptors are not systematically related to the effect size variability. This result was further supported by a meta-regression with the study descriptors included as independent variables, none of them exerting a significant predictive effect on the effect size variability. Taken together, our meta-analysis highlights that the common assumption of lower response rates for Web surveys compared to other modes does not hold true if scrutinized with the aid of meta-analytic research synthesis procedures, taking into account experimentally controlled primary studies. From a practical point of view, these results might contribute to improve the reputation of Web surveys as one survey mode of comparable data quality with respect to response rates.

***Mode Effects in Customer Satisfaction Measurement*, Darby Steiger, The Gallup Organization; darby_miller_steiger@gallup.com, Linda Keil, The Gallup Organization; linda_keil@gallup.com, and Greg Gaertner, The Gallup Organization; greg_gaertner@gallup.com**

This paper describes the results of a field test of a redesigned monthly customer satisfaction measurement questionnaire for the largest business customers of a national package shipping company. In response to a desire to allow these large business customers to provide their feedback via the Web using a mixed-mode design, Gallup conducted a field test experiment to measure response differences between CATI and web versions of the survey. One concern in multiple-mode studies is that the method of data collection can affect the distribution of responses. Recent studies suggest that when a scale is visually presented (as in a Web survey), respondents are more likely to choose the middle alternatives, than when it is exclusively oral (as in CATI) (see Dillman, et al, 2000). Thus a main research question was whether the introduction of a Web option would tend to reduce the proportions of responses in the most favorable and least favorable categories (using a Likert scale). In a customer satisfaction survey, this shift could have significant effects on assessment of the organization or company. The design of the field test allowed us to test this and other hypotheses about mode effects. · What mode effects, if any, were observed in the data? · What are respondents' preferences with respect to mode of survey administration? · What mode differences were observed in response rates, breakoff rates, and survey length? Results of the field test indicate no consistent mode effects on favorability (top 2 box scores) towards the company's products and services. However, web responses are significantly more likely to be clustered in the middle categories than CATI responses, which tend to be more extreme. Thus, inclusion of the Web option would tend to decrease the percentages of most favorable and least favorable responses.

***Prompting Efforts to Raise Response Rates for a Web-Based Survey*, Lekha Venkataraman, NORC; venkataraman-lekha@norc.uchicago.edu, and Maggie Parker, NORC; parker-maggie@norc.uchicago.edu**

As populations become increasingly computer literate, Web-based surveys will become a less problematic and more cost efficient medium to reach larger audiences. While extensive research has been conducted on prompting efforts for mail surveys, limited focus has been placed on effective prompting procedures for web-based surveys. It has been shown that using multiple contacts and modes yields higher survey response rates. Email prompting offers convenient accessibility to completing an online questionnaire that is not possible when using other contact mechanisms. This paper will focus on the results of the prompting efforts for the Gates Millennium Scholars Tracking and Longitudinal Study First Follow-up Web survey. It will examine the relative effects of a targeted email prompting effort throughout the last month of data collection, after more conventional prompting techniques were implemented. It will answer the question "Does email prompting help to increase response rates when conducting a Web based survey?" The paper will also discuss features unique to email prompts which can further encourage respondents to complete an online questionnaire. Preliminary results of the email prompts indicate that emails are effective at increasing response rates. These findings may help to inform future prompting efforts for Web-based surveys and shed light on the value of email prompting.

***An Investigation of the Impact of Departures of Standardized Interviewing on Response Errors in Self-Reports about Child Support and Other Family-Related Variables*, Jennifer L. Dykema, University of Wisconsin, Madison; dykema@ssc.wisc.edu**

***Mobilizing Information Online: The Effects of Primary-Source and Secondary-Source Website Use on Political Engagement*, Bruce Hardy, Cornell University;**

INTERVIEWERS

***Person-Based Data Collection in Practice: An Evaluation of Interviewer/Respondent Interactions*, Jennifer Hunter, US Census Bureau; jennifer.e.hunter@census.gov, and Ashley Landreth, U.S. Census Bureau; ashley.denele.landreth@census.gov**

Often survey researchers make the assumption that interviews will be conducted in the way they are scripted, particularly when interviewers are trained in standardized interviewing techniques. Although conversational interviewing techniques have been studied and are recommended in some situations (see Conrad, F. G. and Schober, M. F., 2000, and Schober, M. F. and Conrad, F. G., 1997), interviewers normally receive training in such techniques prior to implementing them. When an interviewer spontaneously changes the script of a standardized questionnaire, the interviewer may turn a standardized interview into a conversational one without having the

proper training. We expect that this would lead to differing data quality. This paper investigates a person-based method of data collection in which a series of demographic questions is asked for each person in the household. Interviewers are instructed to read the questions as worded, using standardized interviewing techniques. However, in field observations it became apparent that the more people that were in the household, the less likely the interviewer was to continue reading the questions as scripted. We are interested in whether or not this practice leads to more respondent behavior problems. To address this, we behavior-coded a sample of interviews using a demographic questionnaire. We analyzed the interviewer/respondent interactions for the first person in the household and compared that to the interactions for the same questions about additional people in the household. The hypotheses were that when the questions were asked the first time, they would be asked in a more standardized way than when they were asked subsequent times and that deviating from the standardized wording would lead to more respondent behavior problems. We report differences in both interviewer and respondent behavior that stemmed from using a more conversational method of interviewing for later persons in the household.

***Impact of Interviewer Effects on Survey Inference*, Trivellore Raghunathan, University of Michigan; teraghu@umich.edu**
The face-to-face survey is often considered as the mode of data collection that produces highest quality data and usually high response rate. It is also recognized that such surveys are perhaps more susceptible to interviewer effects where the quality of responses may be affected by both the interviewers and the respondents characteristics. This could be an important problem when asking sensitive questions such as drug use, racial discrimination etc as well as nonsensitive questions affected by the social desirability of certain responses preferred by the respondents depending up on the race or gender of the interviewers. In addition to investigating the interviewer effects, the statistical inferences based on the face-to-face survey data must also incorporate these errors in the uncertainty assessment. In a large cohort study of 1,200 African-Americans from Pitt County, North Carolina, geographical segments were randomly assigned to interviewers based on race and gender. This paper will report on the findings about the differences by the respondent/interviewer characteristics on both sensitive and nonsensitive questions in the survey. The results to date suggest that the differences by respondent/interviewer characteristics could be substantial for sensitive questions; modest for certain nonsensitive questions and small or nonsignificant for other nonsensitive but "routine" questions. This paper will also discuss methods for incorporating the interviewer effects in the final analysis.

***Interviewer Judgments about the Quality of Telephone Interviews*, John Tarnai, Social & Economic Sciences Research Center; tarnai@wsu.edu, and M. Chris Paxson, Washington State University; cpaxson@wsu.edu**
Telephone interviewers are sometimes asked to judge the quality of interviews that they conduct, and to describe problems that they may have encountered during the interview. This kind of data is rarely reported, but is a core component of the kind of survey quality that is the theme of this conference. In this paper we present the results of asking interviewers to judge the quality of telephone interviews for 10 different telephone surveys conducted during the last several years. Interviewers rated each interview (excellent, good, fair, poor, inadequate), assessed whether respondents where interested in the survey topic, whether they had strong views on the topic, and the reasons that the quality of information obtained was less than excellent (language, hearing problems, interruptions or distractions, poor phone connection, competency, infirm, intoxication, respondent was rushed, respondent did not take interview seriously, R did not understand meaning of some questions, R was offended by interview, someone else was listening). Additionally, we present the results of an analysis that compares the data from higher quality versus lower quality interviews, with a goal of identifying the consequences of including less than excellent interviews in a survey dataset. We also compare these interviewer judgments of quality by type of respondent and survey topic. We discuss the implications these results have for interviewer training, and for design of telephone surveys.

***Interviewer Intervention for Data Quality in the 2004 Survey of Consumer Finances*, Yongyi Wang, NORC at the University of Chicago; wang-yongyi@norc.org, and Steven Pedlow, NORC/University of Chicago; pedlow-steven@norc.uchicago.edu**
The 2004 Survey of Consumer Finances (SCF), sponsored by the Federal Reserve Board (FRB), and fielded by the National Opinion Research Center (NORC), collects financial information from a national area probability sample of housing units, with an additional list sample from the same areas. As part of the continuous effort to improve data quality in SCF, an automatic SAS program has been created to examine two important aspects of data quality that were identified as measurable. The two quantities tracked are 1) the amount of comments left by interviewers, and 2) the percentage of financial data for which we receive only a Don't Know or Refusal answer. These two quantities are a first attempt to give immediate feedback to field interviewers in order to "nudge" them to improve their data quality, as necessary. Comments left by interviewers are crucial to understand the peculiarities and special circumstances of each interview, as well as to diagnose further improvements that could be made to the SCF questionnaire. Our simplistic approach is to count the number of characters of comments left by the interviewer. In order to minimize the amount of missing data, SCF 2004 has a series of questions designed to allow the interviewers to probe for a range response when the interviewer receives a "Don't Know" or "Refused" answer. In order to measure the quality of the financial information collected, our approach is to calculate the percentage of financial variables asked that yielded no actual range or value. An actionable report has been created on a weekly basis for each Field Manager (FM) that contains each "problem" among the FM's interviewers. This paper describes the specifications to create these FM Reports, presents a summary of the data quality effort, and reviews some learned lessons from the project.

MEASURING RACE AND ETHNICITY

***Can a Survey Change One's Race? An Experiment on Context Effects and Racial Self-Classification*, Thomas Craemer, Stony Brook University; tcraemer@ic.sunysb.edu**
A consensus exists in the social sciences that 'race' represents a social construct not a biological reality. This means that we should expect 'race' to change as the social context changes. The common practice in the social sciences to treat 'race' and 'ethnicity' as exogenous independent variables, however, belies this creed. This paper investigates the susceptibility of demographic racial and ethnic self identification to survey context effects. A sample of N=555 college students responded to two self-administered questionnaires about race related policy opinions and about their racial and ethnic backgrounds. One questionnaire was administered before a set of demographic and political frames were presented and the other one afterwards. The demographic frame referred to the success of genetic family tree research in identifying ancestors of previously unknown racial or ethnic backgrounds. The political frames discuss different affirmative action programs. Overall 11.7 percent of the participants changed their racial or ethnic self-identification in the course of the study. Ancestral ambivalence, both preexisting and induced, significantly increase the likelihood of changes in demographic self-identification. Political frames discussing affirmative action programs, on the other hand, reduce the likelihood of self-identification changes. These findings raise important questions about the placement of demographic questions within a survey. Their customary placement at the end of a survey may

increase the risk of survey induced context effects. The results further show that changes in racial or ethnic self-identification elicit political attitude changes regarding raced targeted policies such as affirmative action. This interplay between political context and demographic self identification raises questions about the general practice in political science to treat race and ethnicity as exogenous independent control variables. It suggests that they may serve future political science and survey research as important dependent variables as well.

***A comparison of single and multiple race category choices in NLSY97, round 6*, Parvati Krishnamurty, NORC; KRISHNAMURTY-PARVATI@norc.uchicago.edu, and Lidan Luo, NORC; luo-lidan@norcmail.uchicago.edu**

Inspired by Census 2000, the National Longitudinal Survey of Youth 1997 allowed respondents to select more than one race category in round 6. In one section of the questionnaire, respondents were asked to select a single race while in another they were allowed to choose more than one race. Responses to the two race questions were compared to address the following questions. What proportion of respondents are multiracial? Which race do they select if they can only choose one race? Does allowing for selection of multiple races give us better data on race? We also compared the data on race from proxy reports on the respondent's race obtained from the round 1 screener interview. Our results indicate that there was a low incidence of don't knows and refusals on both types of race questions in round 6. Allowing multiple race selection enabled 2.3% of respondents to report more than one race and many respondents who selected 'something else' in the single race question chose to select more than one race. However, a considerable proportion of respondents still refused to classify themselves except as Hispanic. The race chosen in the single and multiple race questions were consistent for those who chose only one race in the multiple race question. For those who selected multiple races, when 'White' and another race were selected, the other race was picked when they were constrained to pick only one race. Those who selected 'Black' or 'Asian' and another race were more likely to select 'Black' or 'Asian' in the single race question. The round 1 and 6 selections were consistent when only one race was selected for all groups except Hispanics. The round 1 and round 6 multiple race selection was consistent for Black and non-Hispanic non-Black categories. We explore various explanations for our findings.

***Measuring Ethnic Identity Within a Community-Based Sample of Black American Adults*, Rachel E. Davis, University of Michigan; reda@umich.edu, Abdul R. Shaikh, University of Michigan; arshaikh@umich.edu, Guangyu Zhang, University of Michigan; guangyuz@umich.edu, and Ken Resnicow, University of Michigan; kresnic@umich.edu**

Among racial and ethnic populations in the United States, Black Americans bear a disproportionate burden of cancer incidence and mortality. Many public health programs have sought to promote increased fruit and vegetable intake among Black Americans in an attempt to reduce cancer risk. These programs typically involve extensive efforts to develop culturally targeted materials. However, these efforts have generally failed to accommodate the heterogeneity of culturally affiliated behaviors and beliefs among Black Americans. This paper will discuss how the Eat for Life project is measuring ethnic identity in order to provide participants with ethnically tailored newsletters to increase fruit and vegetable consumption. Although several well-regarded measures exist to examine racial identity among Black Americans, none of these instruments has been utilized to categorize individuals for participation in public health interventions. These measures have been seldom used with community samples and have rarely, if ever, been administered over the telephone. In addition, pre-existing measures lack the capability to identify individuals with bicultural identities. This presentation will discuss how the Eat for Life program adapted existing items and created new items to develop a telephone-administered survey to measure Black ethnic identity among a community-based sample of Black American adults. This process involved many survey design as well as political challenges, particularly given the constraints of measuring Black ethnic identity among a sample of members from a health maintenance organization. Many steps were taken to ensure the acceptability of the survey, which was refined and tested through the use of four focus groups, multiple expert reviews, 16 qualitative pretests, and a psychometric pilot study with a community-based sample of 306 Black adults in Detroit. This paper will present the results of these pilot data, which will be used to construct a baseline questionnaire for a tailored, print intervention for 800 Black Americans.

***Black, White, Other? Why Racial Identifications within Hispanic Ethnicity Matter; A Look at Professional and Social Outcomes Among Research Doctorate Recipients*, Vincent Welch, Jr., NORC; Welch-vince@norc.org, and Kim Williams, NORC; Williams-Kim@NORC.org**

It is a well-established fact that different racial/ethnic groups within the United States have different levels of accomplishment in attaining professional success and high socio-economic status (SES). Indeed, this is one of the primary reasons for measuring race/ethnicity on surveys. However, the current reporting practice of combining all persons who indicate Hispanic ethnicity into the same racial/ethnic category on federal surveys may well be masking the important influence of race on attainment within the Hispanic ethnic group. Using the data from the Survey of Earned Doctorates and Survey of Doctorate Recipients we examine the unique relationship between race and professional and SES outcomes (i.e., salary, professional advancement, professional productivity, and professional and overall satisfaction) in a sample of highly educated U.S. citizen scientists and engineers. We hypothesize that the effect of race within the Hispanic ethnicity will closely match the effect of race on the larger population. That is, there will be an interaction between Hispanic ethnicity and race such that highly educated Hispanics, overall, will show diminished levels of professional and social attainment relative to their non-Hispanic White counterparts and within the Hispanic ethnicity Whites will show higher levels of attainment than their non-White counterparts. Implications of these findings on racial/ethnic reporting, sample design and the societal context in which these results occur will be discussed.

***New Research on the Differences Between Hispanic "Origin" and Hispanic "Identity" and Their Implications*, Paul J Lavrakas, Nielsen Media research; paul.lavrakas@nielsenmedia.com, Matt Courser, PIRE; mcourser@pire.org, and Lillian Diaz-Castillo, Westat; LillianDiaz-Castillo@westat.com**

As shown many times, small differences in the wording of questions can lead to large changes in the how that question is answered. The 1990 Census question wording measuring whether someone was Hispanic was: "Is this person of Spanish/Hispanic origin?" In the 2000 Census, it was: "Is this person Spanish/Hispanic/Latino?" Although the addition of "Latino" to the 2000 Census wording may have made it slightly more likely that someone would have said "Yes", dropping the word "origin" from the 2000 question wording most likely whether someone said "Yes." Lavrakas, Courser, and Diaz-Castillo (2002) reported that this small change in wording of the 2000 Census item appeared to lead to more 10 million people to not be counted as being Hispanic in the 2000 Census, compared to what would have resulted had the wording included the word "origin." This paper follow-ups of our 2002 paper, reporting the results from new studies that confirm/strengthen our earlier findings. A 2002 national mail survey (n=1,100) used an experimental design to test the 1990 and 2000 wording in measuring whether someone was Hispanic (one question version was randomly assigned to half the sample). This experiment found that respondents were more than twice as likely to say "No", when asked the 2000 version compared to the 1990 wording. The

second study was an RDD survey of the NYC area conducted in 204 (n=1,518), which found that 13% of respondents said they were Hispanic when asked the 2000 wording, but an additional 4% said "Yes" when asked about being Hispanic origin. These and other findings, including a multivariate profile of who is more likely to report having a "Hispanic Identity" vs. a "Hispanic Origin," will be presented and discussed, especially as related to how Census might operationalize the Hispanic construct in 2010.

POLITICAL ISSUES II

***Interpreting the Meaning of the Presidential Approval Question*, George Bishop, University of Cincinnati; george.bishop@uc.edu**
A cardinal assumption in asking any survey question is that it should mean the same thing to all respondents. A corollary is that when a survey question is repeated over time, it should mean the same thing the second time as it did the first time. If these vital assumptions cannot be met, valid comparisons across time and respondents become extremely difficult, if not impossible. The author argues that differences in how respondents interpret survey questions have become a significant source of systematic measurement error in public opinion polls. The research presented here focuses on the meaning-and-interpretation of a standard question on presidential approval: "Do you approve or disapprove of the way George W. Bush is handling his job as president?" The analysis draws upon data from: (1) polls conducted by the Gallup Organization,(2) Ohio Poll data, including responses to open-ended questions that provide indirect evidence on how respondents interpret the presidential approval question; and (3) pilot study data from a Cincinnati survey that directly probes what respondents mean when they say they approve or disapprove of the way George W. Bush is handling his job as president. The analysis demonstrates that the meaning of the standard presidential approval question varies across respondents and over time, calling into question a cardinal assumption of survey measurement. The paper concludes with a discussion of what can be done about this chronic source of measurement error.

***Thinking and Feeling the Partisan Gender Gap*, Barry Burden, Harvard University; burden@fas.harvard.edu**
I suggest that the gender gap in party identification is partly an artifact of question wording, which itself is incongruent with the original theory of partisanship. The Michigan model assumes that party identification is fundamentally affective or emotional. But surveys ask respondents to "think" when answering. A new survey experiment reanalyzes the gender gap by comparing the standard partisan battery to an alternative version where feelings rather than thoughts are primed. The experiment is implement in a telephone survey of 800 Ohioans and a Knowledge Networks survey of 2000 American adults. Bringing question wording into closer alignment with theory causes the gender gap essentially to shrink considerably. This happens because the "feel" questions find women to be less Democratic than did the "think" questions, all but erasing the gap between the sexes. Moreover, the disappearance of the gender gap occurs mostly among highly sophisticated women not those usually susceptible to question wording effects. Women also become more Independent in the "feel" condition, essentially eliminating the "independence gap." The results suggest that men and women are actually more, not less, alike when feelings are primed. Women are not simply providing more immediate, visceral responses to the affective questions as conventional wisdom might suggest. Taken together, the findings provide additional insights on why the gender gap emerged at all.

***Behavioral Self-report Measures of Non-political and Political Contributions*, Randall Thomas, Harris Interactive; rthomas@harrisinteractive.com, Jonathan Klein, University of Rochester; jklein@harrisinteractive.com, Mitchell Sanders, Harris Interactive; msanders@harrisinteractive.com, and Lisa Wilding, Harris Interactive; lwilding@harrisinteractive.com**
Results concerning self-report of behaviors have been found to be significantly affected by response format (Thomas, Behnke, Lafond, & Smith, 2003; Thomas, Klein, Behnke, 2004). In this study, we were interested in whether response format would also significantly affect self-report of non-political and political contributions. In a series of 7 waves of survey administration from October 2003 to October 2004, respondents were asked about contributions in the past 5 years using either a Yes-No grid or a Multiple Response format to the following entities: a political candidate, a political party, an organization involved in political issues, a non-profit non-political organization, and a religious organization (e.g., church, temple, mosque, etc.). We found that for all contribution entities, the Yes-No Grid led to higher reporting of contributions than did the Multiple Response format. We also found some intriguing differences in political versus non-political contribution entities across time.

***Political Knowledge Under Respondent-Friendly Conditions*, Markus Prior, Princeton University; mprior@princeton.edu, and Arthur Lupia, University of Michigan; lupia@isr.umich.edu**
When survey interviewers ask people to answer political knowledge questions, their response is often unimpressive. These findings lead researchers to doubt the quality of citizens' voting decisions. Such negative inferences, however, rest on the assumption that these responses measure what people know when they enter the voting booth accurately. This may not be the case. This paper examines if people perform poorly on political knowledge tasks not because they are incapable of answering the questions, but because they are caught unprepared and unmotivated to gather such information. We examine experimentally how changes in motivation and preparation time affect responses to otherwise standard political knowledge questions. Our approach is experimental. We vary two factors-time and reward-within a set of survey interviews. We give one randomly selected half of our sample only one minute to answer each question, whereas the other half has 24 hours to come up with the answers. Independently, we offer one randomly selected half of our sample a monetary reward for each correct answer. Since the variations are independent, each respondent is equally likely to be in one of four conditions: one minute no pay; one minute with pay, 24 hours no pay, 24 hours with pay. These variations are important because the standard way of measuring political knowledge is to place all respondents in the "one minute no pay" context. Our work will clarify the extent to which answers solicited in this context are robust to contexts where motivation and opportunity are greater.

***The Role of Opinion Polls in the 2004 Philippine Elections*, Mahar Mangahas, Social Weather Stations; mahar.mangahas@sws.org.ph, Linda Luz Guerrero, Social Weather Stations; guerrero@sws.org.ph, and Gerardo Sandoval, Social Weather Stations; jay.sandoval@sws.org.ph**
Opinion polls concerning the May 2004 Philippine Presidential election are essential to an understanding of recent Philippine political history. Incumbent Gloria Macapagal-Arroyo (GMA) made a surprise, and well-received, declaration in December 2002 that she would not run in 2004, but sacrificed much popularity by eagerly joining the; coalition of the willing; in the Iraq War in early 2003. In October 2003, GMA changed her mind about running for office, recruited poll topnotcher Noli de Castro as Vice-Presidential running-mate, and eventually caught up, poll-wise, with film star Fernando Poe Jr. (FPJ) who narrowly survived a legal challenge to his citizenship. Support for veteran senator Raul Roco fell dramatically upon disclosure of serious illness midway in the campaign. The integrity of the opinion polls was repeatedly assailed by another opposition candidate, Panfilo Lacson, whose refusal to give way allowed GMA a narrow

win over FPJ, by the smallest percentage margin (3.5%) since independence in 1946, with voters becoming polarized between North and South, rather than by socioeconomic class as in the 1998 election. Although the day-of-election poll, conducted in voters; homes, correctly predicted the candidates; ranks at the national level, it failed to do so in three of the countries 16 regions, including the politically-sensitive capital region, as heavy rains on election day caused unexpectedly high proportions of non-availability of sampled voters. At the national level, the exit poll understated FPJ-votes by 2% and those of the three other candidates by 1% each, consequently overstating GMA-votes by 5%. A scientific committee was tasked to conduct an independent review of the exit poll; it concluded that discrepancies were within sampling error, and found no evidence to support critics; charges that the exit poll had been sabotaged.

POLITICAL PARTY IDENTIFICATION

***Projection and Partisanship: The effects of party identification on opinion processing*, Kate Kenski, University of Pennsylvania; kkenski@asc.upenn.edu, and Dannagal Young, University of Pennsylvania; dyoung@asc.upenn.edu**
Using data from the National Annenberg Election Survey (NAES) 2004, this project tests two competing explanations of the candidate favorability and issue position relationship: policy-based evaluations and the projection hypothesis. While policy-based evaluations would require that an individual gauge the distance between herself and a candidate prior to a judgment of candidate favorability, the projection hypothesis "proposes that democratic citizens systematically distort their perceptions of competing candidates' positions on controversial issues of public policy" (Krosnick, 2002, p.116). Both of these hypotheses focus on three variables: candidate favorability ratings, respondents' issue positions, and respondents' perceptions of the candidates' issue positions. In this study, the causal direction of the relationship between these variables is investigated with panel data collected around the 2004 Republican national convention (n=807). The conditional effects of party identification in relationship to issue-based evaluation and projection are explored.

***Attitude or Demographic? Measurement Effect on Party Identification Estimates in Preelection and Public Policy Polls*, Robert Daves, Star Tribune; daves@startribune.com**
Each year pollsters take it on the chin from partisans who try to find fault with preelection polls' "internals" or the methodology. Partisans - Democrats and Republicans - criticized state and national polls during the 2004 election for the same reason: the perception that there were too many, or too few in any given poll. In Minnesota, that perception of bias resulted in one party calling for the dismissal of the newspaper's in-house pollster. Their claim? The polls included too many Democrats or too few Republicans. This paper builds on earlier papers presented at AAPOR dealing with party ID measures. One of those papers (Daves, 2000) examined the effect of a new, then-powerful third party in the state, and how different party identification measures affected party ID estimates. Because of election angst about party identification, the Minnesota Poll conducted a split-ballot experiment among likely voters in October, 2004, to determine the effect of party ID measures on candidate support. About half the respondents selected randomly heard the party ID question as the first question on the questionnaire, followed by the trial-heat question, then were asked party ID at the end of the questionnaire to test changes in identification. Party ID was measured only at the end of the questionnaire for the remainder. Initial results showed little difference between party ID measures on the two ballots. Although there was little difference in candidate support in either ballot, there was a hint that candidate choice affected party ID among a few voters, especially among weak partisans and independents. This paper explores those findings in detail. The paper also will report the findings of a split-ballot experiment planned for a January poll to help understand party identification measure placement in the questionnaire and its relationship to presidential job approval in a nonelection political environment.

***Party Identification and the 2004 Election*, Larry Hugick, Princeton Survey Research Associates; larry.hugick@psra.com, and Stacy DiAngelo, Princeton Survey Research Associates; Stacy.DiAngelo@psra.com**
The party identification distributions in national pre-election surveys came under unusual scrutiny in election year 2004. During the fall campaign, poll critics who maintained that party ID should be treated like a demographic constant highlighted deviation from the "correct" party ID distribution as evidence of flawed methodology - leading to charges that the CNN/USA Today/Gallup poll, the NEWSWEEK poll, and other major media polls were biased toward the Republicans. In the end, however, the national exit poll results showed a more Republican electorate than the two previous presidential elections, strongly supporting the view that party ID is a fluid, not a fixed, characteristic. Looking back, polling data for election year 2004 provides an opportunity to better understand short-term variation in party ID and look for indications of more lasting movement in party allegiances. Primarily based on data from the NEWSWEEK poll conducted by Princeton Survey Research, this paper examines movement in party ID at various points in the election cycle, including the primary election period, around the political conventions, and the fall campaign in the context of polling data from 2000, 1996, and earlier presidential election years. It also compares the shift toward the GOP in the fall of 2004 with the movement toward the Republicans observed in the months immediately following the 9/11 terrorist attacks. National exit poll data as well as data from The Pew Research Center and other sources will be used to supplement the NEWSWEEK poll data.

***Weighting by Party ID: When It Comes to Elections, Can It Really Be Tossed Out the Door?*, Tatiana Koudinova, TIPP Poll/Technometrica Market Intelligence; tkoudinova@technometrica.com, and Raghavan Mayur, TIPP; mayur@techometrica.com**
TIPP, the political polling arm of TechnoMetrica Market Intelligence, conducts monthly polls for its media partners, the Investors Business Daily and the Christian Science Monitor. Since the 2000 Presidential Election, TIPP has adhered to rigorous polling standards and has amassed a detailed database of party affiliation information spanning more than 4 years. In the recent 2004 Presidential Election, when many other pollsters were saying that weighting by party affiliation was an inappropriate method, TIPP continued to adhere to the philosophy that weighting by party ID was the only way to accurately represent reality. The result: TIPP correctly predicted President Bush's reelection and was one of the closest to predicting his actual margin of victory. In this presentation, we will review existing publications and present an analysis that endeavors to uncover the dynamics and consequences of party stability, taking both a short and long-term view of the data. Demographic variables will be classified into two major groups: 1. Groups wherein party ID is stable over time. 2. Groups wherein party ID may experience volatility. The motive force behind this study is three-pronged: 1. Question conventional AAPOR wisdom that holds - as a central tenet of political polling - that party ID is not a stable indicator and should therefore not be used as a basis for weighting pre-election polls. 2. Analyze the stability of party affiliation at three different levels: a. The voting age population and, b. Registered voters c. Likely voters 3. Test different weighting methods using polling data from the 2004 and 2000 presidential

election. An evaluation of the different procedures for both efficacy and robustness will be carried out using data sourced from the informational storehouses of TIPP and Pew.

COGNITION AND HEALTH

***Using Screening Questions to Identify Persons with Mobility Impairment: Field Test Results*, Vickie Stringfellow, Center for Survey Research, University of MA Boston; vickie.stringfellow@umb.edu, and Patricia Gallagher, Center for Survey Research; patricia.gallagher@umb.edu**

As part of the ongoing development of the Consumer Assessment of Health Plans Survey (CAHPS®), researchers noted the need for a set of screening questions to identify persons with mobility impairment (PWMI). Although there is an established set of screeners to identify adults with chronic health conditions, previous research suggests that PWMI are a distinct subgroup of this population. Health plans in both the private and public sectors have an interest in learning more about the experiences of PWMI in order to improve the services they provide to this group. The first step toward achieving this goal was the development and testing of a series of eleven screening questions. The screening questions were tested tested on a sample of 1124 adults enrolled in the Massachusetts Medicaid program (MassHealth). Sampled adults were sent a questionnaire and fact sheet that answered some commonly asked questions about the instrument. Reminder postcards were sent to everyone two weeks later, and a second mailing was sent to nonrespondents about two weeks after that. In accordance with the CAHPS® data collection protocol, professional interviewers then attempted to complete the questionnaire with nonrespondents by telephone. A total of 564 questionnaires were completed for a response rate of 51% (RR1, AAPOR 2004). In order to validate and refine the new PWMI screening questions, medical claims data including ICD-9 codes were obtained for the survey sample. Analyses of the data from both the questionnaire and medical claims allow us to address several research questions: Do the new screening questions appropriately identify PWMI? Is this group truly different from the group identified by the chronic condition screener? Can the 11-item series be shortened to reduce respondent burden without degrading its effectiveness? Can the same PWMI be identified by using claims data alone, negating the need for screening questions?

***Qualified answers and other doubt expressions as indicators of cognitive problems in a health survey*, stasja draisma, Free University Amsterdam, dpt of Social Research Methodology; ar.draisma@fsw.vu.nl, Wil Dijkstra, Vrije Universiteit; w.dijkstra@fsw.vu.nl, and Yfke Ongena, Vrije Universiteit Amsterdam; YP.Ongena@fsw.vu.nl**

Questions that pose problems to respondents can result in verbal and paralinguistic doubt expressions (e.g. filled pauses, speech rate et cetera). Such doubt expressions indicate inaccuracy and uncertainty, shown earlier in a validity study of Draisma & Dijkstra (2004). For the present paper, verbal interactions in CATI health interviews are analyzed, to identify specific cognitive problems in answering health questions. Two different verbal doubt categories are discerned: on the one hand hedges like 'I think', 'I believe', on the other hand qualified answers that signify necessary cognitive elaborations (think aloud) to arrive at an answer. A qualified answer may be: 'No, I don't smoke. It is impossible with my lung problems.' Paralinguistics, hedges and qualifications may indicate specific problems related to cognitively different answer phases. For instance, Schober (2004) used hedges as indicators for misinterpretation of question concepts, but hedges also denote memory and judgment difficulties in our health survey. In the survey, four different question types occur: frequency estimates of (un)healthy behavior, own health perceptions, opinions about responsibility for health-problems and factual questions concerning own (un)healthy behavior. A preliminary analysis showed that frequency estimate answers were much more often preceded than followed by verbal qualifications. This indicates that judgmental problems in answer construction form a larger problem than justification afterwards, as could also be concluded from an analysis of the content of the qualifications. Furthermore, opinion questions led to more and different verbal qualifications than behavioral questions. The paper will demonstrate how analysis of qualified answers, hedges and paralinguistics can be used to detect cognitive problems with health questions and how it can be used in the early pilot stages of large scale surveys, to eventually enhance survey quality.

***Identifying mismatches between common sense and technical definitions on an HIV risk behavior survey*, Lisa Moses, NCHS; lisamoses@gmail.com, and Paul Beatty, Center for Disease Control and Prevention; pbb5@cdc.gov**

Researchers often erroneously assume that terms utilized in survey questions will be universally interpreted. Periodically, terms researchers consider to be common sense take on diverse meanings among members of varying populations, as was the case with a survey by the National Center for HIV, STD, and TB Prevention (NCHSTP) addressing HIV risk behaviors. A major risk of contracting HIV is through sexual contact. It has been suggested that safe sex behaviors differ based on characteristics of relationship of the partners, i.e., length of relationship, knowledge of partner's sexual history, and commitment level. Researchers at NCHSTP theorized that sexual relationships can be broken down into two categories: "steady partner" and "non-steady partner," and that these terms would be universally agreed upon. Researchers at the National Center for Health Statistics (NCHS) utilized cognitive interviewing techniques in an attempt to determine whether questions utilizing NCHSTP's definitions of these categories lined up with individuals' conceptualizations of these terms. This was accomplished by exploring in depth responses to assess problems with the intention and comprehension of the terms. This presentation will focus on the methodology used for assessing these questions among intravenous drug using (IDU) males, IDU females, and men who have sex with men (MSMs), as well as findings and suggestions.

***Applying Cognitive Psychological Principles to the Improvement of Survey Data: A Case Study from the National Survey on Drug Use and Health*, Joel Kennet, Substance Abuse and Mental Health Services Administration; joel.kennet@samhsa.hhs.gov, Dicy Painter, Substance Abuse and Mental Health Services Adminis; dicy.painter@samhsa.hhs.gov, Peggy Barker, Substance Abuse and Mental Health Services Adminis; peggy.barker@samhsa.hhs.gov, and Jeremy Aldworth, RTI International; jaldworth@rti.org**

The National Survey on Drug Use and Health (NSDUH) collects data on Medicare and Medicaid enrollment status as part of a general interview conducted after the core drug use measures have been administered. While the overall prevalence estimates derived from the NSDUH Medicare and Medicaid enrollment questions have generally appeared credible, it became apparent that among people under 65 years old, Medicare enrollment was overreported and Medicaid enrollment was underreported. Among people over age 65, Medicaid enrollment appeared to be highly overreported. These judgments were based on "eyeball" comparisons with estimates from the SIPP and CPS, both of which administered highly detailed modules on health insurance coverage. Expert review of the NSDUH question wordings suggested that inadequate establishment of context (defining terms after using them in the questions), and other syntactic difficulties created excessive demands on working memory. Correction of these problems in time for the 2003 NSDUH resulted in age-group coverage estimates that more closely matched those obtained in the other surveys, which targeted this topic more specifically and in greater depth. These results are discussed within the context of Tourangeau, et al's (2000) Response Process Model.

THE GOVERNMENT AND SURVEYS

***Making Health Policy at the Ballot-Box: Public Attitudes Toward Five Health Propositions and the Future of Health Policy*, Kristy Michaud, Public Policy Institute of California; michaud@ppic.org, Renatta DeFever, Public Policy Institute of California; defever@ppic.org, and Jon Cohen, ABC News; jon.cohen@abc.com**

This paper explores the California public's views on making and changing health policy at the ballot box, voter attitudes about five state health policy propositions, and the opportunity for future health policy. In November 2004, California's ballot included four citizens' initiatives and one referendum that focused directly on major healthcare issues: employer mandates, children's hospitals, emergency medicine, mental health, and stem cell research. The high number of major health policy propositions provided a unique opportunity to assess the relative impact of voters' healthcare experiences and outlook, coverage status, and a range of socioeconomic factors on their approach to ballot box policymaking on health policy issues. In this paper, we report results from an RDD survey of 2,500 California voters conducted in five languages (English, Spanish, Chinese, Korean, and Vietnamese) and from a series of statewide focus groups, both of which we designed to reveal the motivations of Californians' voting choices and their expectations toward health policy and ballot box decision making more broadly. Specifically, this paper investigates Californians' base levels on knowledge and personal concerns about health coverage and health policy issues, their attitudes towards making health policy decisions at the ballot box, the perceived impacts of the five ballot measures on level of importance of health policy issues, their overall awareness of the five ballot measures, and their individual voting decisions. Overall, these data confirm Californians' commitment to the citizens' initiative process, their distrust in elected officials, and their abiding interest in direct ballot box control over the state's public policy. These data also reveal opportunities for fundamental change in health research and the delivery of healthcare, including but not limited to Californians' passage of the stem cell research bonds.

***A quantitative and qualitative approach to gauging public confidence in official statistics in a British environment.*, Amanda Wilmot, Office for National Statistics; amanda.wilmot@ons.gov.uk, and Jacqui Jones, Office for National Statistics, UK; jacqui.jones@ons.gov.uk**

Authors: Amanda Wilmot and Jacqui Jones Abstract One of the key objectives of National Statistics is to improve public confidence in official statistics by demonstrating that they are produced to best professional standards and free from political interference. In order to do this a greater understanding of the nature of 'trust'and public confidence in statistics was required. The Office for National Statistics and the Statistics Commission, UK, are working together on different aspects intended to identify the key issues relating to this subject. Various strands of the project combine both quantitative and qualitative methods of inquiry to inform the debate. This paper discusses: i) the conceptualisation, development and operationalisation of a structured questionnaire module administered to the general public via the National Statistics Omnibus survey; and, ii) a qualitative exploration of the general public's views of official statistics and how they formulate these views, principally captured through a series of focus groups. These two strands ran along-side each other, informing, supporting and complimenting the findings from both.

***Elements behind trust in Government, An explanatory model*, Rene Bautista, Gallup Research Center- University of Nebraska-Lincoln; rbautis1@bigred.unl.edu, and David Palmer, Gallup Research Center; dpalmer1@bigred.unl.edu**

Several surveys conducted in the US since the 1960's have shown a decreasing trend in political trust. Implications of this trend include lower levels of political participation and voting, and limited support for public policies. Some authors have argued that the democratic design usually gives rise to disagreement, which in turn gives rise to political distrust. On the other hand, it has been argued that politicians' performance is the main source of political distrust. At the heart of this debate is what elements compose the construct of political trust. In this debate the political trust construct has been considered an independent variable in political models. However, in order to better understand political trust's role as an independent variable in these models, we should investigate it is as a dependent variable. This paper addresses political trust as a dependent variable and its meaning in political surveys. It is imperative to distinguish between two theoretical elements: specific support (satisfaction with authorities' performance) and diffuse support (perception of the political system). By means of linear regression analysis, several components of trust in government are modeled. Preliminary results suggest that trust in government is mainly a measure of diffuse (institutional) rather than specific (personal) support. In this regard, our statistical analysis suggests an increase of trust in government when people approve Congress' performance, they have a "warm" position on institutional thermometers (Supreme Court, Congress and Presidency), and when people are satisfied with the democratic system. Additionally, trust in government increases, though to a lesser degree, when people have a positive perception of the President's traits. In addition, our model suggests that elements such as personal and national economy perceptions have no significant impact on trust in government. The dataset used for this analysis comes from the American National Election Study 2002.

***Neighbourhoods Matter: Fixes and Random Effects in Police Satisfaction Surveys in the UK*, Rachel Dinkes, AIR; rdinkes@air.org**

INDEX

D

E

F

L

M

N

O

P

Q

R

S

T

U

V

W

X

Y

Z

No single data collection mode matches the methodology of every research project. Should the project be fielded on the Internet, phone or both? Western Wats doesn't limit you to just one survey mode. Our proprietary multi-mode data collection system, W.I.R.E. (Wats Integrated Research Engine), is used by more than 250 clients worldwide to create competitive advantage resulting from thoughtful mode selection.

Using the right tools can mean the difference between a successful project and one resulting in a sore thumb. Call us today and start using the data collection solutions that are right for you.

(801)373-7735

Western Wats
Data collection you can trust
Partners you can rely on

www.westernwats.com

SurveySpot is the online panel you can trust. Our well managed, proprietary panel includes more than 3.3 million households, representing over 8 million household members worldwide. Enhanced specialty selections allow you to easily reach your target markets.

SSI is the sampling company you can trust. SSI has cultivated the most dependable, diverse line of sampling solutions in the industry because sampling is our only focus – and has been for nearly 30 years.

- **SurveySpot Panel** — USA, Canada, UK, Australia, and Asia
- **Global Panels** — 30 countries through Bloomerce, an SSI company
- **B2B Sample** — Internet and Telephone
- **Global RDD Telephone Sample** — 21 countries
- **SSI-SNAP** — 24/7/365 online ordering system

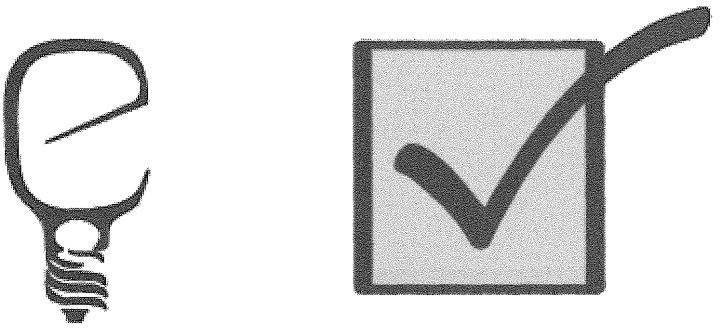
Online or telephone, business or consumer, SSI is the most reliable resource for your sampling needs the world over. Contact your SSI account team by email at info@surveysampling.com or by phone at 1.203.255.4200.

Survey Sampling International

surveyspot

Phone: 1.203.255.4200 ▮ E-Mail: info@surveysampling.com ▮ Web: www.surveysampling.com

Fairfield ■ London ■ Paris ■ Rotterdam ■ Sydney ■ Tokyo



Edison/Mitofsky

www.exit-poll.net

Exit polls from Edison Media Research and Mitofsky International: The exclusive source for election day information.

We wish **AAPOR** a most successful 2005 conference

Edison Media Research
www.edisonreserch.com
908 707-4707

Mitofsky International
www.mitofskyinternational.com
212 980-3031

Special Offer to

AAPOR Attendees!

Get a free ASA publication by joining the American Statistical Association during this special promotion!

Enjoy a year of ASA membership and your choice of:

Journal of the American Statistical Association (JASA)—*JASA* is the preeminent source of statistical knowledge in the economic, social, physical, engineering, and health sciences. The peer-reviewed journal emphasizes real-world applications. Readers benefit from new applications, revealing case studies, new/useful data sets, and practical indicators. Published quarterly by the ASA. **OR...**

The American Statistician (TAS)—Statistical educators, practitioners, researchers, and others interested in the impact of statistics will enjoy *TAS*. Subscribers benefit from challenging discussions about applications and problems in the field and form resources and reviews for teachers at every level. Published quarterly by the ASA.

Join today to enhance your statistical knowledge right away!

www.amstat.org/join

ASA MEMBERS ENJOY:

- *Amstat News*, the monthly membership magazine of the ASA, and *ASA Member News*, our monthly electronic newsletter.
- Members Only discounts on all ASA publications, meetings, and products.
- Career-enhancing opportunities through the JSM Career Placement Service, *Amstat News*, and online JobWeb postings.
- Free web subscription to the *Current Index of Statistics (CIS)*.

YES! I would like to join ASA for $85 and get a free subscription to *(please circle one publication)*: ***JASA*** *or* ***TAS***

Name _____ Organization _____

Address _____

City _____ State/Province _____ Zip/Postal Code _____ Country _____

Phone _____ Email _____

❑ Check/money order payable to American Statistical Association (in U.S. dollars drawn on U.S. bank)

Credit Card: ❑ VISA ❑ MasterCard ❑ American Express

Card Number _____ CVS# (3 digit # on reverse of card) _____ Exp. Date _____

Name of Cardholder _____

Authorizing Signature _____

AAPOR05

MAIL: American Statistical Association, Dept. 79081, Baltimore, MD 21279-0081
FAX: (410) 626-7509 **CALL:** 1 (888) 231-3473

ASDE
SURVEY SAMPLER

729 St-Joseph, Suite 201
Gatineau (Hull), Qc
J8Y 4B6

Phone 1-888-323-3651
Email info@surveysampler.com
Web www.surveysampler.com

Look no further.

For faster service, quality data and competitive pricing call Randa Bell today at 888-323-3651.

- **Phone Samples for Canada and the USA**
  - Sampling software or on demand
  - Census overlay
  - Business/medical lists
  - Listed or RDD samples
  - Custom geographies
- **List Matching**
  - manual or automated phone number look up from any list
- **IVR In-bound and out-bound**
- **Pre-Dialing of phone numbers to identify line status**
- **Geocoding of samples using various official sources**
- **National 'Do Not Call' cooperative**

ASDE is certified ISO 9001:2000 for sampling and related services for survey professionals

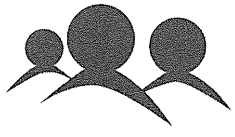
Brand bland?
Logo a no-go?

Use Blackstone's ideation tools, including our **high-creative classifier** and **Crazy Salad**, to make over an existing brand from dud to stud.

Test value propositions, positions, and logos with **Chimeric**, a hybrid that lets you get to the heart of your brand without losing your head.

Mikey likes it, but does Miguel – or Michelle? In domestic, multicultural, and global markets, **Persona IPB** finds the segments where your brand is champion.

Let's talk brand building. Contact Ashref Hashim or Kathi Rose:



BLACKSTONE GROUP
research insights for strategic advantage

The Blackstone Group, Inc.
www.bgglobal.com info@bgglobal.com 312.419.0400
360 N. Michigan Avenue, Suite 1500, Chicago, IL 60601

California Survey Research Services, Inc.

- Telephone Survey Research (CATI)
- Bi-lingual Spanish/English capabilities
- Data Processing Services
- Mail Surveys
- Internet Surveys

For 24 years, CSRS has provided quality research services to Academic, Commercial, and Government clients.

Founded by a U.S. Census Bureau statistician, CSRS provides data collection and data processing services for clients who require skilled execution of complex studies.

Recent clients include:

USC Annenberg Center for the Digital Future **– Internet Project: Surveying the Digital Future (2000-2005)**
www.digitalcenter.org

USC Annenberg School for Communication **– Metamorphosis: Transforming the Ties that Bind (1998-2005)**
www.metamorph.org

UCLA Jonsson Cancer Center **– California Asian Indian Tobacco Study (2004)**
www.sscnet.ucla.edu/aasc/caits

UCLA Department of Medicine **– Communications in Medical Care (1999–2005)**

U.S. Department of Veterans Affairs **– Well Being Among Veterans Enhancement Study (2003-2005)**

Veterans Administration of Greater Los Angeles **– Women Veteran's Ambulatory Care Use Project (2003-2004)**

Jewish Community Federation of San Francisco **– Community Study (2004)**

Ken Gross, President
California Survey Research Services, Inc.
15350 Sherman Way, Suite 480
Van Nuys, CA 91406
818-780-2777
KGross@CalSurvey.com
www.CalSurvey.com

NEW TITLES IN SOCIAL MEASUREMENT

DEMOGRAPHY: ANALYSIS AND SYNTHESIS

FOUR-VOLUME SET

Editors-in-Chief

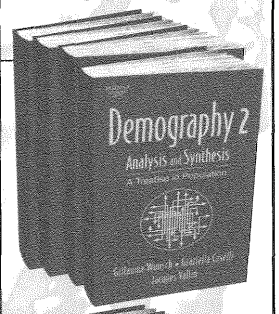
Gillaume Wunsch
University of Louvain, Belgium

Graziella Caselli
Università degli Studi di Roma, Italy

Jacques Vallin
Institut National d'Etudes Démographiques, Paris, France

This four-volume collection of 120 original chapters covers virtually everything of interest to demographers, sociologists, and others. Over 100 authors present population subjects in ways that provoke thinking and lead to the creation of new perspectives, not just facts and equations to be memorized. The articles follow a theory-methods-applications approach and so offer a kind of "one-stop shop" that is well suited for students and professors who need non-technical summaries, such as political scientists, public affairs specialists, and others.

November 2005, c. 3,500 pp., ISBN: 0-12-765660-X
Introductory price: $800.00
List print price: $995.00
*Introductory price valid for three months after month of publication.



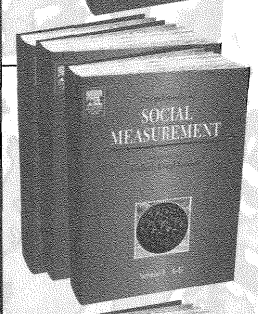
AVAILABLE BOTH IN PRINT AND ONLINE!

ENCYCLOPEDIA OF SOCIAL MEASUREMENT

THREE-VOLUME SET

Editor-in-Chief

Kimberly Kempf-Leonard
University of Texas at Dallas, U.S.A.



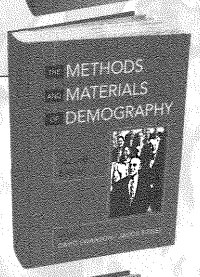
"...the Encyclopedia of Social Measurement will be highly valued by general readers, beginning social science researchers, and experienced practitioners alike."

—PETER MARSDEN, HARVARD UNIVERSITY

"The enyclopedia... provides fully developed chapters—by recognized authorities—that easily replace a substantial library of sociological methodology books"

—KARL VAN METER, CENTRE NATIONAL DE LA RESHERCHE SCIENTIFIQUE

2005, 3,008 pp., ISBN: 0-12-443890-3
List print price: $950.00
For online pricing, please visit: http://books.elsevier.com/measurement



THE METHODS AND MATERIALS OF DEMOGRAPHY

SECOND EDITION

Jacob S. Siegel
J. Stuart Siegel Demographic Services, Bethesda, Maryland, U.S.A.

David Swanson
Helsinki School of Economics and Business Administration, Finland

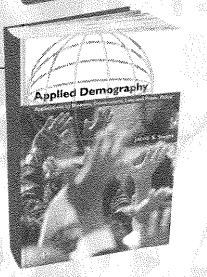
"All demographers and demography students will have to have this book, which is current, timely, accessible, and chock full of all the demographic methods, data sources, and references needed to do demography."

—DUDLEY L. POSTON, JR., TEXAS A&M UNIVERSITY, U.S.A.

"This book is an indispensable reference tool for professional and academic scholars in population studies... It is essential reading for demographers as well as geographers, economists, sociologists, and health scientists."

—FILOMENA RACIOPPI, DIPARTIMENTO DI SCIENZE DEMOGRAFICHE, UNIVERSITA DEGLI STUDI DI ROMA "LA SAPIENZA", ITALY

2004, Casebound, 819 pp., $89.95, ISBN: 0-12-641955-8



APPLIED DEMOGRAPHY

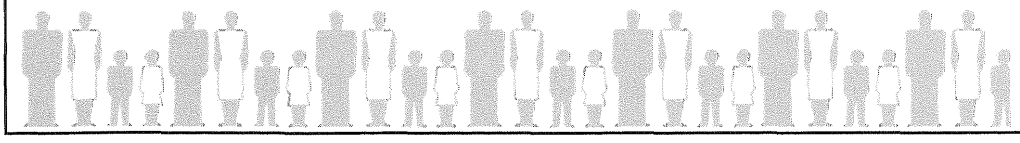
Applications to Business, Government, Law, and Public Policy

Jacob S. Siegel
J. Stuart Siegel Demographic Services, Bethesda, Maryland, U.S.A.

"Siegel has delivered a winner (again!). Intended for the demographic practitioner and those who seriously wish to engage the field of applied demography, this book is The Userguide *to the demographer's toolkit. The scope and breadth of the concrete examples is awesome!"*

—PAUL R. VOSS, UNIVERSITY OF WISCONSIN-MADISON, U.S.A.

2002, Paperback, 686 pp., $84.95, ISBN: 0-12-641840-3



Order directly from:

In the U.S. and Canada:
ELSEVIER
Toll Free: 1-800-545-2522
Tel: 1-314-579-3300
Fax: 1-800-535-9935
or 1-314-523-5940
email: usbkinfo@elsevier.com
http://books.elsevier.com

All other countries:
ELSEVIER
Tel: +44 1865 474110
Fax: +44 1865 474111
email: eurobkinfo@elsevier.com
http://books.elsevier.com

ELSEVIER
ACADEMIC
PRESS

GROUP DIMENSIONS INTERNATIONAL

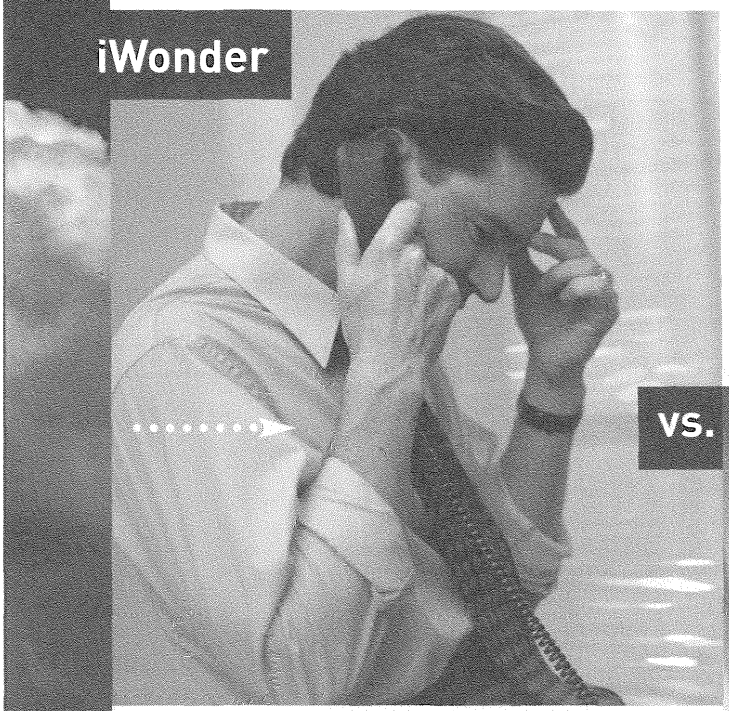
Research, Facilitation, and Training for Group Effectiveness

With a distinguished client list and a global reach, GDI is a recognized leader in social and policy research...

- **Quality Research**
  - Focus Groups and Surveys
  - Program Evaluations
  - Needs Assessments
  - Structured Transcripts ™
- **Expert Facilitation**
  - Strategic Planning
  - Meeting and Retreat Facilitation
- **Practical On-Site Training**
  - "Conducting Professional Focus Groups"
  - "Making Sense of SPSS and Surveys: From Design to Analysis"
  - "TeamWork: Meeting Facilitation, Communication, and Conflict Management"
- **Comprehensive Training Manuals**
  - *Making Sense of SPSS and Surveys: From Design to Analysis (2005)*
  - *The Power of Focus Groups for Social, Policy, and Market Research (4th ed., 2004)*
  - *The Power of Focus Groups: Focus on Health (2004)*
  - *The Power of Focus Groups: Focus on Education (2004)*
  - *The Power of Focus Groups: Focus on International Development (2003)*
  - *El Poder de los Grupos Focales: Un Manual de Entrenamiento para Investigación Social y de Políticas Públicas, e investigación de mercados (2004)*

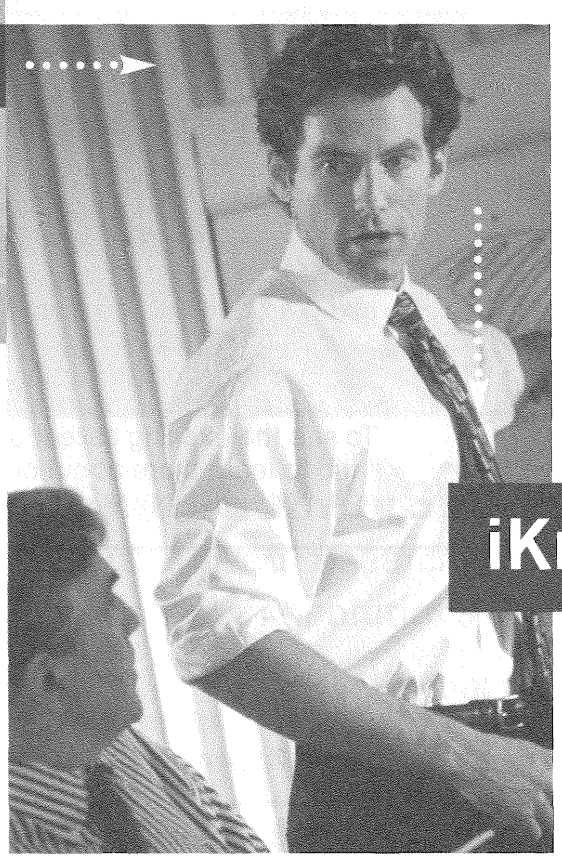
Janet Mancini Billson, PhD, CCS, Director
Norman T. London, EdD, Co-Director

300 Narragansett Avenue, Barrington, Rhode Island 02806
401-246-0797 401-465-6004
jmbillson@earthlink.net
www.focusgroupdimensions.com



At last there's a way to resolve the uncertainty behind your quantitative data...

VS.



iModerate.

Real-time, one-on-one interviews conducted by highly trained moderators within your online survey, so you no longer have to wonder what your respondents meant when they said something — **you can ask them.**

iModerate℠
research technologies

Broaden Your Expectations

Visit our booth. AAPOR Annual Conference www.iModerate.com

Market knowledge at your fingertips

Snap Survey Software

Design, process and analyze your entire survey

» Paper
» Scanning
» Phone
» Web
» E-mail
» PDA

First class marketing surveys have never been easier

Snap SurveyShop

Customized research and data outsourcing solutions

» Design
» Analysis
» Reporting
» Web-hosting
» Data Processing
» Technical Solutions

snap® surveys

Phone: 603-610-8700

Web: www.snapsurveys.com

1. How accurately do the following statements describe you?

	Very Inaccurately				Very Accurately
You want to focus your resources on content, sampling, and analysis	☐	☐	☐	☐	☑
Your time and resources are valuable	☐	☐	☐	☐	☑
You would like to eliminate administrative hassles	☐	☐	☐	☐	☑
Paper-based surveys are important to your organization	☐	☐	☐	☐	☑
Your organization would benefit from reduced mailing costs	☐	☐	☐	☐	☑
Cost savings are important to your organization	☐	☐	☐	☐	☑
You need an experienced staff to assist you with your printing and distribution	☐	☐	☐	☐	☑
Your organization would benefit from our expert staff processing and coding your data	☐	☐	☐	☐	☑
You would feel more comfortable knowing your data capture is guaranteed accurate	☐	☐	☐	☐	☑

NPC's Survey Processing division builds custom data capture and collection solutions for organizations like yours. Our services include...

- Form design and layout
- Printing, inserting, and low-cost mailing
- High-volume ICR, OCR, OMR capture and key entry capabilities
- Data capture, coding, and information management
- Customized data formatting, including all popular database and spreadsheet formats, CSV, delimited text, XML and SPSS
- Image archiving

NPC®

For more information about NPC's survey processing services, please contact:

Bob Beiter
Director of Government Services
NPC
1-814-239-8787 ext. 1277
bob.beiter@npcweb.com

AAPOR has a new Logo...

AAPOR

AMERICAN ASSOCIATION FOR PUBLIC OPINION RESEARCH

to celebrate we will be offering a few great products so you can show your AAPOR pride.

The following products will be offered:

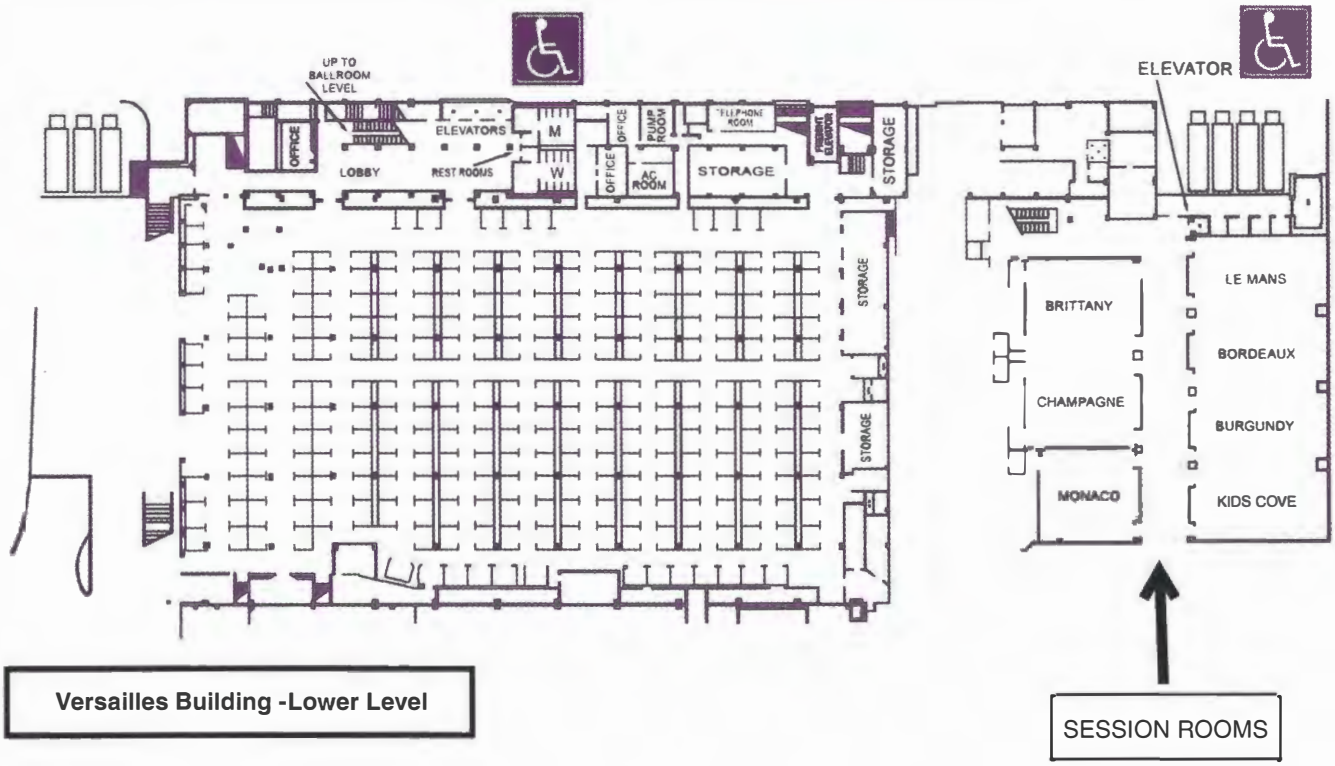
Mouse Pad - $5.00
Coffee Mug - $5.00
Men's Golf Shirts - $20.00 (light green and tan) Sizes: S, M, L, XL, XXL
Women's fitted tee - $15.00 (white and black) Size: S, M, L, XL, XXL

All items will have the new logo in "AAPOR green". Order forms will be included in each attendee's registration packet and items will be available for purchase beginning Friday afternoon. All sizes and colors will be made available at the conference.

NO SHIPPING COST

when you order at the conference

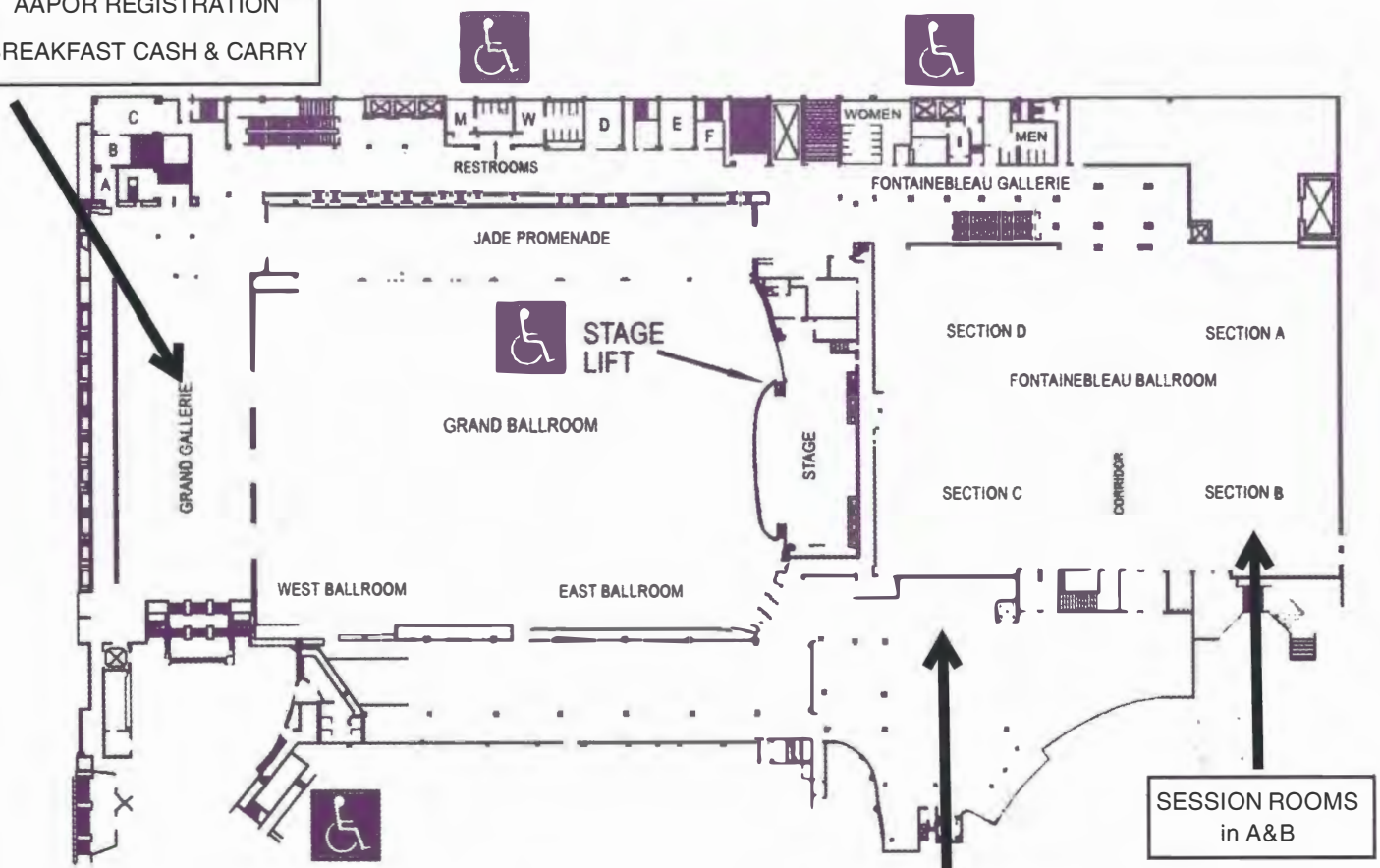
NOTES



Versailles Building -Lower Level

SESSION ROOMS

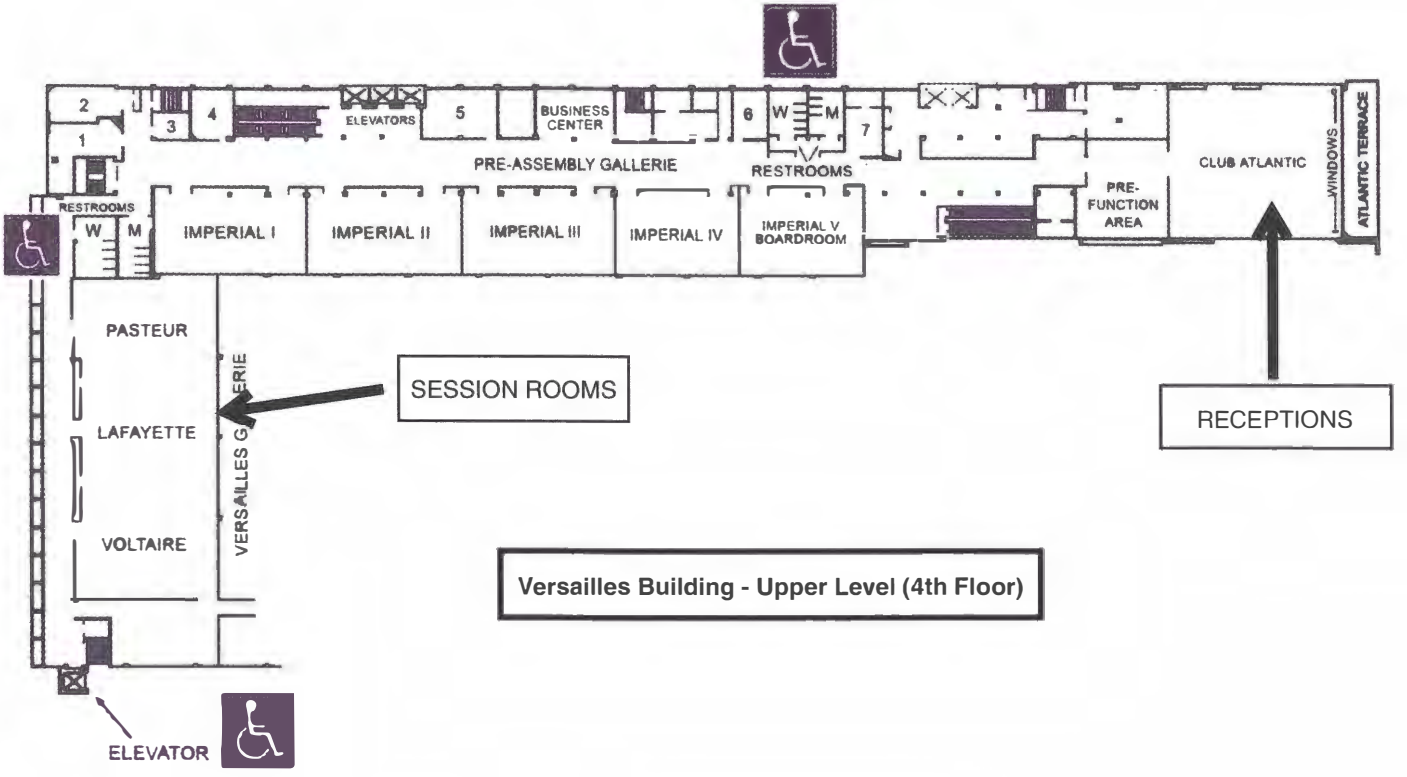
AAPOR REGISTRATION
BREAKFAST CASH & CARRY



Versailles Building Main Level

SESSION ROOMS in A&B

EXHIBITS, BREAKS, BOOK SALE, POSTER & MTA sessi onsi nC&D



Versailles Building - Upper Level (4th Floor)