

**SUPPORTING STATEMENT**  
**U.S. Department of Commerce**  
**National Oceanic & Atmospheric Administration**  
**Weather and Society Survey and Using Quick Response Surveys to Build a Public Perception and**  
**Response Database**  
**OMB Control No. 0648-XXXX**

**SUPPORTING STATEMENT PART B**

**1. Describe (including a numerical estimate) the potential respondent universe and any sampling or other respondent selection method to be used. Data on the number of entities (e.g., establishments, State and local government units, households, or persons) in the universe covered by the collection and in the corresponding sample are to be provided in tabular form for the universe as a whole and for each of the strata in the proposed sample. Indicate expected response rates for the collection as a whole. If the collection had been conducted previously, include the actual response rate achieved during the last collection.**

*The Weather and Society Dashboard Project*

The target population for the Wx Survey is all adults (age 18+) who reside in the U.S. Approximately 255,200,373 persons populate this universe. A web-based opt-in sampling approach will be used to construct samples of survey participants that match the demographic and geographic characteristics of the target population. Census estimates of sex, age, ethnicity, race, and NWS region will be used to define the demographic and geographic strata of the target population. Quota sampling will be used to construct samples that match these strata. Table 1 provides an example list of Census estimates and corresponding quotas from the 2021 Severe Weather and Society Survey, a previous data collection in this series.

Table 1: Example List of Census Estimates and Corresponding Quotas

	Census Estimates %	Quotas n	Participants n (%)
<b>Sex</b>			
Female	51.3%	770	795 (51.3%)
Male	48.7%	730	754 (48.7%)
<b>Age</b>			
18 to 24	12.0%	181	186 (12.0%)
25 to 34	18.0%	270	279 (18.0%)
35 to 44	16.3%	244	253 (16.3%)
45 to 54	16.4%	246	253 (16.3%)
55 to 64	16.7%	250	259 (16.7%)
65 and up	20.6%	309	319 (20.6%)
<b>Ethnicity</b>			
Hispanic	16.3%	244	253 (16.3%)
Non-Hispanic	83.7%	1256	1296 (83.7%)
<b>Race</b>			
White	77.9%	1168	1207 (77.9%)
African American	13.0%	195	201 (13.0%)
Asian	5.9%	88	91 (5.9%)
Other Race	3.2%	49	50 (3.2%)
<b>NWS Region</b>			
Eastern	31.6%	474	492 (31.6%)
Southern	27.1%	406	420 (27.1%)

Central	20.7%	310	321 (20.7%)
Western	20.6%	310	316 (20.6%)
<b>TOTAL</b>	<b>255,200,373</b>	<b>1500</b>	<b>1549</b>

Due to the opt-in and fluent nature of the sampling process, the research team will not track survey invitations, making it impossible to calculate ordinary survey response rates (response rate = number of people who complete the survey / number of people who are invited to participate in the survey). In place of response rates, researchers who use web-based opt-in sampling approaches typically report completion rates, which indicate the proportion of people who complete the survey once they start (completion rate = number of people who complete the survey / number of people who start the survey) (Callegaro and DiSogra 2008). In previous data collections, such as the 2021 Severe Weather and Society Survey, approximately 70% of people who began the survey went on to complete the survey. The research team will use this as a benchmark to measure success on this dimension of the study.

### *Quick Response Surveys Project*

The target population for the Quick Response Surveys is adults over 18 who have experienced a specific flash flood, tornado/high wind, or winter weather event. The potential respondent universe is estimated to be 47,169,776 persons (Table 2). This number represents the population of the County Warning Areas (CWAs) covered by National Weather Service (NWS) forecast offices that intend to participate in this research. Surveys will be targeted to people residing in the counties impacted by a specific weather hazard. Respondents will be surveyed 1 day to 3 weeks after a select severe or winter weather event occurs to limit recall bias by the respondent and provide the NWS with rapid results. Respondent selection will focus on stratifying the population of the counties in question by race, gender, and age groups.

Table 2. Potential Respondent Universe

<b>Forecast Office Name</b>	<b>CWA Total Population</b>
Amarillo AMA	427,550
Boston BOX	8,942,549
Fort Worth FWD	8,979,846
Little Rock LZK	1,679,275
Mt. Holly Philly PHI	12,055,800
Nashville OHX	2,690,557
Peachtree Atlanta FFC	8,247,031
Pittsburgh PTZ	3,728,392
San Angelo SZT	418,776
<b>Total</b>	<b>47,169,776</b>

The primary sampling method will be a convenience sample, and an internet-based probability sample will also be used. Surveys will be disseminated on participating Weather Forecast Offices' (WFOs) social media pages. In addition, each WFO will enlist core partners (local weather, Emergency Managers (EMs), community groups) to share the survey links on their social media pages and email listservs to reach broader segments of the population. The goal is to create a high quality, geodemographically stratified convenience sample, called a quota sample.

We estimate a conservative 0.2% response rate based on our previous surveys in the Dallas Fort Worth metroplex where a similar approach was used. We apply this response rate to the estimated followers or subscribers of the different types of social media sites to compute the number of respondents. See Table 2. We use Facebook followers as a proxy for the number of people reached.

Table 3 shows the number of followers for 1) each WFO Facebook page (rounded to the nearest 10,000); 2) other public safety/weather social media pages (e.g., Broadcast meteorologists, EMs, and sheriffs); and 3) community pages (e.g., NextDoor, faith-based organizations, schools). Based on a review of these pages, we estimate that Non-NWS Weather pages have approximately the same number of followers, and community groups represent about half of the NWS followers. Assuming a 0.2% response rate across all page types, the estimated number of respondents is calculated for Year 1 collection (6 surveys for each WFO) and Year 2 collection (12 surveys for each WFO) **for a total of 99,000 respondents.**

Table 3. Convenience Sample Estimated Respondents

Survey Dissemination	WFO	NWS Facebook 1	Non-NWS Weather Facebook (EM, TV mets) 2	Community Facebook (Nextdoor, Faith, etc.) 3	Total Reach	Response Rate 4	Responses per survey	Year 1 Responses 6 surveys	Year 2 Responses 12 surveys	Total Estimated Responses
<b>Followers Estimated</b>	FWD	210,000	210,000	105,000	525,000	0.20%	1,050	6,300	12,600	18,900
	BOX	140,000	140,000	70,000	350,000	0.20%	700	4,200	8,400	12,600
	PHI	140,000	140,000	70,000	350,000	0.20%	700	4,200	8,400	12,600
	FFC	70,000	70,000	35,000	175,000	0.20%	350	2,100	4,200	6,300
	LZK	140,000	140,000	70,000	350,000	0.20%	700	4,200	8,400	12,600
	AMA	80,000	80,000	40,000	200,000	0.20%	400	2,400	4,800	7,200
	SJT	40,000	40,000	20,000	100,000	0.20%	200	1,200	2,400	3,600
	PBZ	110,000	110,000	55,000	275,000	0.20%	550	3,300	6,600	9,900
	OHX	170,000	170,000	85,000	425,000	0.20%	850	5,100	10,200	15,300
<b>Total</b>		<b>1,100,000</b>	<b>1,100,000</b>	<b>550,000</b>	<b>2,750,000</b>		<b>5,500</b>	<b>33,000</b>	<b>66,000</b>	<b>99,000</b>

<sup>1</sup> Number of Facebook followers for each NWS WFO, rounded to nearest 10,000.

<sup>2</sup> Estimated non-NWS Weather followers for each WFO's County Warning Area (Media, Emergency Management, Sherriff, etc.)

<sup>3</sup> Estimated Community-based Facebook followers for each WFO's County Warning Area

<sup>4</sup> Response rate is estimated to be 0.2% based on similar surveys conducted by the grantees, which were distributed in the DFW Metroplex across select social media pages (weather, media, community)

<sup>5</sup> If each WFO sends out 6 surveys in Year 1, and 12 surveys in Year 2, the total number of responses over the 2-year collection period is estimated to be 99,000. This is annualized to 33,000.

Although a probability-based sample has historically been the gold standard for this type of survey, it would be cost prohibitive to conduct the multiple surveys required for this effort. Therefore, an important topic for this research is how to collect and weight the quota sample to allow for generalizability of results (see question 2). In order to address this topic, we will conduct a sampling experiment where we will compare a probability-based sample and a convenience sample for 1 – 3 events (see question 3). The probability-based sample is being purchased from Ipsos KnowledgePanel. Ipsos uses probability sampling to recruit a representative U.S. online panel. Panel members are recruited using probability selection algorithms for both random-digit dial (RDD) telephone and address-based sampling (ABS) methodologies. The Ipsos response rate is expected to be 55-60% (OMB). Table 4 estimates the number of expected respondents. IPSOS will guarantee 385 responses per survey. Therefore, at an expected response rate of 55%, IPSOS will recruit up to 700 people per survey.

Table 4. Probability Sample Estimated Respondents

Information Collection	Population or Potential Respondents Universe (a)	Number of Respondents Selected (b)	Maximum Ipsos surveys (c)	Expected Completion Rate (d)	Expected Number of Respondents (e) =(b) x (c) x (d)
Ipsos KnowledgePanel	All of the population within the selected 9 NWS WFO County Warning Areas = 47,169,776	700	3	55%	1,155
<b>Total</b>					<b>1,155</b>

## References

OMB 2021. Available from <https://omb.report/icr/202105-0990-004/doc/111804401>

Pew Research 2020. Available from <https://www.pewforum.org/2020/01/22/methodology-31/>

Pew Research 2021. Available from <https://www.pewresearch.org/internet/2021/04/07/social-media-U.S.e-in-2021/>

IPSOS 2015. Documentation for Human Subject Review Committees: IPSOS company information, past external review, confidentiality, and privacy protections for panelists. Available from <https://www.ipsos.com/sites/default/files/Documentation%20for%20IRBs.pdf>

We Are Social 2021. Digital 2021: Global October Snapshot Report. Available from <https://wearesocial.com/blog/2021/10/social-media-users-pass-the-4-5-billion-mark/>

## 2. Describe the procedures for the collection of information including:

### *Weather and Society Dashboard Project:*

#### 2.1 Statistical methodology for stratification and sample selection

The study will recruit survey participants by contracting with companies that maintain large panels of people who sign up to complete internet surveys, such as Qualtrics and Survey Sampling International. During the recruitment phase, the team will use demographic quotas for sex, race, ethnicity, and geographic region to ensure that survey respondents match the target population on these dimensions. To initiate the quota sampling process, the team begin each data collection by compiling Census estimates to identify the target number of respondents for each demographic/geographic category. According to Census estimates in Table 1, for example, 51.3% of adults were female when that survey began. The target sample size of that survey was 1,500, so the female quota was 770 ( $0.513 * 1500 = 769.5$ ). After identifying these quotas, the team will work with Qualtrics to execute a phased sampling process. We will begin by identifying eligible panelists and inviting them to participate in a survey through direct contact, primarily by emails and notifications on phone applications. At first, we will send invitations to a large and diverse group. As panelists complete the survey, we will carefully monitor the quotas and send invitations to select groups based on preliminary imbalances. For example, if most of the early participants are female, we will target male panelists when sending select follow-up invitations. In addition to targeting invitations, we will restrict eligibility criteria when quotas are filled or nearly filled. For example, when 770 female respondents complete the survey, females will no longer be eligible to participate. As shown in Table 1, this process results in a mix of survey participants that closely match demographic and geographic characteristics of the target population.

In addition to quotas, post-stratification weights will be used to enhance the demographic representativeness samples and generalizability of the results. Weights will be calculated for each respondent to adjust for slight imbalances in sex, age, race, and Hispanic ethnicity within each of the four NWS regions that divide the contiguous U.S. (CONUS)—the Eastern, Southern, Central, and Western regions. The weighting process will involve three steps: (1) calculate the proportion of the U.S. population that shares the demographic characteristics of each respondent (population proportion); (2) calculate the proportion of the sample that shares the demographic characteristics of each respondent (sample proportion); and (3) divide the population proportion by the sample proportion to calculate a weight for each respondent. This process will result in a survey weight for each respondent that indicates how much each case will “count” in weighted analyses. A weight that is greater than one means that a participant with a given set of demographic attributes is underrepresented in the survey sample (relative to the target population), and responses from that participant will receive greater statistical emphasis than responses from survey participants who are represented in direct proportion to the adult population. Conversely, a weight that is smaller than one means that a respondent with a given set of demographic attributes is overrepresented in the sample (relative to the target population), and responses from that participant will receive less emphasis. Weights will be calculated within NWS regions to facilitate generalization within and comparison across the regions.

#### 2.2 Estimation procedure

Following data collection, the team will utilize modern techniques in small area estimation (SAE) to estimate statistics of interest for CWAs and counties in the U.S., over time. Currently, there are two

primary SAE techniques: disaggregation and multilevel regression and poststratification (MRP). When applying disaggregation, researchers compile as many comparable datasets as possible, and then use responses from survey participants who live in the same geographic area (e.g., county) to calculate a given statistic within that area. While intuitive, disaggregation is data intensive—it requires sufficient sample size in each geographic unit to produce reliable estimates. Most large population surveys do not collect enough observations in each geographic area to produce these estimates; this is especially true in low population areas. MRP is less data intensive than disaggregation and it allows researchers to account for nesting. It uses regression analysis to identify demographic and geographic patterns in areas where data are available to produce estimates in areas where data are relatively sparse. There is an emerging consensus among survey researchers that MRP is a viable alternative to disaggregation when demographic and geographic patterns are evident in the data (Lax and Phillips 2009; Buttice and Highton 2013). As such, researchers from many different fields and agencies are using this technique to estimate a wide variety of community statistics. For example, scientists at the CDC are using MRP to estimate the prevalence of public health outcomes in census blocks, tracts, districts, and counties across the country (Zhang et al. 2014, 2015; Wang et al. 2018) and opinion analysts are using it to forecast election outcomes in U.S. states (Wang et al. 2015; Kiewiet de Jonge et al. 2018). More importantly, and of direct relevance to this study, the research team has shown that MRP techniques provide reliable estimates of forecast and warning reception, comprehension, and response (Ripberger et al. 2020).

### 2.3 Degree of accuracy needed for the purpose described in the justification

While the team will certainly strive for accuracy (i.e., precise estimates that match independent observations), this study does not require highly accurate estimates of the measures for each geographic location (i.e., 82% of people in Oklahoma County receive tornado warnings). Rather, it requires that estimates provide enough information to facilitate relative comparison (more people in Oklahoma county receive tornado warnings than Tulsa County). As previous work demonstrates, the MRP techniques this study will employ provide this level of accuracy (Ripberger et al. 2020).

### 2.4 Unusual Problems and Use of Less Frequent Data Collection Cycles

The research team does not anticipate any unusual problems that will require specialized sampling procedures. All survey data will be collected on an annual basis.

## *Quick Response Surveys Project*

### 2.1 Statistical methodology for stratification and sample selection

This research will use a demographically stratified, convenience sample, also called quota sample. Our intent is to create a high-quality convenience sample that can mitigate the selection bias issues that are common with convenience samples, and therefore enable us to use standard statistical inference approach as an approximation, as suggested in Vehovar et al. 2016.

In this project, we will introduce probability-like sample properties to our convenience sample using the following steps. First, collecting a convenience sample from multiple social media sites and email lists has been demonstrated to introduce some degree of randomness into the data because different sites address different populations. Second, using quota sampling helps to create a stratified sample which can be matched to American Community Survey (ACS) data through sample matching. Third, we will construct sample weights from the matched ACS sample with known population parameters using ranking ratio estimation and, perhaps, propensity score matching to make the sample representative of



the populations within each WFO. Lastly, we plan to evaluate these methods by comparing the quota sample results to the results from the Ipsos probability sample (See Question 3).

## 2.2 Estimation Procedure

The data collected in the survey will primarily be categorical as survey respondents select specific actions or perceptions through multiple choice questions, or they indicate the importance of varied factors using a Likert scale. Data from surveys will be aggregated at the hazard level. The primary outcome variables are: **1)** actions taken before the event (for example, Did you cancel a planned trip based on predicted flood risk?) and **2)** actions taken during the event (for example, Did you drive on a flooded roadway?). The independent variables include environmental, cognitive, situational, and demographic factors. These data will be used for two primary purposes:

1. Descriptive Statistics and tests of association. The data will be used to examine measures of central tendency, standard deviation, and distributions. It will also be used to conduct Chi-square analyses to test the association between the outcome variables and the various factors. This information will be generated on an automated basis soon after a data collection effort is complete for evaluation by NWS forecasters to understand its usefulness.
2. Logistic and Multinomial regression. Since the outcome variables are categorical or binary, non-parametric regression will be used to analyze how the different factors influence the odds of taking protective actions before or during the event. Once a model is estimated, the internal validity of the model will be evaluated using Receiver Operating Curve (ROC) analysis, looking at the area under the receiver operating curve (AUC). AUC scores indicate the ability of the model to correctly predict outcomes in data while avoiding false negative and false positive predictions. In addition, once the internal validity of the model is demonstrated, marginal probabilities will be computed for taking protective action given a specific factor. Another approach to analyzing internal validity is k-fold cross-validation. K-fold cross-validation divides a sample into  $k$  number of subsamples, each of which is tested against the larger remaining sample for similarity between analysis results. Coupling this method with AUC scoring provides a dual approach to ensuring internal validity (Kohavi 1995).

## 2.3 Degree of accuracy needed for the purpose described in the justification.

The analysis will focus on understanding how different factors influence protective action decisions. We will address the degree of accuracy needed to analyze different groups through power analysis. Power analysis evaluates the ability to detect differences between two groups in a given dataset. Put another way, it enables researchers to correctly reject the null hypothesis that two samples are taken from the same distribution. In current practice, most data collection efforts look for 0.8 power, meaning there is an 80% probability that the difference between the two groups will be detected, if it exists. Power is a function of sample size, the level of precision desired, the anticipated prevalence of the group in the sample, and effect size. Effect size measures the marginal distance between groups and is usually obtained empirically through prior studies or a pilot project (Durlak, 2009). This logic can be applied to any subgroup that we might want to analyze on its own, and the proportion can be changed to suit other hazards. The benefit is being able to target certain populations before survey collection begins.

## 2.4 Unusual Problems and Use of Less Frequent Data Collection Cycles.

There are no unusual problems requiring specialized sampling procedures and we will not use less frequent than annual data collection cycles to reduce burden. In this study, we will field 54 surveys the first year and 108 surveys the second year (i.e., 6 and 12 surveys, respectively, at each of the 9 WPOs). Each survey will address a hazard type: 1) flash floods; 2) tornado/severe thunderstorms; or 3) winter events. WFOs will be instructed to initiate surveys for the nth hazard occurrence based on the historical frequency at the WFO level.

## References

- Baker, R., Brick, J. M., Bates, N. A., Battaglia, M., Couper, M. P., Dever, J. A., ... & Tourangeau, R. (2013). Summary report of the AAPOR task force on non-probability sampling. *Journal of survey statistics and methodology*, 1(2), 90-143.
- Durlak, J. A. (2009). How to select, calculate, and interpret effect sizes. *Journal of pediatric psychology*, 34(9), 917-928.
- Ender, P. B. (2011). *STATA: Power Logistic Regression (Powerlog): Stata Statistical Software: Release 12*. College Station, TX: StataCorp LP.
- Hays, R. D., Liu, H., & Kapteyn, A. (2015). Use of Internet panels to conduct surveys. *Behavior research methods*, 47(3), 685-690.
- Kim, J. K., & Wang, Z. (2019). Sampling techniques for big data analysis. *International Statistical Review*, 87, S177-S191.
- Kohavi, R. (1995, Aug U.S.t). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai* (Vol. 14, No. 2, pp. 1137-1145).
- MacInnis, B., Krosnick, J. A., Ho, A. S., & Cho, M. J. (2018). The accuracy of measurements with probability and nonprobability survey samples: Replication and extension. *Public Opinion Quarterly*, 82(4), 707-744.
- Madigan, D., Stang, P. E., Berlin, J. A., Schuemie, M., Overhage, J. M., Suchard, M. A., ... & Ryan, P. B. (2014). A systematic statistical approach to evaluating evidence from observational studies. *Annual Review of Statistics and Its Application*, 1, 11-39.
- NOAA (2021). Storm Events Database. Available from <https://www.ncdc.noaa.gov/stormevents/>
- Vehovar, V., Toepoel, V., & Steinmetz, S. (2016). Non-probability sampling. *The Sage handbook of survey methods*, 329-345.

**3. Describe methods to maximize response rates and to deal with issues of non-response. The accuracy and reliability of information collected must be shown to be adequate for intended U.S.es. For collections based on sampling, a special justification must be provided for any collection that will not yield "reliable" data that can be generalized to the universe studied.**

### *Weather and Society Dashboard Project*

The study will maximize response rates by recruiting survey participants through contracting with companies that maintain large panels of people, such as Qualtrics and Survey Sampling International. Pre-tests on the complexity of questions and length limits will be used to increase the probability that

participants who begin the survey, complete the survey. This will reduce the burden on the public and maximize the quality of the data. In test iterations of the survey, approximately 70% of people who begin the survey go on to complete the survey. The research team will use this as a benchmark to measure success on this dimension of the study.

### *Quick Response Surveys Project*

**Response Rates.** Through Qualtrics and a Survey Dashboard that will be created as part of this project, we will monitor the demographic characteristics of the survey responses in real time compared to benchmark county-based statistics to ensure that we are getting the demographic representation needed for increased generalizability and number of responses required for statistical power. As the research progresses, we will evaluate the quality of the response and adjust frequency of reposting links, and the types of organizations that post the links to the survey. This approach is based on several studies (see Antoun et al., 2016; Perrotta et al., 2021; Vehovar et al., 2016) that found quota sampling through Facebook is a systematic strategy that can be used to obtain a stratified sample approximating American Community survey results when considering multiple sampling frames and randomization.

After we develop the process for monitoring survey response, we will ask WFOs to track the time it takes for forecasters to initiate and monitor data collection to understand the burden it will place on WFOs.

**Sampling Experiment.** In addition to ensuring representativeness in the data collection process as outlined in the section above, the quality and accuracy of the quota sample will be assessed through a sampling experiment where we compare the quota sample results and IPSOS sample results for the same weather event.

The sampling experiment in this study will consider the IPSOS sample as the benchmark given the quality of the information. We will compare our quota sample with and without demographic weights to the IPSOS sample, computing the deviation from the mean (percentage) of key variables using the absolute deviation measures above as well as chi-square tests. In addition, we will also compare the coefficients for logistic regression models based on the different samples. The sampling experiment will provide valuable information on biases that may exist in the quota sample and the source of the bias that might lead new weighting strategies or data collection strategies; the sampling experiment can also reveal population groups that might be missing or underrepresented in the quota sample; however, it may also show that the probability sample does not have sufficient power for certain kinds of analyses. Based on the results of this test, we can begin to assess the potential for inference for this data collection effort and what kinds of caveats should be associated with the data.

### **References**

Antoun, C., Zhang, C., Conrad, F. G., & Schober, M. F. (2016). Comparisons of online recruitment strategies for convenience samples: Craigslist, Google AdWords, Facebook, and Amazon Mechanical Turk. *Field methods*, 28(3), 231-246.

Callegaro, M., Baker, R. P., Bethlehem, J., Göritz, A. S., Krosnick, J. A., & Lavrakas, P. J. (Eds.). (2014). *Online panel research: A data quality perspective*. John Wiley & Sons.

CFI Group (2018a). *National Weather Service Webmonitor Results Q3 FY2018 (April – June 2018)*. Retrieved with permission from the NWS Performance Management Web Portal.

Chang, L., & Krosnick, J. A. (2009). National surveys via RDD telephone interviewing versus the Internet: Comparing sample representativeness and response quality. *Public Opinion Quarterly*, 73(4), 641-678.

Dutwin, D., & Buskirk, T. D. (2017). Apples to oranges or gala versus golden delicious? Comparing data quality of nonprobability internet samples to low response rate probability samples. *Public Opinion Quarterly*, 81(S1), 213-239.

Perrotta, D., Grow, A., Rampazzo, F., Cimentada, J., Del Fava, E., Gil-Clavel, S., & Zagheni, E. (2021). Behaviours and attitudes in response to the COVID-19 pandemic: insights from a cross-national Facebook survey. *EPJ data science*, 10(1), 1-13.

Malhotra, N., & Krosnick, J. A. (2007). The effect of survey mode and sampling on inferences about political attitudes and behavior: Comparing the 2000 and 2004 ANES to Internet surveys with nonprobability samples. *Political Analysis*, 15(3), 286-323.

NOAA 2021. Storm Events Database. Available from <https://www.ncdc.gov/stormevents.html>

Nunley, C., & Sherman-Morris, K. (2020). What people know about the weather. *Bulletin of the American Meteorological Society*, 101(7), E1225-E1240.

Rivers, D. (2007, Aug U.S.t). Sampling for web surveys. In *Joint Statistical Meetings* (p. 4).

Vehovar, V., Motl, A., Mihelič, L., Berčič, B., & Petrovčič, A. (2012). Zaznava sovražnega govora na slovenskem spletu. *Teorija in praksa*, 49(1), 171-189.

Yeager, David S., et al. "Comparing the accuracy of RDD telephone surveys and internet surveys conducted with probability and non-probability samples." *Public opinion quarterly* 75.4 (2011): 709-747.

**4. Describe any tests of procedures or methods to be undertaken. Testing is encouraged as an effective means of refining collections of information to minimize burden and improve utility. Tests must be approved if they call for answers to identical questions from 10 or more respondents. A proposed test or set of tests may be submitted for approval separately or in combination with the main collection of information.**

The Weather and Society Dashboard Project will not collect test data because the research team has been collecting test data, assessing the validity and reliability of survey measures, and honing the small area estimation methodology for the past three years (see Ripberger et al. 2019 and Ripberger et al. 2020).

The Quick Response Surveys Project has piloted similar surveys through public response surveys conducted in the past by the Principal Investigator and Co-Investigator. Based on this experience, the surveys will not be tested on more than 9 people.

**5. Provide the name and telephone number of individuals consulted on statistical aspects of the design and the name of the agency unit, contractor(s), grantee(s), or other person(s) who will actually collect and/or analyze the information for the agency.**

Weather and Society Dashboard Project contacts who were consulted on statistical aspects of the design and will collect and analyze the survey data for the agency include:

Joseph Ripberger, Makenzie Krocak, Carol Silva, and Hank Jenkins-Smith from the University of Oklahoma's National Institute for Risk and Resilience - 405-325-1720; [risk-info@ou.edu](mailto:risk-info@ou.edu)

Quick Response Surveys contacts who were consulted on statistical aspects of the design and will collect and analyze the survey data for the agency include:

Brenda Philips, (413) 577-2213; [bphilips@engin.umass.edu](mailto:bphilips@engin.umass.edu);

Cedar League, (406)202-8167; [cedarleague@gmail.com](mailto:cedarleague@gmail.com);

Nathan Meyers, (734)771-7323; [npmeyers@soc.umass.edu](mailto:npmeyers@soc.umass.edu);

Quinnehtukqut McLamore, (325) 260-6760; [qmclamore@umass.edu](mailto:qmclamore@umass.edu);

David Westbrook, (413) 522-1409; [westy@cs.umass.edu](mailto:westy@cs.umass.edu).