American National Election Studies (ANES)
http://www.electionstudies.org

Board Report

Future Considerations for Instrumentation and Measurement on the ANES

James Druckman (Northwestern University), chair
Shanto Iyengar (Stanford University)
Melissa Michelson (Menlo College)
Stephen Nicholson (University of California, Merced)
Randy Stevenson (Rice University)
Joshua Tucker (New York University)

April 3, 2017

April 3, 2017

Future Considerations for Instrumentation and Measurement on the ANES

Committee: James Druckman (Northwestern University), Shanto Iyengar (Stanford University), Melissa Michelson (Menlo College), Stephen Nicholson (University of California, Merced), Randy Stevenson (Rice University), Joshua Tucker (New York University)[1]

Introduction

How to best measure politically relevant concepts has been a foundational issue for the ANES. The ANES must regularly balance maintaining consistent measures for the integrity of the time series with taking advantage of advances in measurement and instrumentation. When it comes to considering the future, at least two approaches are possible. First, one could discuss developments in how to best measure core constructs that currently appear or have appeared on the survey.[2] Second, one could focus on novel instrumentation that would expand upon what the ANES currently measures.

In this report, we take the second approach. We do this not to minimize the importance of continued discussion about topics such as question wording and survey mode. These topics will continue to be debated with each new ANES, and there are sizeable research communities (e.g. AAPOR) which provide cutting-edge knowledge of these types of measurement innovations. These issues are, in some sense, built into "future" discussions about each new wave of the ANES.[3]

We thus offer ideas that would expand measurement and instrumentation into new domains. As Aldrich's introductory essay notes, the first goal of the ANES "focuses on the citizens. The questions are what factors promote or inhibit the citizen from turning out to vote in the election, and what factors influence the choices voters make between or among the various parties and candidates." There have been at least two notable developments, over the past quarter-century, which speak directly to this goal. First, advances in the opinion formation research point to a number of processes that largely go beyond extant ANES instrumentation. This includes implicit (e.g., unconscious) psychological processes and health related variables that capture physiological reactions and fundamental health status (e.g., well-being). Second, the electoral communication environment has fundamentally changed in recent years – information is now ever-present, and people receive much of their information via the internet.

In what follows, we present a set of six instrument innovations for the ANES to consider. The first set entail measures taken *during the conventional survey interview* and involve capturing the aforementioned psychological processes and/or biological variables; specifically: physiological

[1] We also thank Ethan Busby and Adam Howat (both Northwestern Ph.D. students) for research assistance.
[2] This involves the consideration of "best measurement practice within surveys" (e.g., Schaeffer and Dykema 2011 *Public Opinion Quarterly*) and of alternative measurement approaches from distinct fields (e.g., Montgomery and Cutler 2013 *Political Analysis*).
[3] Moreover, we do not feel we are in a position to offer definitive advice about how to proceed or what criteria to use in assessing the tradeoffs between this type of new science and time series continuity.

measures, implicit measures, and biomarker measures. The second set of innovations capture behaviors *beyond the conventional survey interview,* aimed at assessing information acquisition. This includes web browsing measures, social media measures, and measures taken via micro-surveys.

Our list is not exhaustive and we recognize there would be costs involved in adding new instrumentation. However, we believe these ideas are worthy of consideration/discussion as the ANES moves forward in light of continued scientific advancement and communication evolution. These also are ideas that may facilitate connections with other scholarly communities (and potentially other funders). Given that the innovations differ from one another in terms of applicability and prior usage in the discipline, we discuss each in slightly different ways. Yet we do, for most, attempt to 1) describe the main theoretical construct, when relevant (in some cases there a large number of constructs or no particular constructs *per se*, which will be clear) and common operationalization(s), 2) touch on extant and potential applications, 3) detail hurdles to implementation, and 4) offer next steps in the way of advice, recommended readings, and/or expert suggestions.[4] All recommended reading references appear at the end of the document, prior to an extended appendix on social media application. That appendix details what was done with social media with the 2016 ANES.

In what follows, we start with our discussion of new psychological/health measures, and then turn to the communication/information oriented measures. To be clear, this report offers recommendations on topics for further discussion rather than any clear statement on what "should" be done. We also do not offer a formal conclusion as we view the point of the report as generating discussion on each possible technique introduced. As such, there is no overall conclusion to be drawn at the end of the report.

Measuring Novel Processes/Variables During the Survey Interview

The ANES is widely regarded as the most accurate and comprehensive survey of the American public's political attitudes, participation, and vote decisions. As noted, however, recent scholarly developments suggest that it may be worth considering incorporating additional measures to capture psychological and biological dynamics often missed in extant surveys. We focus here on three promising possibilities. It turns out that all three share two characteristics: they involve measures taken during the survey interview itself, and they avoid self-reports as all focus on constructs that are either beyond conscious awareness or likely to be mis-reported due to desirability pressures or lack of knowledge (e.g., about one's own health).

[4] The overall next step for any of these techniques, should they stimulate interest, would be to more clearly development instrumentation, consider pragmatics of adding it to the ANES, engage in pilot testing, and possibly do so with the consultation of experts from other fields that more regularly use the measures. We should note that we further recommend careful consideration on how to assess the validity of any new measure. In the past, the focus has been on predictive validity; while this is important, it also could potentially lead to a path dependency such that extant measures have established predictive validity which is part of the reason they have remained in the time series. In considering novel approaches, one should also consider content and construct validity and whether the new measures may add information beyond existing approaches (e.g., is there a payoff of keeping an extant measure for time-series purposes but adding a new one with better content or construct validity).

Physiological Measures[5]

Physiological responses to environmental stimuli are grounded in the autonomic nervous system. When encountering a threat, for example, our heart rates increase, we perspire, and our muscles contract. For the most part, responses of the autonomic nervous system happen outside a person's conscious awareness or control. Yet, not all people experience environmental stimuli in the same way. Some people will exhibit higher heart rates and sweatier palms than others.

Physiological measures have certain properties that make them attractive for studying political attitudes, preferences, and participation. As an involuntary response, a measure of physiological arousal captures an affective reaction to stimuli that cannot be censored by the respondent. If asked about a sensitive matter regarding race, for example, participants can misreport their opinions in a survey but (largely) cannot control affective responses. Physiological measures may also provide insight into affective orientations that respondents are incapable of reporting. They may also help with survey measurement potentially validating items (e.g., self-reported emotions). Lastly, affective responses to political stimuli have the added benefit of evading social desirability biases in reporting since they are largely outside the control of individuals).

When it comes to operationalizing measures of autonomic nervous system activity, researchers collect measures of electrodermal activity, electromyography, and/or cardiovascular activity. In brief, electrodermal activity involves recording variations in skin conductivity, electromyography involves recording variation in muscle movement, and cardiovascular activity involves recording heart rate. Given the tradeoffs between the different types of measure (as will be shortly discussed), electrodmeral activity (or EDA) represents the most promising avenue for measuring physiological response in the ANES.

When it comes to applications, physiological response measures may help us better understand the *affective* foundations of political attitudes and behavior. Since physiological measures capture a generalized affective response, the nature of the stimuli informs what is being captured. A primary distinction in the literature is between *political* and *non-political* stimuli. We discuss each in turn.

In terms of political stimuli, physiological readings are taken during the survey to assess responses to relevant stimuli. Mutz and Reeves (2005) provide a highly informative application of physiological response to political stimuli. In an experimental study of the effects of incivility on trust in government, Mutz and Reeves (2005) manipulated the levels of incivility expressed by two congressmen in a mock political talk show and collected measures of skin conductivity (EDM). In the uncivil condition, participants experienced significantly higher levels of arousal compared to the civil condition suggesting that viewing uncivil discourse is an emotionally stimulating experience. Although the authors did not claim to be capturing a specific emotional response, they suggest the increase in EDM is negatively valenced.

With this example in mind, we feel that a generalized measure of physiological measure would likely be informative for a good deal of ANES items. We offer a few (far from exhaustive)

[5] Special thanks to Vin Arceneaux, John Hibbing, and Kevin Smith for valuable advice and guidance on this section.

possibilities of how collecting physiological measures during the interview process might provide additional insight into ANES instruments.

An obvious and potentially valuable use of physiological measures would involve matching physiological readings to items about political parties and candidates. If the ANES collected physiological responses while answering questions on party identification and attitudes toward parties, researchers could look at differences between those exhibiting high or low arousal. Most likely one could not identify a specific emotional state such as anger or anxiety but given that many questions concern identity they could potentially capture the affective foundations of party identity. For example, research on affective polarization might be enriched by collecting measures of physiological arousal when respondents answer questions about parties, both their own and the out-party. To capture affective polarization, scholars have relied heavily on feeling thermometers, self-reported measures of “warm” or “cold” feelings toward parties. Despite that affect is central to this research, self-reported feelings are ambiguous. A physiological measure of affect captured during the interview process could potentially be used to validate whether the feeling thermometer measures are indeed capturing affect. In addition, it might be the case that some respondents exhibit greater physiological arousal than others during these questions and others about the parties (likes and dislikes) allowing researchers to examine differences in respondents who are high and low in affect toward the parties. Such measures, as discussed below, might be used as moderators to explore motivated reasoning, party loyalty, participation, and activism.

Similar measures might also help reveal more about the affective foundations of candidate evaluations. For example, assuming President Trump runs for reelection in 2020, imagine two Democratic respondents expressing strong disapproval of his job as president. One respondent might exhibit substantial physiological arousal when answering questions about Trump whereas another may not. Measuring levels of physiological arousal, whether high or low, in response to this question and other items about him and his Democratic opponent may shed light on the affective foundations of candidate (and party) evaluations and potentially explain, among other things, outcome variables such as political participation. For instance, those exhibiting high physiological arousal in response to these items might be more likely to participate in political discussions, put a candidate sign in their yard, and vote. In addition, a measure of physiological arousal might moderate the effects of candidate or party evaluation on outcome variables such as economic evaluations or trust in government. Democratic respondents exhibiting high physiological arousal when questioned about Trump might be more likely to report distrust in government or that the economy is doing poorly.

Other possibilities include racial attitudes, trust in government, conspiracy beliefs, and policy attitudes on “culture war” issues.

Although most ANES items likely would not lend themselves to understanding discrete emotions such as anger or anxiety, some ANES instrumentation might. For example, items specifically invoking emotional content (e.g., does President Trump make you feel angry, hopeful, afraid, proud, disgusted, sympathetic, uneasy?) might provide insight into what emotion the physiological response measure is capturing. In other words, measures of physiological arousal *might* potentially provide some insight into questions that invoke a specific emotion. For instance, a physiological arousal measure may illuminate the depth or intensity of emotional response potentially serving as a moderator for self-reported emotions. Alternatively, physiological measures may provide validation for items that purport to capture a specific emotion.

The examples to this point focus on adding to what we know about political reactions. Physiological measures also can be utilized to gauge fundamental individual characteristics that may explain political outcomes or moderate reactions to political stimuli – this would entail collecting physiological measures that are exogenous to politics by including *new, non-political* stimuli. Using pre-tested stimuli such as non-political images or video (e.g., a growling dog showing its teeth) that invoke specific emotional responses such as threat, disgust, or anxiety could provide valid measures of a respondent's sensitivity in those domains. In an example from the literature, Coe et al. (forthcoming) use a series of threatening images from the International Affective Picture System (IAPS) (Lang, Bradley, and Cuthbert 1999) to create a measure of threat sensitivity. Differentiating between those high and low in threat sensitivity, they examine susceptibility to frames that invoke threat versus frames that do not. Borrowing Nelson, Clawson, and Oxley's (1997) framing experiment on political tolerance in response to a Ku Klux Klan rally, Coe et al. found that when exposed to a public order frame invoking danger participants high in threat sensitivity were significantly less likely to express tolerance of the Klan compared to those low in threat sensitivity. On the other hand, there was no difference in susceptibility to framing effects between participants high and low in threat sensitivity when exposed to a free speech frame (that did not invoke danger). This study suggests that some people are physiologically predisposed to find some arguments more persuasive than others. Since people often lack the self-awareness to report their predispositions toward threat or disgust, physiological measures offer a unique opportunity to examine the effect of gut-level process on relevant political outcomes. Similar applications of using non-political stimuli that invoke specific emotions include looking at the role of anxiety in shaping immigration attitudes (Renshon, Lee, and Tingley 2015) and disgust sensitivity in explaining attitudes toward same-sex marriage (Oxley et al. 2008). Although space is scare on the NES, these new items would likely take very little time to administer (much less than an IAT, for example).

Of course, despite the potential benefits, even the most straightforward collection of physiological measures is likely to encounter difficulties. Both electrodermal and cardiovascular activity are promising in terms of cost and effectiveness but researchers have not found much relationship between cardiovascular measures and variables of interest to political scientists (hence the lack of discussion on such measures here).[6] Nevertheless, given that cardiovascular and electrodermal measures can be collected using the same device, it might be worth collecting both given the significantly larger N in the ANES.

A clear challenge is that any physiological measure should be collected in a quiet environment with the respondent physically stable. The measures are worthless if the respondent is moving or distracted by noise. To accomplish this, it might be worth providing noise-canceling headphones though this would pose obvious problems in face-to-face interviews. The first couple minutes of

[6] Personal communications with John Hibbing and Kevin Smith.

the interview might be confounded but after respondents have settled into the interview the readings are likely to be useful.

Empatica (E4) and iCalm both have wristbands that provide continuous measures of EDA. These are low-cost, comfortable items to wear and, importantly, can be attached by the respondent and (hopefully) quickly forgotten. The wristbands would likely need to be connected to a laptop so that the physiological readings could be synched with survey stimuli.

In terms of what to do next, it is critical to obtain the help of experts working in the field. More information about the use and cost of the wristbands, noise-canceling headphones, and whether separate laptops would be needed before proceeding. Pilot testing is a must. The obvious next step is to more thoroughly review key works, as listed in the reference section and consult relevant experts.[7]

[7] Some potential expert advisors include Kevin (Vin) Arceneaux, Temple University; John Hibbing, University of Nebraska-Lincoln; Matt Hibbing, UC Merced; Jaime Settle, College of William & Mary; Kevin Smith, University of Nebraska-Lincoln; and Stuart Soroka, University of Michigan. Mutz and Nicholson, Members of the Board, have also used these types of measures.

Implicit Attitude Measures

Implicit attitudes are evaluations that occur without conscious awareness towards an attitude object. There are many implicit attitude measures and an extensive literature evaluating them and applying them to politically relevant topics (see e.g., Gawronski and De Houwer (in press), Uhlmann et. al. 2012). Further, a previous ANES team authored a 2008 report accessing the usefulness of implicit measures for the ANES, which resulted in the inclusion of one implicit measure (the AMP) in the 2008 study. Given these developments, we by no means attempt to offer an exhaustive summary of the literature or re-produce analyses of the 2008 ANES measures (although we discuss them). Instead, as stated in the report's introduction, we review various measurement approaches, touch on applications, and offer some guidance for further discussion (including touching on hurdles).

When it comes to measurement and operationalization, there is some disagreement; however, Gawronski and De Houwer (in press) offer a compelling characterization:

> "measurement outcomes may be described as implicit if the impact of the to-be measured psychological attribute on participants' responses is unintentional, resource independent, unconscious, or uncontrollable. Conversely, measurement outcomes may be described as explicit if the impact of the to-be-measured psychological attribute on participants' responses is intentional, resource-dependent, conscious, or controllable … For example, a measure of racial attitudes may be described as implicit if it reflects participants' racial attitudes even when they do not have the goal to express these attitudes (i.e., unintentional) or despite the goal to conceal these attitudes (i.e., uncontrollable). An important aspect of this conceptualization is that the terms implicit and explicit describe the process by which a psychological attribute influences measurement outcomes rather than the measurement procedure itself or the underlying psychological attribute…"

Applied to the study of voting behavior, Lodge and Taber's "The Rationalizing Voter" (2013) is a leading example of this work in political science. In contrast to models of the vote that argue people draw on and assemble the considerations available in memory, Lodge and Taber turn this argument on its head by proposing that any such considerations follow, or rationalize, decisions that were arrived at instantly through deeply-ingrained affective responses to candidate stimuli. Using a variety of priming experiments, Lodge and Taber demonstrate that automatic, affective associations largely structure candidate evaluations. In other research using implicit measures, scholars have also shown that implicit racial (Kam 2007, Valentino, Hutchings and White 2002), gender (Mo 2015), and religious (Albertson 20111) attitudes may affect choice. There is also some discussion of the role of implicit and explicit attitudes in shaping the voting decisions of decided and undecided voters in pre-election polls. Galdi and colleagues (2008) found that explicit, but not implicit, attitudes affected the vote of decided voters whereas implicit, but not explicit, attitudes affected the vote of undecided voters. However, Friese and colleagues (2012) found that explicit attitudes better predicted the vote choice of both decided and undecided voters compared to implicit attitudes.

Given the central importance of partisanship to the study of elections and voting behavior, a highly promising application involves implicit measures of party identification (Theodoridis

2017) or party affect (Iyengar and Westwood 2015). Despite that explicit and implicit measures of party identification are highly related, Theodoridis (2017) shows that an implicit measure of party identification (using the IAT) better captures intensity than an explicit party identification measure (the standard 7-point NES question). Although both implicit and explicit party identification distinguish Democrats and Republicans from each other, the implicit measures was better at distinguishing between levels of intensity within a party in predicting outcomes such as party affect (e.g., feeling thermometer differences), differential candidate evaluation, and motivated reasoning processes. In a related application, Iyengar and Westwood (2015) used an IAT to measure party attitudes (rather than identity) and found that it was a strong predictor of affective polarization. Since implicit party measures provide greater insight into partisan intensity, they could help advance understanding of partisan motivated reasoning on a variety of ANES questions such as economic and candidate evaluations.

Implicit attitudes are often used when scholars believe that people are unwilling (social desirability bias or impression management) or unable (lack of self-awareness) to report their attitudes. In a systematic analysis of attitudes towards a number of different objects (political, social, etc.) Nosek (2005) found the highest correlation between explicit and implicit attitudes for political objects, r=0.7 (whereas the general relationship across domains was .36).[8] The strong correlation between many implicit and explicit political measures suggests that social desirability biases and lack of self-awareness may be less troublesome for political attitudes compared to other types of issues such as those involving racial attitudes. Nevertheless, implicit measures may provide greater insight into some concepts as party intensity (Theodoridis 2017). Furthermore, there are applications in which implicit and explicit political measures are largely unrelated. For example, Intawan and Nicholson (2016) show that despite the low levels of trust in government reported in national surveys, the modal respondent nevertheless exhibits an implicit trust in government. They find that both implicit and explicit trust in government predict system outcome variables such as justification and trust in government during crisis events.

There are a large number of implicit measures, and so here we detail the ones that are most likely to be of interest to political scientists, have been shown to be reliable, and have at least some chance of meeting the practical requirements of administration in a face-to-face or online survey format (many measures are usually administered in labs using special software). In addition, we also focus on implicit measures that are observational rather than experimental. The priming studies in Lodge and Taber (2013), for instance, are experimental and therefore largely inappropriate for use in a survey.

One of the better-known measures is the traditional Implicit Association Task (IAT). This is a computer-based task in which participants rapidly sort items from two target categories each with two attributes (e.g., white and black faces and good and bad things). The trick is that various pairs of these four attributes share a common response key so that differences in response latencies for different pairings show the strength of category attribute associations (e.g., white paired with good vs. white paired with bad). Trials requiring classification with all four possible pairings are presented to respondents in blocks.

[8] Also see Smith, C. T., & Ratliff, K. A. (in press). Implicit measures of attitudes. In T. Ortner & F. van de Vijver (Eds.), Behavior-Based Assessment: Going Beyond Self-Report in the Personality, Affective, Motivation, and Social Domains (pp. 113-132). Goettingen, Germany: Hogrefe

There are number of variants of the IAT, some of which have been applied to political questions. The Personal IAT came about in response to critiques that the traditional IAT reflects cultural expectations rather than personally held beliefs. The personal IAT substitutes evaluative categories like “good” and “bad” or “pleasant” and “unpleasant” with more personal ones like “I like” and “I dislike”. The Single block IAT (SB-IAT) eliminates the block structure of pairings by presenting all trials in a single block with instructions on pairings integrated into the trials. In one application “the relationship between the SB-IAT and outcome measures, SB-IAT scores showed a moderate correlation (r = .43) with explicit political attitudes and predicted voting intentions; predictive ability of the SB-IAT disappeared when entered into a simultaneous regression which included explicit attitudes (Teige-Mocigemba et al., 2008).” The Single Category Implicit Association Test (SC-IAT) only pairs a single target category (e.g., blacks) with both evaluative criteria (good, bad) but not to opposing target categories (white vs. black). In one study, the “SC-IAT has successfully predicted behavior in several different contexts. For example, more positive attitudes on a nuclear power SC-IAT predicted support for increasing governmental reliance on nuclear power and less reluctance to have a nuclear plant placed near their home (Truelove, Greenberg, & Powers, in press), a SC-IAT measure of trust in government showed that implicit trust increased support for system justification and trust in government during crisis events (Intawan and Nicholson 2016), and a political party SC-IAT predicted voting in the 2002 German Parliamentary elections (Friese, Bluemke, & Wanke, 2007).” Finally, the Brief IAT (BIAT) is “like the IAT but shorter.” “To date, the BIAT has only been used to predict voting behavior; a political party BIAT predicted voting intention in the 2008 Serbian Parliamentary elections (Pavlović & Žeželj, 2013) and implicit race bias on the BIAT predicted voting behavior in the 2008 U.S. Presidential race between John McCain and Barack Obama (Greenwald, Smith, Sriram, Bar-Anan, & Nosek, 2009).”

There are some non-IAT approaches as well.

The Go-No Go Association Task (GNAT) presents both target and distracter stimuli for brief periods of time and requires a "go" (press the space bar) for items that belong to instances of a target category (e.g., black faces) and a target evaluative attribute (e.g., good). No response "no-go" (do not press any key) is required when items appear that do not belong to the target category and target attribute. The extent to which the target category and attribute are associated is reflected in differences in sensitivity (a function of the number of correct and incorrect responses) between pairing conditions (e.g., black faces and good vs. black faces and bad). Thus, the GNAT can be scored without measuring response latencies (though these can also be used). One representative political study using the GNAT is by Knowles, Lowery, & Schaumberg (2010). They show that “increasingly negative attitudes toward Black people [as revealed by the GNAT] predicted less willingness to vote for Barack Obama in the 2009 Presidential election and less support for his health care plan, even when controlling for explicit prejudice” (Knowles, Lowery, & Schaumberg, 2010).

The Affect Misattribution Procedure (AMP) asks participants to make judgments (like whether the figure is positive or negative) about affectively “neutral” images like Chinese characters or Rorsarch type pictures. However, these figures are preceded by a prime such as a white or black face. Often, subjects are explicitly told not to let the prime impact their judgement of the character. Nevertheless, a large literature demonstrates that judgements are also affected. Pasek, Krosnick, and Thompson (2012) and Ditonto, Lau, and Sears (2013) both examine the impact of implicit attitudes toward blacks on vote choice using the AMP. As discussed below, both showed implicit measures to have some predictive power but not over and above explicit measures of racism.

When it comes to additions to the ANES, the starting point is recognition that most political applications of explicit measures have been to predict/explain vote choices, which is of course central to the mission of ANES. One set of such studies argues the value of implicit measures is that they can reveal support for candidates and parties even which the respondent is unaware that they have an implicit preference. Thus, this work tends to focus on the usefulness of implicit measures of support for understanding undecided voters. It is unclear whether such measures, even if effective for this purpose, would be more effective than the explicit technique of pushing respondents to "lean" one what or the other, which of course has proven quite strongly predictive of choice.

Finally, given the frequent findings that implicit measures are most strongly correlated with behaviors or attitudes that are produced quickly and/or under time or resource pressures, political scientists might reasonably ask how they can be associated (as they certainly are in some studies) with deliberative future behaviors like reasoned policy choices, votes, or participation decisions?

One promising answer to this question is work that shows that implicit attitudes predict selective exposure and thus that is a potentially useful application. Specifically,

> "Whereas selective exposure in decided participants showed stronger relations to explicit compared with implicit measures, selective exposure in undecided individuals showed stronger relations to implicit compared with explicit measures. Such biases in information processing explain why implicit measures are capable of predicting future choices and decisions that seem highly deliberate, such as voting behavior and other political decisions (e.g., Galdi, Arcuri, & Gawronski, 2008; Payne, Krosnick, Pasek, Lelkes, Akhtar, & Tompson, 2010). For example, undecided voters may selectively expose themselves to information that is consistent with their implicit preference, and this biased set of information may ultimately provide the basis for their deliberate decision to vote for a particular candidate. Thus, to the extent that deliberate choices are based on the information that is available to an individual and the representations captured by implicit measures predict processing biases in the acquisition of this information (e.g., biased interpretation, selective exposure), implicit measures can be expected to make a unique contribution to the prediction of future decisions even when these decisions are highly deliberate"(Gawronski and De Houwer, pp.28-29).

In 2008 Krosnick and Lupia reviewed various implicit measures for use in the ANES and adopted the AMP for use in the ANES. One of the main justifications that Krosnick and Lupia provided for the choice of the AMP over the IAT was that the IAT required a two-group comparison rather than a being relevant to a single group. That issue has been solved with the development of the IAT-SP (or the GNAT) which does not require a comparison group.

Further, the performance of the AMP derived implicit measures of racial attitudes was disappointing in that they showed little added explanatory power for vote choice relative to explicit measures (Ditonto et al 2013). Similarly, Pasek, Krosnick, and Thompson (2012) show that implicit attitudes towards African Americans, measured using the ANES AMP, explain no additional variance in either vote choice or presidential approval when explicit attitudes were controlled for. These items may have some use, however, as a study of only the 2008 election published in 2010 by the same authors concluded that:

> "Both explicit and implicit prejudice were significant predictors of later vote choice. Citizens higher in explicit prejudice were less likely to vote for Barack Obama and more likely to vote for John McCain. After controlling for explicit prejudice, citizens higher in implicit prejudice were less likely to vote for Obama, but were not more likely to vote for McCain. Instead, they were more likely to either abstain or to vote for a third-party candidate rather than Obama."

Additionally, Stanley Feldman has a working paper shows implicit measure mattered for liberals, not conservatives. This suggests some heterogeneity in implicit measures and political outcomes, which might explain differences in previous findings.

Many implicit measures require special software to administer. While this can be done easily in an in-person interview (by allowing the respondent to answer on a laptop), until recently this was difficult to do on an online survey unless the respondent was willing to download special software (that can, among other things, measure response times precisely). However, recent studies have overcome this problem using web-based platforms (see Hansford, Intawan, and Nicholson 2017; Intawan and Nicholson 2016). Of course, some measures (like the AMP and one version of the Go no Go task) do not require response times and so are easier to implement on online surveys.

Another type of implicit measure involves mouse-tracking. In contrast to the IAT and other types of implicit measures covered thus far, mouse tracking does not involve a separate task. Instead, mouse tracking can be used to measure various aspects of how people respond to political stimuli when using the computer illuminating aspects of information processing as respondents answer on-line or computer survey items. Mouse tracking can provide measures for latency, velocity, acceleration, and pull towards non-endorsed items. Duran, Nicholson, and Dale (2017) provide an application of mouse tracking to the study of conspiratorial political beliefs. Here, they find evidence that, compared to baseline measures, people who endorse conspiracy items are tempted by accuracy motivations and those who disavow conspiratorial items are tempted by motivated reasoning processes (wanting to endorse something negative about the out-party). In contrast to prior studies of conspiratorial beliefs that rely exclusively on self-reports, this research provides insight into how people process information while answering items.

In terms of next steps, the starting point is, as mentioned, that the literature on implicit measures is well developed and, for many measures, includes applications to core political constructs like vote intention, party identification, and the political relevance of attitudes about race, gender immigration, and more.[9] Further, as just noted, in 2008, the ANES investigated such measures (focusing ultimately on AMP and the IAT, but initially exploring more widely). That analysis resulted in the inclusion of the AMP on the 2008 study and several resulting investigations of the usefulness of the measures (which should be carefully consulted as the results seem mixed, as noted above).

We recommend the ANES undertake additional testing to evaluate the usefulness or an implicit measure of party identification or party affect. Furthermore, given that mouse-tracking would not take any additional time on the survey, it too represents a promising next step for further evaluation.

[9] In terms of experts to consult, there are many psychologists (e.g., Nosek) and ANES board members (e.g., Iyengar, Nicholson, Valentino) who are well versed in implicit measures.

Biomarkers

Biomarkers are "directly measured traits that provide insight into the functioning of biological systems" (McDade 2010). From the perspective of social scientists, biomarkers provide an opportunity to pinpoint the specific physiological processes through which contextual (e.g., cultural, economic) factors shape people's well-being. This obviously overlaps with, and in some sense envelopes, the previously discussed physiological measures, but we focus on distinct instruments that focus more on general health constructs.

Specifically, we have in mind biomarker measures that involve specimens (e.g., blood, saliva). Methodological advances in recent years have made the collection of biomarker data more practical—a number of population-based surveys in the US and beyond have made use of them. As summarized in McDade et al. (2007), these include the Health and Retirement Study (http://hrsonline.isr.umich.edu/); the National Longitudinal Study of Adolescent Health (http://www.cpc.unc.edu/projects/addhealth); and the National Social Life, Health, and Aging Project (http://www.norc.org/Research/Projects/Pages/national-social-life-health-and-aging-project.aspx). Notably, all of these studies concern a range of social and behavioral outcomes as well.

A popular method is dried blood spot (DBS) samples, obtained through a finger prick and collected on filter paper. The process is fairly easy; non-medically trained interviewers can be trained to do it, as can the participants themselves (McDade 2013). DBS samples can initially be transported and stored at room temperature (though they must be kept in laboratory-grade freezers long-term). DBS samples can be used to measure a variety of substances and processes in the body, such as proteins associated with stress (e.g., McClure et al. 2015; see McDade et al. 2007, Table A1, for a summary of DBS uses in the literature). In general, these biomarker data provide objective measures of respondent health, which present some advantages over self-reports (e.g., respondents need not be aware of a health issue for the analysis to detect it).

A second alternative is saliva samples, from which a variety of *genetic* data may be extracted. The Wisconsin Longitudinal Study (WLS) has conducted a pilot study using such samples (see http://www.ssc.wisc.edu/wlsresearch/about/Genetic_Data.pdf), enabling researchers to test for predispositions to conditions such as heart disease, cancer, diabetes, and more. The National Longitudinal Study of Adolescent to Adult Health has also collected genetic data, which researchers have connected to a variety of health and behavioral outcomes.

In terms of applications, biomarker data—and health data more broadly—have been linked to a variety of social and political outcomes. McDade et al. (2006) use DBS samples to find that psychosocial stressors contribute to the production of C-reactive protein, a predictor of cardiovascular disease. Although they find no differences between demographic groups once they control for stressors, these stressors appear to be more prevalent among African Americans and women, as well as those lower in education. Along similar lines, McClure et al. (2015) employ biomarker data to examine the effects of psychosocial stressors on health among immigrants—a population whose health has been shown to deteriorate the longer they reside in the US. Alongside blood pressure, cholesterol, and a few other measures, immune function using DBS samples constitute the authors' index for "allostatic load" (AL)—a general measure of physiological wear and tear. They find that Mexican immigrant women living in majority-white communities, with low levels of social support, were eight times as likely as women with high social support to have higher AL scores. They do not find this same difference among immigrant women living in Mexican enclaves. Taken together, these studies suggest that social context has a substantial impact, through stress, on some groups' physical wellbeing. The same may prove true for more directly political phenomena.

Genetic data, such as that obtained from saliva samples, has also been connected to behavioral and political outcomes. McDermott et al. (2013), for example, find an interaction between one's genetic disposition and life experience: Individuals with the low-activity form of monoamine oxidase-A (MAOA) who are exposed to violence in youth (e.g., in a conflict zone) are more likely to engage in physical aggression themselves during adulthood.

With regard to the ANES, biomarker data could potentially be collected from respondents during the survey interview itself. The resulting ability to measure biological processes directly provides a number of advantages over self-reported health measures, which may not always be reliable (McDade et al. 2007; McDade 2010). Health measures, in general, have proved relevant to a number of political and social phenomena:

- ***Turnout:*** Gollust and Rahn (2008), find that chronic health conditions impact turnout, with cancer diagnoses *increasing* turnout while heart disease and generally poor health *decrease* turnout. Pacheco and Fletcher (2015), along similar lines, find that those who report "excellent" health are both more likely to turn out to vote and to identify as Republicans.

- ***Vote choice:*** Recent analysis finds that, even after controlling for a variety of demographic factors, worse health outcomes in a region predicted a greater vote margin for Donald Trump in the 2016 presidential election (Economist 2016).

- ***Social capital:*** Putnam (2000) links declining social connectedness to poorer (self-reported) health outcomes. His work primarily frames health as a *consequence* of reduced social capital; however, the possibility remains that decreased health might engender less social engagement and trust. Use of objective health measures might untangle this causal pathway.

In addition, it seems probable that objective measures of health could influence more specific policy preferences. Most obviously, those in poorer health may express greater support for health care programs such as Medicaid and the Affordable Care Act. It remains to be seen, however, whether direct measures via DBS samples would more strongly predict these attitudes.

Causal processes may flow in the opposite direction as well. Given the demonstrated impact of social context on disease risk (e.g., McClure et al. 2015; McDade et al. 2006), it seems probable that more overtly political phenomena might exert a similar impact. Measuring these outcomes using DBS samples in the ANES would enable researchers to test this possibility. Much remains to be learned, as well, about the impact of health *on* other political attitudes (e.g., policy preferences). However, in the latter case, it remains somewhat unclear which outcomes would be better predicted by direct measurement via biomarkers as opposed to self-reported health (i.e., can health conditions of which respondents are unaware affect their attitudes?).

There are clear hurdles as biological samples present additional difficulties for survey researchers. They impose a greater burden/risk on the study participants (though these are minimal), as well as additional logistical concerns (e.g., training of interviewers, transport and storage of samples, safety concerns, response rates; see Dykema et al. 2017). For example, although DBS constitutes a relatively inexpensive option compared to other methods, the costs are non-trivial: for each participant, $1.50-$2.00 for collection supplies and $5-$20 for lab analysis (McDade et al. 2007). The collection of any biomarker data also necessarily broadens the range of ethical concerns in a study (ibid.). Moreover, the laboratory processing involved in genetic data collected from saliva samples may be prohibitive. The *relatively* low-cost traditional DNA sequencing employed by WLS (see http://www.ssc.wisc.edu/wlsresearch/about/Genetic_Data.pdf) costs a minimum of $100 per subject and often much more. Moving forward, we recommend consulting with scholars who have used or are currently using biomarker data and considering whether these measures are worth the investment from a political science perspective.[10] One potential benefit, however, is connecting with the growing community of social and biological science scholars who have interest in such measures.

[10] Thomas McDade, Northwestern University, is an expert on the use of DBS samples and was a great resource during the creation of this initial report. Rose McDermott, Brown University, possesses expertise in the use of physiological and biomarker data, including in specific reference to political phenomena. Barry Burden, University of Wisconsin-Madison, in a personal communication, indicated that he has not yet conducted research using biomarker data but has an interest in doing—thus, he could potentially comment on the usefulness of such measures for a range of political outcomes.

Measuring Communications/Variables Beyond the Survey Interview

The electoral information environment fundamentally differs from what it was a quarter century ago. Citizens obtain their information from radically different sources via the internet. Politically relevant news is ever-present not only via news outlets but also entertainment sources and social networks. This is a challenge for the ANES since it means the origins of opinions and behaviors extend beyond the traditional electoral season and come from so many sources that it is unreasonable to expect survey respondents to remember each source. In this section, we discuss three measurement innovations meant to expand the ANES to better address modern campaign context. Each involves collecting data outside of the traditional survey interview, using relatively new technologies, and involves capturing multiple constructs (and thus we do not so strictly define particular constructs in this section). It is worth noting too that the social media discussion builds on the pilots conducted with the 2016 study (led by Josh Tucker who wrote the below social media section) and the web browsing behavior idea is already being planned for 2020 (spearheaded by Shanto Iyengar who wrote the below web browsing section).

Behavioral Measures of Media Exposure: Web Browsing Behavior

Media exposure has been part of the ANES survey mission from the very beginning. In recent years, to counter widespread evidence of inflation in self-reported/recalled exposure, ANES adopted new items that ask respondents to select from a list of media sources (Dilliplane, Goldman, and Mutz, 2013). While the "list" approach may bypass problems associated with exaggerated recall and response set, it too is susceptible to criticism on multiple grounds, most fundamentally, that it is unable to differentiate between frequent and infrequent users of particular news outlets (see Prior, 2015). Moreover, as mentioned, a changed communication/media environment has altered the nature of when exposure occurs and how it processed – thus, beyond measurement concerns, it is vital the ANES work to accurately capture varying types of exposure, much of which happens via the web.

When it comes to media/communication exposure, the gold standard for evaluating any measurement is actual media/information consumption. Fortunately, it is now possible to track media usage (and general information consumption; this is the general construct of interest), at least in the case of online news, and there are several large-scale studies of web browsing behavior (Gentzkow and Shapiro, 2011; Flaxman, Goel, and Rao, 2016; Hannak et al., 2013; Goel, Hofman, and Sirer, 2012). These studies have investigated, among other questions, the extent to which individuals gravitate to biased sources and the prevalence of so-called "echo chambers" inhabited by partisans seeking like-minded news providers. To date, however, the behavioral browsing data have yet to be merged with survey evidence on individuals' political attitudes and preferences. Fortunately, the technology for merging surveys with web browsing behavior is now widely available. Several companies have recruited web panelists who agree to install an application that tracks their web browsing activity. The database generated by the application includes every URL visited by the survey respondent and the time spent at the URL. For instance, we can observe the number of times respondents clicked on news reports from foxnews.com and can match this with survey responses measuring self-reported exposure to Fox News. It will be possible, in short, to compare self-reported media exposure with actual exposure. Of course, the behavioral measures of media consumption can be used itself to predict a variety of important outcomes including political knowledge, issue salience, partisan polarization, and candidate preference.

A group of Stanford researchers (Sharad Goel, Shanto Iyengar (who wrote this section), and Erik Peterson) have recently worked with one of the web browsing applications, developed by Wakoopa (a Dutch market research firm). YouGov has recruited a subset of their national panel to install the Wakoopa application. These panelists have granted informed consent for their anonymized browsing behavior to be made available to researchers. YouGov clients are now able to purchase survey respondents' web browsing behavior. These Stanford researchers conducted a two-wave panel study over the course of the 2016 presidential campaign. Interviews occurred in August and November of 2016, and the researchers also obtained respondents' web browsing activity over this three-month period. The data not only have the complete set of news organizations that respondents visited, but also the full text of the news reports (scraped from the URL). The team then uses machine learning to classify the text of news reports. They are therefore able to examine both volume of exposure to media sources as well as exposure to different types and genres of news reporting (e.g. hard news versus soft news, political coverage versus sports coverage, issue-based versus horse race news etc.). Their results suggest that patterns of media usage differ significantly by news content. For news reports with little political content, party identification has only weak effects on preferred (visited) sources. However, when the news coverage is election-relevant, there is clear evidence of divergence in the media behavior of Democrats and Republicans.

Next steps here are relatively straightforward insofar as a subset of the current PIs plan to incorporate web browsing behavior in the proposal as part of the online mode for the 2020 study. Potential respondents might be incentivized to install a tracking application and it can be programed to record only visits to a list of news organizations, thus assuring respondents that other aspects of their web browsing are not available. Alternatively, the ANES can subcontract with Yougov to recruit a fresh panel with web browsing enabled for the purposes of the 2020 study. The YouGov costs appear modest, at least in comparison with what the ANES has been paying respondents in 2016. For a sample of 1400 and two fifteen minute interviews, the per capita cost for monitoring web browsing activity between August and November amounted to $45.[11] The data would be collected presumably from the first week of the pretest survey through the completion of the post-election wave

.

Clearly, there are potential limitations to observing online news consumption. Monitoring web browsing activity is obtrusive; respondents know that their behavior is being tracked. This may lead to Hawthorne effects and other forms of reactivity in the data. In the case of the 2016 study, such biases seem very minor as the ranking of web sites based on the amount of traffic matches other research in which participants did not know that their browsing activity was to be recorded (as in the case of the studies based on Internet Explorer users). Another limitation is the

[11] In moving forward, experts to consult include Sharad Goel, Stanford; Doug Rivers, Stanford; Erik Peterson, Stanford (soon to be Dartmouth); and Lynn Vavreck, UCLA.

possibility of selection bias; people who consent to have their online activity tracked might be outliers on several traits. Once again, the 2016 data do not support such concerns. Survey data for panelists who installed the application and those that did not was obtained. Comparisons across the two groups indicated only minor differences on gender, party, education, and other relevant factors. One lingering issue, though, is that this is likely only feasible to do with the online sample given the invasiveness of trying to get the face-to-face respondents to access their computers and install the application.

Social Media: **Approaches and Measures**

There are three big picture ways we could think about social media measures/data and the ANES, which we can loosely call the "alternative" option, the "survey platform" option, and the "supplemental" option.

We define all three of these options, but focus most extensively on the third due to the fact that it is closest to the actual remit of the committee and because we have actually taken some preliminary steps in this direction as part of the current wave of the ANES. We do, however, also offer some preliminary thoughts on the first option which, although beyond our current remit, is something that should at least be on the ANES's radar moving forward.

The *first option* is to use social media as an alternative ANES. Indeed, according to the most recent Pew Research report, 79% of online adults in the United States currently use Facebook, 32% use Instagram, and 24% use Twitter. [12] Thus the most transformative way we might think of the relationship between social media and the ANES would be to actually conduct an "Election Study" utilizing digital trace data that is posted to social media platforms. Tools and techniques would be need to be developed to measure the relevant quantities of interest on social media platforms, but it is worth noting that these platforms already have some structured "interview" questions – in the form of respondents' biographies – as well as virtually limitless open ended time-series answers to the question "what are you thinking about today that you'd like to share with other people?".

Due to the fact that the mission of this particular report instructs us to keep in mind that "the primary goal of the ANES, which is to produce 'high quality data from its own surveys on voting…'", we will not engage in a drawn-out discussion here of the pros and cons of moving from traditional survey methods to more of a data mining approach from social media platforms, but we want to make two points here. First, there are "pros" to this approach that would allow the ANES to do things that are currently impossible using standard survey methodologies. One class of advantages would revolve around the ability to collect fine-grained time-series data in a way the ANES as currently constructed cannot. To give an example from the 2016 campaign, the ANES was in no way able to track the life-cycle effects of the release of the Access Hollywood video or the various "fake news" pieces that appeared especially towards the end of the campaign. These are research questions that are of great interest to political scientists and those outside the academy, and a social media based Election Study could provide rich new sources of data with which to address these types of questions. Another class of advantages would revolve around the detection of new politically relevant topics. As currently structured, the ANES requires us to choose the topics for our questions months ahead of time (and, if we are being honest, the time-series component of the study means that most of these questions were decided on years if not decades earlier). A social media based ANES, in contrast, by relying on primarily open-ended responses (i.e., individuals' "posts") could incorporate elements of topic discovery as part of its analysis. Second, the cost of a social media based ANES would be much, much

[12] http://www.pewinternet.org/2016/11/11/social-media-update-2016/. Pew estimates that 86% of Americans are currently online.

lower than the cost we currently pay to conduct interviews. If we are entering an era where funding available to the ANES is going to decrease dramatically, harnessing digital trace data – and especially open-source digital trace data – as a form of future election studies would seem to be an important avenue to consider.

This is of course *not* to say there would not be significant hurdles to such a study, including technical, legal, and ethical concerns. But for now, we leave such considerations to future study due to the focus of our mandate here to focus on the current survey structure of the ANES.

The second option for using social media is to use it as a *platform for conducting interviews*. Of the three options discussed here, this is potentially the least exciting/interesting, and probably the least important due to the fact there are so many different internet survey options available today. Nevertheless, we would be remiss if we did not at least include a mention in this report of the possibility that social media platforms could be used as a venue for conducting surveys. Indeed, Facebook has what appears to be a suite of tools available to let users do just this (https://apps.facebook.com/my-surveys/), as does Twitter (https://about.twitter.com/company/polls), although the latter is more for micro-polls of a small number of questions. Further, there are scholars who have used Twitter as a platform for inviting people to take surveys on other platforms (such as Qualtrics).[13]

The third and most promising (and already pursued) social media opinion is to use it as a supplement to the existing ANES survey. The basic idea is to builds on the existing ANES survey structure, with the added dimension being that we can also ask survey respondents for their Twitter handle and/or access to their Facebook data. These additional sources of data can then be used to calculate additional individual specific variables that could be included with the ANES files alongside survey responses. These measures would supplement traditional survey response questions by providing a set of non self-reported "objective" variables. While such variables could be especially useful for measuring social media usage, they could provide all sorts of other variables of interest. In short, this would capture a host of constructs and provide measures of media behavior that relate but are distinct from the just discussed web browser tracking instrument.

Examples of variables could include:

- Number of politics related posts on Twitter (Facebook)
- Number of friends/followers on Twitter (Facebook)
- Estimated political ideology from Twitter follower network[14]
- Reported political ideology from Facebook
- Likes of "fake news" websites on Facebook
- Shares of (Fox News/Breitbart News/CNN/NY Times) links on Twitter
- Ratio of political to non-political posts on Facebook

[13] See for example Bode et al. 2011, Vacarri et al. 2013.
[14] See Barberá 2015; Barberá et al. 2015.

How would this work? Data could be collected from the first week of the pre-test survey through the completion of the post-election and possibly beyond. On Twitter, this is relatively straightforward. It involves simply asking a respondent during an interview for their Twitter handle. With the handle in hand, the ANES could then collect the respondent most recent 3200 tweets as well as their entire Friend and Follower network using an R library like StreamR.[15] On Facebook, the process is a bit more complicated. We need to first build a Facebook App that will essentially transfer people's personal records from Facebook to a secure server. Once the App has been built and approved, then people taking the survey can log in to it to have their FB data downloaded. This could work for both online and face-to-face participants as long as the interviewer/respondent has a mobile device connected to the internet. Technically, such an App violates the current FB terms of service because it does not provide any real value to the user,[16] but FB seems to be making an exception to this rule for research studies. We know – at least for now – that this approach is actually feasible, because we did just this for the 2016 ANES. Our experiences to date are detailed below in the section, appended to the report, *What We Did With Social Media In the 2016 ANES*. (This was spearheaded by Joshua Tucker, who wrote this section, and the PIs. We thus do not list "experts" here since the approach is far into development already within the confines of the ANES.)

There are a number of very legitimate concerns with augmenting ANES surveys with social media data, which we break down below loosely into three categories: ethical, logistical, and analytical. Ethically, people's identity can be revealed through their social media data. This is easy to accomplish with Facebook data, and, while more complicated, certainly also possible with Twitter data. Thus, if the raw FB or Twitter data was released along with the survey answers, anonymity would be compromised. Therefore, it is absolutely imperative that the raw data *not* be publicly released along with the survey data. Additionally, people may not realize when being asked for their Twitter handle or to log in to FB that this means there could be/will be identifying information connected to their survey answers. Thus, care should be taken to ensure that separate consent is given when providing social media data, and to make sure respondents understand exactly what they are doing.

When it comes to logistics, the ANES would have to deal with the logistics of identity protection and thus data from the survey and from the social media platform should be stored separately and only linked with some strong form of encryption key. The data should only be linked *once variables have been created from the social media data*. Also, IRB review will be needed, which might lead to different types of data being permitted at different institutions and in different time

[15] https://CRAN.R-project.org/package=streamR . Indeed, there is nothing to stop the ANES from at that point creating a two-hop network (friends of friends), which would permit more serious forms of network analysis; see for example Larson et al. (2017).

[16] E.g., if the user got a "chart of the day" about the election from the ANES, that would be a service provided to the user. FB used to allow apps that would do anything in terms of extracting data, but a while ago the policy changed to require apps to enhance the user experience, not just gather data for the people who created the app. Whether apps meet this requirement appears to be determined on a case by case basis.

periods. Finally, getting this data is dependent on FB and Twitter not changing the way they distribute data publicly.

Analytical concerns include: not everyone in the ANES will have social media accounts, and not everyone who has a social media account will share it with us (and thus response rates will matter and there will be bias in the data we do manage to collect). Additionally, the ANES will likely never be able to release all the raw data, which means that it will be up to ANES to determine which variables are created and released (although in the long term some sort of Online Commons approach could be used to allow users to create variables). This could also introduce bias in terms of what is studied, although likely not any more than the type of bias that is introduced by the PIs choosing which questions to ask in a given survey.

Next steps are clear and involve building on the 2016 effort which again is detailed in an appendix.

Micro-Surveys

Our last measurement innovation does not focus on a particular construct (similar to the last two in that regard) but rather as pure extension of the general ANES instrument. Specifically, it involves micro-surveys which use cell phones to collect information from respondents, asking only one or a few questions at a time. Health care providers in a variety of settings have used micro-surveys to collect data from patients. Micro-surveys could be used by the ANES to study campaign effects and isolate trends with-in and between respondents. Respondents could be compensated for responding.

While most consumers use smartphones, some mobile phone customers still use feature phones (aka dumb phones), as shown in Table 1. According to a January 2017 Pew report (based on a November 2016 survey), about 77% of Americans now own a smartphone.[17] But ownership is not evenly distributed: 92% of 18- to 29-year olds own a smartphone, 74% of 50- to 64-year olds, but only 42% of those age 65 and older and only 64% of those in households earning less than $30,000/year. At the same time, some 95% of American adults own a cell phone of some kind.[18]

Feature phone users can respond to micro-surveys sent by text message that prompt users to reply with a short response. This can include predefined ("text 1 for yes, 2 for no") or open-ended responses. Examples from the health care field are provided below. Smartphones, of course, can be used in much more sophisticated ways, but as shown in Table 1 there are segments of the public with relatively low rates of smartphone ownership, especially older, less educated, and lower income adults.

[17] http://www.pewresearch.org/fact-tank/2017/01/12/evolution-of-technology/
[18] http://www.pewinternet.org/fact-sheet/mobile/

Table 1. Mobile Phone Ownership among U.S. Adults

	Any mobile phone (%)	Smartphone (%)	Feature phone (%)
Total	95	77	18
Men	96	78	18
Women	94	75	19
White	94	77	17
Black	94	72	23
Latino	98	75	23
Ages 18-29	100	92	8
Ages 30-49	99	88	11
Ages 50-64	97	74	23
Ages 65+	80	42	38
Less than high school graduate	92	54	39
High school graduate	92	69	23
Some college	96	80	16
College graduate	97	89	8
Income <$30,000	92	64	29
Income $30,000-$49,999	95	74	21
Income $50,000-$74,999	96	83	13
Income $75,000+	99	93	6
Urban	96	77	17
Suburban	96	79	16
Rural	94	67	27

Source: Pew survey conducted Sept. 29-Nov. 6, 2016, http://www.pewinternet.org/fact-sheet/mobile/.

The use of micro-surveys via mobile phones offers several advantages for data collection. One is that the surveys could be pre-translated into a variety of languages, allowing for respondents to choose to participate in their preferred language and potentially increasing response rates and compliance among language minority groups. Another is that data would be incoming on a rolling basis, allowing for immediate changes and/or re-surveys if a particular micro-survey failed or if circumstances on the ground during the campaign suggested the need for new information. It would be a flexible tool that would allow for rapid response by the ANES team to on-the-ground campaign and world events. It also would address the aforementioned extant media environment where information is constantly available and at moves at incredible speeds.

To see how this might work, consider various applications from work by health care providers. Studies conducted in developed country settings have investigated the use of cell phones on the patient end to generate feedback for improved chronic illness care and monitoring (Cho et al. 2009, Shapiro et al. 2008, Anhøj 2004), increased medication compliance (Cocosila et al. 2009) and smoking cessation (Rodgers et al. 2005).

Cho et al. 2009 randomly assigned diabetes patients to submit blood glucose information via a mobile phone with a glucometer integrated into the battery pack (the 'Diabetes Phone') or via the Internet-based glucose monitoring system (IBGMS). Patients could also communicate with medical staff via their assigned device. After three months' intervention, HbA_{1c} levels of both groups had decreased significantly and levels of patient satisfaction and adherence to medical advice were similar. Mobile, bidirectional communication between doctors and patients using the diabetes phone was as effective for glucose control as the previously-studied Internet-based monitoring system and it was good for patient satisfaction and adherence. One point here too is these technologies could potentially be used to collect biomarker type data (see section on biomarkers).

Shapiro et al. (2008) randomly assigned phones to families with obese children (age 5-13), with a control group of families asked to report information using paper diaries. Each family in the SMS condition was instructed to send 2 SMS per day (one for parent and one for child), daily for the full 8 weeks of the study, and for each SMS sent, they would each receive an immediate, *automated* SMS feedback message from the program hosted on a secure server. The feedback message was automated to provide instant responses to the participants regardless of the time of day. Hundreds of feedback messages were developed to avoid duplicate messages; algorithms were based on (1) how many goals were met and (2) enhancement or deterioration from the previous day (e.g., "Wow, you met your step and screen time goals—Congratulations! What happened to beverages?"). Families in the paper diaries (PD) condition used self-monitoring forms to record the 3 behaviors daily for both parent and child, turned in their forms at each session, and received weekly verbal feedback. Families in SMS and PD completed daily responses to 3 questions: (1) what was the number on your pedometer today? (2) how many sugar-sweetened beverages (SSB) did you drink today? and (3) how many minutes of screen time did you have today? Children in SSM had somewhat lower attrition and significantly greater adherence to self-monitoring.

A major takeaway from the Cho et al. study is that participants may be willing to have their phones automatically report data to researchers. A major takeaway from the Shapiro et al. study is that participants are willing to answer multiple mini-surveys every day over a fairly long period of time.

In contrast, Anhøj and Møldrup (2004) found that participants did not want to answer multiple messages each day. They tested the use of SMS messages to monitor Dutch asthma patients. Over a period of 2 months, participants received 4 SMS messages each day, including a medication reminder, a request to enter peak flow, data on sleep loss, and medication dosage. Participants were asked to reply to a minimum of 3 of the messages per day. Half the participants reported more than about two thirds of the requested diary data. Furthermore, response rates were relatively steady during the study period with no signs of decreasing usage over time. From the subsequent focus group interview with 9 users we learned that, in general, the participants were enthusiastic about the SMS diary – it became an integrated part of their everyday life. However, the participants wished for a simpler diary with only one SMS message to respond to.

Cocosila et al. (2009) conducted a randomized trial where participants were to take 1 vitamin C pill per day for 1 month for preventive reasons; those assigned to the treatment group received text message reminders and were asked to acknowledge receiving their messages after taking the vitamins, whereas control group subjects had no text messaging activity. Adherence was higher in the treatment group, but the difference was not statistically significant.

Rodgers et al. (2005) used text messaging to encourage smoking cessation among young adults in New Zealand. Participants were allocated to either a control group or to a group that received a support program. Participants allocated to the intervention group were sent regular, personalized text messages providing smoking cessation advice, support, and distraction. Several other text message based services were provided for the intervention group: Quit buddy (participants with similar characteristics and quit days were put in touch with each other); TXT crave (participants could ‘‘pull’’ text messages on demand by sending a text message to a short code number and they would receive a tip on how to get through the cravings); TXT polls (for example, messages sent to all participants on current topics, and the answers were sent back to all); and TXT quizzes (questions were sent out, followed by answers the next day). Control group participants only received one text message every two weeks, thanking them for being in the study, reminding them of the incentive at the end, and providing information about how to contact the study center. Overall, those in the treatment group were more likely to quit. Unfortunately, the authors do not provide any information about the TXT polls and quizzes, but could potentially be reached and asked. Of note is that there were five car crashes during the study that occurred just after participants were texting, emphasizing the importance of reminders to participants to no reply to the micro-surveys while driving or operating heavy machinery.

Ostojic et al. (2005) explored the use of SMS for monitoring of young (average age 24.6) asthma patients. Patients randomly assigned to the treatment group were instructed to send a daily text reporting their peak expiratory flow (PEF) and symptoms. Patients in the control group were told to keep a paper diary of the same data and bring them to the clinic at the end of the 16-week study period. Compliance was nearly perfect with the SMS, while the paper diaries were incomplete. Patients thought SMS was convenient, and that it did not intrude into their daily activities. Ferrer-Roca (2004) used SMS for monitoring of diabetes patients of all ages. Participants reported their daily blood glucose levels and body weight. The researchers found that while compliance was high the elderly participants needed assistance from younger relatives to submit their responses.

Tomlinson et al. (2009) used mobile phones to conduct a survey among lay community health workers in a peri-urban settlement in South Africa. In this research, the data was entered by 24 local women who were hired and trained in how to collect data using “Mobile Researcher,” a system developed in cooperation with a private company (Clyral). Phones used had to be able to run Java. The advantage of this system, as shown in the figures below, is that the program allowed for more than simple text message replies, such as branching, skip logic, and enforced validation.

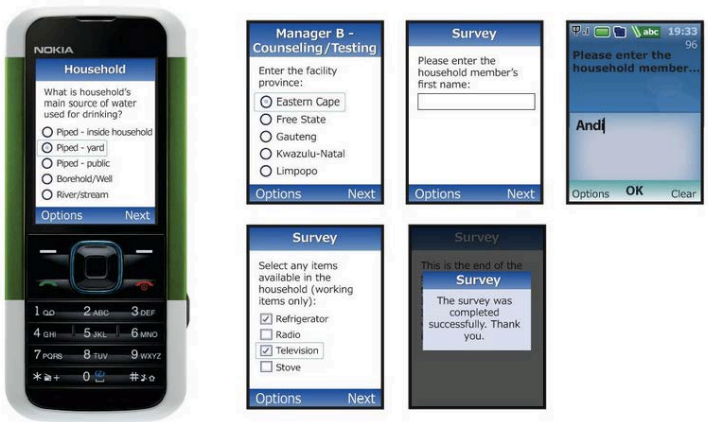

Figure 1
Screen shots of survey on the mobile phone.

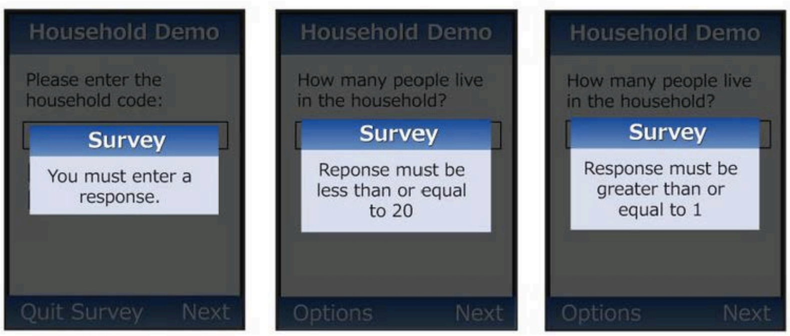

Figure 2
Enforced validation.

As mentioned, micro-surveys conducted during the election season could supplement the ANES in terms of exposure to and reactions to ongoing events in the election campaign. For example, respondents might be asked their reaction to a political debate, or to the release of a campaign ad, or to the announcement by a candidate about their running mate or what they would do if elected. Rather than relying on overall recall, respondents could indicate weekly their news consumption and exposure to campaign material, as well as their evolving political attitudes and ongoing political behavior. Pre-election micro-surveys could also ask respondents about their exposure to events on the campaign trail or to contextual events. They could be asked to report their daily impressions of the campaign. Post-election micro-surveys could collect data about reactions to election results, such as belief in the integrity of the election, trust in government, and political efficacy.

In terms of hurdles, perhaps the major one is the skewed distribution of smartphone ownership among the U.S. public. Thus, micro-surveys that could be administered using feature phones would have a much higher likelihood of being available to all potential participants. On the other hand, micro-surveys by smartphone would allow for more complex data collection such as that used in the study in South Africa. Other challenges would be the possibility that too many micro-surveys prior to the main surveys may affect response rates and then there is the cost of incentivizing respondents. We do not believe there would be any IRB difficulties here.

Regardless, studies from the health care field suggest asking ANES respondents to respond to micro-surveys throughout the campaign would likely be successful, particularly if participants are in some way compensated for their participation. Next steps might include pre-testing to determine the rate at which we would need to compensate participants, and how that rate might change for fewer and larger numbers of items/micro-surveys. It also could be explored whether it would be plausible to not just employ micro-surveys from the pre-election wave on but actually to get in the field during the summer prior to the election.

Another next step might be making a decision about whether the micro-surveys should be conducted using text messaging only, allowing for recruitment from a broader range of U.S. adults, or if the greater flexibility of an app-based tool is worth the tradeoff in terms of limiting participation to those with smartphones.

Physiology Readings

Aarøe, Lene, Michael Bang Peterson, Kevin Arceneaux. "The Behavioral Immune System Shapes Political Intuitions: Why and How Individual Differences in Disgust Sensitivity Underlie Opposition to Immigration." *American Political Science Review*, forthcoming.

Ahn, Woo-Young et al. 2014. "Nonpolitical Images Evoke Neural Predictors of Political Ideology." *Current Biology* 24(22): 2693–99.

Coe, C.M., K.S. Canelo, K. Vue, M.V. Hibbing, and S.P. Nicholson. Physiology of Framing Effects: Threat Sensitivity and the Persuasiveness of Political Arguments. *Journal of Politics*, forthcoming.

Hibbing, John R, Kevin B Smith, and John R Alford. 2014. "Differences in Negativity Bias Underlie Variations in Political Ideology." *Behavioral and Brain Sciences* 37(03): 297–307.

Mutz, D. C., & Reeves, B. (2005). The new videomalaise: Effects of televised incivility on political trust. *American Political Science Review*, 99(1), 1–15.

Oxley, D., Smith, K., Alford, J., Hibbing, M., Miller, J., Scalora, M., Hatemi, P., & Hibbing, J. (2008). Political attitudes vary with physiological traits. *Science*, 321(5896), 1667–1670.

Renshon, J., Lee, J. J. and Tingley, D. (2015), Physiological Arousal and Political Beliefs. *Political Psychology*, 36: 569–585.

Soroka, S. N. 2014. *Negativity in Democratic Politics: Causes and Consequences*. New York: Cambridge University Press.

Implicit Measure Readings

Recent Reviews

Bertram Gawronski, Jan De Houwer (In press) "Implicit Measures in Social and Personality Psychology" Chapter to appear in: H. T. Reis, & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (2nd edition). New York: Cambridge University Press.

Glaser, Jack and Christopher Finn (2013) "How and Why Implicit Attitudes Should Affect Voting" P.S. July.

Krosnick, Jon and Arthur Lupia (2008). "Decisions Made about Implicit Attitudes Measurement in the 2008 American National Election Studies." September. Report.

Nosek, Brian et. al. (2011) "Implicit social cognition: from measures to mechanisms" *Trends in Cognitive Sciences*. Vol 15. No 4.

Uhlmann, Eric et. al. (2012). Getting Explicit About the Implicit: A Taxonomy of Implicit Measures and Guide for Their Use in Organizational Research. *Organizational Research Methods* 15(4) 553-6

Some Political Applications

Albertson, B. (2011). Religious appeals and implicit attitudes. *Political Psychology*, 32(1), 109-30.

Arcuri L, Castelli L, Galdi S, Zogmaister C, Amadori A (2008) Predicting the vote: Implicit attitudes as predictors of the future behavior of decided and undecided voters. *Political Psychology* 29: 369–387

Duran, Nicholas, Stephen P. Nicholson, and Rick Dale. "Simultaneous Attraction and Resistance to Political Conspiracies in a Real-Time Decision-Tracking Task." Working paper.

Ditonto, Lau, and Sears. "AMPing Racial Attitudes: Comparing the Power of Explicit and Implicit Racism Measures in 2008" *Political Psychology*, Vol. 34, No. 4, 2013

Galdi S, Arcuri L, Gawronski B (2008) Automatic mental associations predict future choices of undecided decision-makers. Science 321: 1100–1102.

Hansford, T.G., Intawan, C. & Nicholson, S.P. 2017. Snap Judgment: Implicit Perceptions of a (Political) Court." *Political Behavior*, forthcoming.

Intawan, C. & Nicholson, S.P. (2016). My Trust in Government is Implicit: Automatic Trust in Government and System Support. Paper presented at Exploring New Frontiers, Forging New Synergies: Bolstering the Links Between Bio-Politics and Political Psychology conference, University of California, Merced.

Iyengar, S., & Westwood, S. J. (2015). Fear and loathing across party lines: New evidence on group polarization. *American Journal of Political Science*, 59(3), 690-707.

Kam, Cindy D. 2007. "Implicit Attitudes, Explicit Choices: When Subliminal Priming Predicts Candidate Preference." *Political Behavior* 29(3): 343-367.

Lodge, M., & Taber, C. S. (2013). *The rationalizing voter*. New York: Cambridge University Press.

Malte Friese, Colin Tucker Smith, Thomas Plischke, Matthias Bluemke, Brian A. Nosek 2012 "Do Implicit Attitudes Predict Actual Voting Behavior Particularly for Undecided Voters?" PLOS

Mo, C. H. (2015). The consequences of explicit and implicit gender attitudes and candidate quality in the calculations of voters. *Political Behavior*, 37(2), 357-395.

Pérez, E. O. (2016). *Unspoken politics: Implicit attitudes and political thinking*. New York: Cambridge University Press.

Raccuia L (2016) Single-Target Implicit Association Tests (ST-IAT) Predict Voting Behavior of Decided and Undecided Voters in Swiss Referendums. PLoS ONE 11(10): e0163872. doi:10.1371/journal.pone.0163872

Teige-Mocigemba, S., Klauer, K. C., & Rothermund, K. (2008). Minimizing method-specific variance in the IAT: A single block IAT. European *Journal of Psychological Assessment*, 24, 237–245.

Theodoridis, A. G. (forthcoming). Me, myself, and (I), (D) or (R)? Partisan intensity through the lens of implicit identity. *Journal of Politics*.

Valentino, Nicholas A., Vincent L. Hutchings, and Ismail K. White. 2002. "Cues that Matter: How Political Ads Prime Racial Attitudes During Campaigns." *American Political Science Review* 96(1):75-90.

Biomarker Readings

Dykema, Jennifer, Kerryann DiLoreto, Kenneth D. Cores, Dana Garbarski, and Jeremy Beach. 2017. "Factors Associated with Participation in the Collection of Saliva Samples by Mail in a Survey of Older Adults." *Public Opinion Quarterly* 81: 57-85.

Economist. 2016. "Illness as indicator: Local health outcomes predict Trumpward swings." http://www.economist.com/news/united-states/21710265-local-health-outcomes-predict-trumpward-swings-illness-indicator

McClure, Heather H., J. Josh Snodgrass, Charles R. Martinez Jr., Erica C. Squires, Roberto A. Jiménez, Laura E. Isiordia, J. Mark Eddy, Thomas W. McDade, and Jeon Small. 2015. "Stress, Place, and Allostatic Load Among Mexican Immigrant Farmworkers in Oregon." *Journal of Immigrant and Minority Health* 17: 1518-1525.

McDade, Thomas W. 2010. "The state and future of blood-based biomarkers in the Health and Retirement Study."

McDade, Thomas W. 2014. "Development and Validation of Assay Protocols for use with Dried Blood Spot Samples." *American Journal of Human Biology* 26: 1-9.

McDade, Thomas W., Sharon Williams, and J. Josh Snodgrass. 2007. "What a Drop of Blood Can Do: Dried Blood Spots as a Minimally Invasive Method for Integrating Biomarkers into Population-Based Research." *Demography* 44: 899-925.

McDermott, Rose, Chris Dawes, Elizabeth Prom-Wormley, Lindon Eaves, and Peter K. Hatemi. 2013. "*MAOA* and Aggression: A Gene-Environment Interaction in Two Populations." *Journal of Conflict Resolution* 57: 1043-1064.

Pacheco, Julianna, and Jason Fletcher. 2015. "Incorporating Health into Studies of Political Behavior: Evidence for Turnout and Partisanship." *Political Research Quarterly* 68: 104-116.

Putnam, Robert D. 2000. *Bowling Alone: The Collapse and Revival of American Community*. New York: Simon and Schuster.

Rahn, Wendy M., and Sarah E. Gollust. 2015. “The Bodies Politic: Chronic Health Conditions and Voter Turnout in the 2008 Election.” *Journal of Health Politics, Policy and Law* 40: 1115-1155.

Web Browsing Behavior Readings

Dilliplane, Susanna, Seth K. Goldman and Diana C. Mutz. 2013. Televised Exposure to Politics: New Measures for a Fragmented Media Environment. *American Journal of Political Science* 57:236-248.

Gentzkow, Matthew, and Jesse Shapiro. 2011. Ideological Segregation Online and Offline. *Quarterly Journal of Economics* 126: 1799–1839

Goel, Sharad, Jake M. Hofman, and M. Irmak Sirer. 2012. Who Does What on the Web: A Large Scale Study of Browsing Behavior. Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media, 1–8. ACM.

Goel, Sharad, Flaxman, Seth and Justin Rao. 2016. Filter Bubbles, Echo Chambers, and Online News
Consumption. *Public Opinion Quarterly*, 80: 298–320.

Hannak, Aniko, Piotr Sapiezynski, Arash M. Kakhki, Balachander Krishnamurthy, David Lazer, Alan Mislove, and Christo Wilson. 2013. Measuring Personalization of Web Search. Proceedings of the 22nd International Conference on World Wide Web, 527–38. International World Wide Web Conferences Steering Committee.

Prior, Markus. 2013. The challenge of measuring media exposure. *Political Communication* 30:620–634.

Social Media Readings

Allcott, Hunt, and Matthew Gentzkow. 2017. *Social Media and Fake News in the 2016 Election*. No. w23089. National Bureau of Economic Research, 2017.

Azari, J.R., 2016. How the news media helped to nominate Trump. *Political Communication*, *33*(4), pp.677-680.

Barberá, Pablo. 2015. “Birds of the Same Feather Tweet Together. Bayesian Ideal Point Estimation Using Twitter Data”, *Political Analysis*, 23(1): 42-58.

Barberá, Pablo, Richard Bonneau, John T. Jost, and Jonathan Nagler, and Joshua A. Tucker. “Tweeting from Left to Right: Is Online Political Communication More Than an Echo Chamber”. 2015. *Psychological Science*, 26(10):1531-42.

Bode, L., K. Dalrymple and D. Shah (2011), *Politics in 140 Characters or Less: Campaign Communication, Network Interaction, and Political Participation on Twitter*, paper presented at the annual meeting of the American Political Science Association, Seattle, 1-4 September.

Enli, Gunn. 2017. "Twitter as arena for the authentic outsider: exploring the social media campaigns of Trump and Clinton in the 2016 US presidential election." *European Journal of Communication* 32.1 (2017): 50-61.

Larson, Jennifer, Jonathan Nagler, Jonathan Ronen, and Joshua A. Tucker. 2017. “Social Networks and Protest Participation: Evidence from 130 Million Twitter Users”, *manuscript.*

Lawrence, R.G. and Boydstun, A.E., 2016. What We Should Really Be Asking About Media Attention to Trump. Political Communication, pp.1-4.

Munger, Kevin Richard Bonneau, Patrick Egan, Jonathan Nagler, Jonathan Ronen, and Joshua A. Tucker. 2017. “Learning (and Unlearning) from the Media and Political Parties: Evidence from the 2015 UK Election”, *manuscript*.

Wells, C., Shah, D.V., Pevehouse, J.C., Yang, J., Pelled, A., Boehm, F., Lukito, J., Ghosh, S. and Schmidt, J.L., 2016. How Trump drove coverage to the nomination: Hybrid media campaigning. *Political Communication*, *33*(4), pp.669-676.

Vaccari, Cristian, Augusto Valeriani, Pablo Barberá, Richard Bonneau, John T. Jost, and Jonathan Nagler, and Joshua A. Tucker. 2013. "Social Media and Political Communication: A survey of Twitter users during the 2013 Italian general election", *Italian Political Science Review*. XLIII(3): 381-410.

Micro-Surveys Readings

Anhøj, Jacob, and Claus Møldrup. 2004. “Feasibility of collecting diary data from asthma patients through mobile phones and SMS (short message service): response rate analysis and focus group evaluation from a pilot study.” *Journal of medical Internet research* 6 (4): e42.

Cho, Jae-Hyoung, Hye-Chung Lee, Dong-Jun Lim, Hyuk-Sang Kwon, and Kun-Ho Yoon. 2009. “Mobile communication using a mobile phone with a glucometer for glucose control in Type 2 patients with diabetes: as effective as an Internet-based glucose monitoring system.” *Journal of Telemedicine and Telecare* 15, (2): 77-82.

Cocosila, Mihail, Norm Archer, R. Brian Haynes, and Yufei Yuan. 2009. “Can wireless text messaging improve adherence to preventive activities? Results of a randomised controlled trial.” *International journal of medical informatics* 78 (4): 230-238.

Ferrer-Roca, Olga, A. Cardenas, A. Diaz-Cardama, and P. Pulido. 2004. “Mobile phone text messaging in the management of diabetes.” *Journal of telemedicine and telecare* 10, (5): 282-285.

Ostojic, Vedran, Branimir Cvoriscec, Sanja Barsic Ostojic, Dimitry Reznikoff, Asja Stipic-Markovic, and Zdenko Tudjman. 2005. “Improving asthma control through telemedicine: a study of short-message service.” *Telemedicine Journal & E-Health* 11, (1): 28-35.

Rodgers, Anthony, Tim Corbett, Dale Bramley, Tania Riddell, Mary Wills, Ruey-Bin Lin, and Mark Jones. 2005. “Do u smoke after txt? Results of a randomised trial of smoking cessation using mobile phone text messaging.” *Tobacco control* 14 (4): 255-261.

Shapiro, Jennifer R., Stephanie Bauer, Robert M. Hamer, Hans Kordy, Dianne Ward, and Cynthia M. Bulik. 2008. “Use of text messaging for monitoring sugar-sweetened beverages, physical activity, and screen time in children: a pilot study.” *Journal of nutrition education and behavior* 40 (6): 385-391.

Tomlinson, Mark, Wesley Solomon, Yages Singh, Tanya Doherty, Mickey Chopra, Petrida Ijumba, Alexander C. Tsai, and Debra Jackson. 2009. “The use of mobile phones as a data collection tool: A report from a household survey in South Africa.” *BMC Medical Informatics and Decision Making* 9 (1): 51.

What We Did With Social Media in the 2016 ANES

We did not collect any Twitter data in conjunction with the 2016 ANES.

However, the ANES team did work together with the NYU Social Media and Political Participation (SMaPP) lab (which is co-Directed by Tucker) to build an App that would allow us to collect a limited subsection of a respondent's Facebook Data. That App was built after several meetings with Facebook personnel to discuss our plans (and, indeed, with their encouragement), and was approved by FB despite not providing a product to users. The cost of building the app was minimal ($2000 in programming expenses) and was jointly funded by the SMaPP lab and the ANES.[19]

The option to log in to the App and allow us to collect the data was given in *post* election wave of the survey. Details of response rates, etc. are pending. However, a similar study done by the NYU SMaPP Social Media and Election survey achieved an over 30% response rate.[20]

The (general) information that we collected from FB fall into the following categories:

- General Likes: Books, Music, Movies, Games, Television, and Likes are all in this category
- Posting: Everything the respondent has written or been tagged in. Includes Feed, Posts, and Tagged
- Profile: Profile itself is its own category, but it includes subcategories such as favorite_teams, favorite_players, inspirational_people, sports, quotes, political, languages, religion, gender and more (these seemed to be the most relevant ones).
- ANES Related: permissions and respondent_id.

A more detailed description of the data we collected can be found in next appendix, "ANES 2016 Facebook Data Collection: File Content Description**"**

The next step is to turn the raw FB data into variables that can be appended to the ANES file. ANES staffers and students are working on this task, but to date the following (tentative!) variables have been created:

[19] The joint funding model was used because we simultaneously built a similar App for the *SMaPP Social Media and the 2016 US Election Panel Survey*; the only important differences between the two Apps were (1) the content of the landing page and (2) the server to which the data was sent. This is relevant because in the future it could provide a model for deferring the (admittedly already low) cost of programming associated with building (maintaining?) an App. Perhaps more importantly, it might provide an opportunity for the ANES to provide a public good to others in the scholarly community. To be very clear, though, since collection the FB data from ANES respondents have been solely in hands of the ANES on servers to which no one in the SMaPP lab has access.

[20] Matt DeBell is in the process of extracting the exact numbers from the ANES pilot. There was, unfortunately, a problem downloading the data this year due to the fact that FB changed its API while the survey was in the field, which means the percentage of people from whom we actually collected data is lower than the percentage of people who agreed to allow us to collect data. The good news is this means it is a more promising method looking forward than the actual amount of data we got this year would suggest. Response rates from the *SMaPP Social Media* election survey – not afflicted by the same API problem -- suggests that expecting a response rate of close to one-third of ANES respondents would not be unreasonable.

resp_id: respondent ID

likes: Full page name of Media outlets or personalities the respondent "likes" (along with pages that have simple political words like "liberal" or "Republican" in their name). If no political likes, “no political likes”. This is a first pass and catches some non-political likes (such as “Banana Republic”)[21]

numposts: Number of posts per year from 2007 onwards. NA means the respondent didn't have any posts that year. Variable names such as numposts_2007

numposts_tot: Total number of posts respondent made since 2007.

permission_action.news: action/news permission. 1=granted, 0 = not granted

permission_likes: likes permission. 1=granted, 0 = not granted

permission_posts: posts permission. 1=granted, 0 = not granted

permission_religion_politics: religion/politics permission. 1=granted, 0 = not granted

public_profile: public profile permission. 1=granted, 0 = not granted

political_label: User’s self-defined political label. If not provided, “not listed”

wordcount: Counts of specific political words in user’s posts. For example, if “obama” is found 3 times in all the posts, the value for variable wordcount_obama is 3. [22]

Please note that this process is still very much ongoing.

Next Steps:

[21] BBC World News","CNN","Fox News","New York Times", "ABC News","Huffington Post","NBC News","The Economist","Yahoo! News","The Guardian","USA TODAY", "Democra", "Republic", "Liberal", "Conservati", "Breitbart", "WorldNetDaily", "Infowars", "Rush Limbaugh", "Glenn Beck", "Rachel Maddow", "Megyn Kelly", "Vox","Slate","Buzzfeed","MSNBC","Washington Post","Newsweek", "Trump","Obama","Clinton","Hillary","Pantsuit Nation

[22] "democrat","democracy","GOP","republican","barack", "obama", "mccain", "mitt", "romney", "hillary", "clinton", "trump", "elect", "campaign", "vote", "voting", "ballot", "president","congress", "supreme court", "government", "constitution", "amendment","palin", "paul ryan", "pence","biden", "china","chinese","russia","mexico","mexican","syria","syrian","israel","iraq","iran","nuclear","nuke", "oil","pipeline", "immigrant","immigration","illegal","undocumented","wall", "terror","jobs","manufacturing","race","racist","racism","police","state","economy","debt","justice","gun", "income","trade","birther","war","health", "obamacare","taxes","foreign","inequality", "crime", "military","security","environment","education","charter","school", "voucher","tpp", "benghazi", "email server", "private server", "muslim", "altright", "birth certificate", "killary", "the Donald", "abortion", "planned parenthood","baby parts" ,"classified", "sandy hook", "false flag","newtown","columbine", "9/11","Fox News","mainstream media","lamestream media","MSNBC","baby killer", "scalia", "truther", "vaccin", "vax", "gay","homosexual","trans","transgender","transsexual"

There are perhaps limitless numbers of next steps that could be taken in this regard, but the obvious ones are (a) to continue the variable creation process and (b) attempt to use some of these new variables in data analysis. More broadly, to the extent that the feasibility of collecting Facebook Data of ANES respondents has been demonstrated, steps should be taken to (c) begin thinking conceptually about the types of questions we would like to use this data to answer[23] as well as (d) showcase the data we've collected to larger numbers of people (beginning with the full ANES board) to consider whether any of the concerns elucidated earlier in this report should prevent us from ultimately including this data as part of the 2016 ANES release (i.e., or whether it should simply be treated as an internal pilot study) and/or going forward with this as a regular part of the ANES in the future.

Conclusions:

Although it is a cliché, incorporating social media data into the ANES presents both challenges and opportunities. We have tried to be explicit as possible about many of those challenges in this report, but at the end of the day the opportunity to augment existing surveys consisting almost entirely of self-reported data with "objective" measures of political behavior online at a time

[23] For example, Munger et al. (2017) looks at whether exposure to information about politics on Twitter during the 2015 British election campaign led respondents to increase or decrease levels of political knowledge, as well as the extent to which the source of the information (media vs. politicians; left vs. right wing) mediated that effect.

ANES 2016 Facebook Data Collection: File Content Description
(PRELIMINARY REPORT)

Overview:

There are several groups of data in this file. Most of the categories in the .json files did not have data attached to them. The ones that did fall into the following groups

- General Likes: Books, Music, Movies, Games, Television, and Likes are all in this category
- Posting: Everything the respondent has written or been tagged in. Includes Feed, Posts, and Tagged
- Profile: Profile itself is its own category, but it includes subcategories such as favorite_teams, favorite_players, inspirational_people, sports, quotes, political, languages, religion, gender and more (these seemed to be the most relevant ones).
- ANES Related: permissions and respondent_id.

No picture URLs or link URLs are included in the posts. Cover photo (in profile) technically gives a URL, but attempts to follow it yield "URL signature expired ".

Almost all variables are free entry – nothing is multiple choice with the exception of permissions and timezone. Gender is nominally multiple choice with a "custom" free entry option.

Detailed Data:

This list has all of the possible categories of data from the json facebook files. Blank categories are listed towards the end.

books: Books or genres of literature the respondent has liked.

Includes three subcategories:

- created_time: date and time the like was made
- name: name of the book/genre
- id: id of the action

feed: Everything the respondent has posted to their wall (links, statuses, etc), as well as anything the respondent was tagged in, since 2008. Except in rare cases (where the respondent kept the link url as the text in the message box), we do <u>not</u> have the actual link url. We only know that the respondent shared a link, but nothing about the link itself.

This contains four subcategories:

- created_time: date and time the item was posted

- message: the words typed by the respondent. Can be blank when respondent is sharing a link or photo or something else
- story: whether or not the respondent shared something in this post (such as a link or video), or if someone tagged the respondent. Can be blank (if item is just a text post)
- id: id of the post

games: Games (video or board) the respondent has liked.

Contains 3 subcategories:

- created_time: Date and time the game was liked
- name: name of the game
- id: id of the action

likes: All pages the respondent has liked. Includes all subcategories of likes (books, games, etc), plus likes that do not fit in a subcategory

Contains 3 subcategories

- created_time: date and time of the like
- name: name of the page liked. Only the name, not the URL.
- id: id of the action

movies: Movies liked by the respondent

Contains three subcategories

- created_time: date and time of the like
- name: name of the movie liked
- id: id of the action

music: Music liked by the respondent

Contains three subcategories

- created_time: date and time of the like
- name: name of the band or music album
- id: id of the action

permissions: Data that the respondent has agreed to share with the ANES

Contains 2 subcategories. Each respondent has 5 entries in the permissions category

- permission: name of the permission.
  The 5 permission names are:
    - user_religion_politics
    - user_likes

    - user_posts
    - user_action.news
    - public_profile
- status: whether the permission has been granted or not

posts: Everything the respondent has posted to their wall (links, statuses, etc) since 2008. Almost the same as 'feed: except for posts where the respondent was tagged by someone else (which appear in feed but not posts). Except in rare cases (where the respondent kept the link url as the text in the message box), we do not have the actual link url. We only know that the respondent shared a link, but nothing about the link itself.

This contains four subcategories:

- created_time: date and time the item was posted
- message: the words typed by the respondent. Can be blank when respondent is sharing a link or photo or something else
- story: whether or not the respondent shared something in this post (such as a link or video). Can be blank (if item is just a text post)
- id: id of the post

profile: Information about the user

Contains 35 subcategories:

- last_name: user's last name
- locale: country
- installed: unknown to ANES (i.e., we do not yet know what this is)
- currency: whether the person pays in USD or other currency
- third_party_id: some form of id number
- favorite_teams: favorite sports teams
- favorite_athletes: favorite athletes
- inspirational_people: people named as "inspirational" by the user
- timezone: timezone written as the difference from utc
- updated_time: last time profile was updated
- id: user id
- first_name: first name
- middle_name: middle name
- name_format: which name goes first
- political: political preferences as described by the user. There is a box on the facebook profile in which the user can fill in their political beliefs. Answers can range from 'Conservative' to 'Democrat' to 'Peace and Love' to 'I vote for the person, not the party'. It's completely free-form. As the user is typing in their answer, they are given a drop down list of suggested political pages that include

the words they're typing – however, they do not have to pick one of those pages, they can type whatever they want.

- sports: sports the user participates in
- languages: languages spoken by the user
- religion: religious beliefs as described by the user
- payment_pricepoints: unknown to ANES
- security_settings: user's security settings
- is_verified: unknown to ANES
- is_shared_login: unknown to ANES
- test_group: unknown to ANES
- link: link to the user's profile
- verified: unknown to ANES
- name: full name
- quotes: favorite quotes provided by the user
- install_type: unknown to ANES
- gender: gender of user
- cover: cover photo. Contains subcategories for the image and the positioning of the image
- devices: devices this account has been accessed on
- viewer_can_send_gift: unknown to ANES
- context: unknown to ANES. Has a subcategory called mutual_likes
- video_upload_limits: unknown to ANES

respondent_id: ANES case ID

tagged: All of the posts in which this person was tagged

Contains four subcategories:

- message: the message that was written in which the user was tagged
- story: if the user was tagged in a photo, or at a place, or via sharing a link, the person who tagged the user is included in the story subcategory. However, if the user was tagged in a status, the story subcategory is blank and no info on the person who tagged Jane Doe is provided in Jane Doe's file, although we still have the message subcategory. Examples:
    - An event when John Doe shares a photo and tags Jane Doe and 4 other friends. We have Jane Doe's file:
        - {message: "dinner with the fam"
          story: "John Doe was with Jane Doe and 4 others"
          id: some number
          tagged_time: some time}

- An event when John Doe writes a status and tags Jane Doe. We have Jane Doe’s file:
  - {message: “Happy Anniversary, Jane Doe!”
    id: some number
    tagged_time: some time}

- tagged_time: the time the user was tagged
- id: the id of the action

television: Television liked by the respondent

Contains three subcategories

- created_time: date and time of the like
- name: name of the tv show or channel
- id: id of the action

Blank Fields:

accounts, achievements, ad_studies, adaccountgroups, adaccounts, adcontracts, admined_groups, adnetworkanalytics, albums, applications, apprequestformerrecipients, apprequests, brand_teams, business_activities, businesses, businesssettinglogs, checkins, commission_splits, conversations, curated_collections, domains, events, family, favorite_requests, friendlists, friendrequests, friends, groups, home, ids_for_business, inbox, insights, integrated_plugin_feed, invitable_friends, leadgen_forms, locations, notifications, notify_me, objects, outbox, ownerapps, payment.subscriptions, payment_transactions, payments, personal_ad_accounts, photos, platformrequests, pokes, privacy_options, promotable_domains, promotable_events, ratings, request_history, scores, screennames, session_keys, stream_filters, taggable_friends, tagged_places, threads, updates, videos,