Data Analysis

Michael Schramm

2023-05-17

# Survey Demographic Profile

Table : Demographic profile (unadjusted) of survey respondents.

| **Characteristic** | **N = 1,1001** |
| --- | --- |
| Sex/Gender |  |
| Male | 529 (48%) |
| Female | 565 (51%) |
| Other | 4 (0.4%) |
| No answer | 2 (0.2%) |
| Age |  |
| 18:24 | 125 (11%) |
| 25:34 | 192 (17%) |
| 35:44 | 204 (19%) |
| 45:54 | 198 (18%) |
| 55:64 | 171 (16%) |
| 65+ | 208 (19%) |
| No answer | 2 (0.2%) |
| Race/Ethnicity |  |
| American Indian/Native American or Alaska Native | 15 (1.4%) |
| Asian | 49 (4.5%) |
| Hispanic or Latino or Spanish Origin of any race | 109 (9.9%) |
| Black or African American | 119 (11%) |
| Native Hawaiian or Other Pacific Islander | 3 (0.3%) |
| White or Caucasian | 723 (66%) |
| Other | 21 (1.9%) |
| Two or More | 57 (5.2%) |
| No answer | 4 (0.4%) |
| Educational |  |
| Some high school | 47 (4.3%) |
| High school graduate or GED | 418 (38%) |
| Associate degree | 178 (16%) |
| Bachelor's degree | 246 (22%) |
| Master's degree | 132 (12%) |
| Doctorate or terminal degree | 28 (2.5%) |
| Other | 40 (3.6%) |
| No answer | 11 (1.0%) |
| 1n (%) | |

# Data Analysis

## Weighting

Prior to analysis, individual survey responses were weighted so that marginal proportions of the survey (Table ) matched national level benchmarks from the 5-year 2021 American Community Survey (ACS) on sex/gender, age group, race/ethnicity, and education level (Table ). Weights on gender were developed by re-coding “female” and “other” responses as “non-male” because the ACS only provides binary response options for sex. Using this approach, responses from both “female” and “other” respondents have the same marginal weight. Kennedy et al. (2022) provide substantial discussion on the treatment of sex and gender in survey adjustment. Due to small subpopulation sample sizes within the Race/Ethnicity variable, Race/Ethnicity were recoded as White or Caucasian and Non-white categories. Weights were developed by poststratification raking using the American National Election Study (ANES) weighting algorithm implemented in the *anesrake* R package (DeBell and Krosnick 2009; Pasek 2018).

Table : Marginal survey and target population proportions and marginal weighted values.

| **Variable** | **Value** | **Unweighted N** | **Unweighted %** | **Target %** | **Weighted N** | **Weighted %** |
| --- | --- | --- | --- | --- | --- | --- |
| Sex/Gender | Male | 529 | 48.1 | 49.0 | 539.1 | 49.0 |
| Not Male | 569 | 51.7 | 51.0 | 560.9 | 51.0 |
| No answer | 2 | 0.2 |  |  |  |
| Age | 18:24 | 125 | 11.4 | 11.9 | 130.6 | 11.9 |
| 25:34 | 192 | 17.5 | 17.7 | 195.1 | 17.7 |
| 35:44 | 204 | 18.5 | 16.6 | 183.1 | 16.6 |
| 45:54 | 198 | 18.0 | 16.3 | 179.2 | 16.3 |
| 55:64 | 171 | 15.5 | 16.8 | 184.4 | 16.8 |
| 65+ | 208 | 18.9 | 20.7 | 227.6 | 20.7 |
| No answer | 2 | 0.2 |  |  |  |
| Race/Ethnicity | White | 723 | 65.7 | 62.4 | 686.3 | 62.4 |
| Non-white | 373 | 33.9 | 37.6 | 413.7 | 37.6 |
| No answer | 4 | 0.4 |  |  |  |
| Education | Some high school | 47 | 4.3 | 7.8 | 85.8 | 7.8 |
| High school graduate or GED | 418 | 38.0 | 49.4 | 543.7 | 49.4 |
| Associate degree | 178 | 16.2 | 8.3 | 91.3 | 8.3 |
| Bachelor's degree | 246 | 22.4 | 19.4 | 213.7 | 19.4 |
| Master's degree | 132 | 12.0 | 8.3 | 91.3 | 8.3 |
| Doctorate or terminal degree | 28 | 2.5 | 1.3 | 14.7 | 1.3 |
| Other | 40 | 3.6 | 5.4 | 59.5 | 5.4 |
| No answer | 11 | 1.0 |  |  |  |

## Models

We developed three different proportional odds models (Agresti 2002) to assess factors associated with a respondent’s (1) self-described knowledge of PFAS (4 repsonses ranging from “I’ve neve heard of it, and don’t know what it is” to “I’m confident I know what it is”); (2) awareness of potential sources of PFAS (5 responses ranging from “Not at all familiar” to “Extremely Familiar”); and (3) intention to change use of items assosciated with PFAS (5 responses ranging from “will never change” to “have already changed”). Dependent variables included Sex/Gender, Age, Race/Ethnicity, Education, and awareness of community exposure to PFAS (Yes, No, Not Sure).

Model results are presented as odds-ratios (with approximate p-values calculated by comparing the t-value against the standard normal distribution). Marginal effects are also presented as population-level predicted probabilities for a given predictor estimated using observed values (Hanmer and Ozan Kalkan 2013). 95% confidence intervals were derived using a parametric bootstrap as implemented in the *svyEffects* R package (Santos 2023). All models were fit using the *survey* package in R version 4.2.1 (Lumley 2004; R Core Team 2022).

# Results

Table : Summary of weighted survey responses to questions about PFAS knowledge and awareness of community exposure.

| **Question** | **Response** | **Percent Response, SE** |
| --- | --- | --- |
| To your knowledge, has your community been exposed to PFAS? | Yes | 11.5, 1 |
| No | 41.1, 1.6 |
| Not sure | 47.4, 1.6 |
| How would you describe your knowledge about PFAS as an environmental contaminant? | I've never heard of it, and don't know what it is | 45.1, 1.6 |
| I've heard of it or seen it somewhere, but don't know what it is | 31.6, 1.5 |
| I think I know what it is | 17.2, 1.2 |
| I'm confident I know what it is | 6.2, 0.8 |

Most respondents had no knowledge (41.1%) or were unsure (47.4%) if their community has been exposed to PFAS (Table ). Only 11.5% responded that they knew their community has been exposed to PFAS. When asked to describe knowledge level about PFAS, 45.1% responded that they have never heard of it and don’t know what it is. An additional 31.6% responded they have heard of PFAS, but don’t know what PFAS are.

Table : Summary of weighted responses to questions about awareness of different potential sources of PFAS.

| **Question** | **Not at all familiar** | **Slightly familiar** | **Moderately familiar** | **Very familiar** | **Extremely familiar** |
| --- | --- | --- | --- | --- | --- |
| Drinking Water | 45.8% (1.6) | 19.7% (1.3) | 17.1% (1.2) | 9.2% (0.9) | 8.2% (0.9) |
| Public waterways near waste disposal sites | 45.2% (1.6) | 18.3% (1.3) | 20% (1.3) | 10.7% (1) | 5.9% (0.8) |
| Soils near waste disposal sites | 46.3% (1.6) | 20.2% (1.3) | 17.5% (1.2) | 10.1% (0.9) | 5.9% (0.8) |
| Dairy products | 51.1% (1.6) | 16.3% (1.2) | 15.3% (1.2) | 10% (0.9) | 7.3% (0.8) |
| Fresh produce | 50.3% (1.6) | 14.5% (1.1) | 16.1% (1.2) | 11.7% (1) | 7.4% (0.8) |
| Freshwater fish | 48.7% (1.6) | 16.4% (1.2) | 17.6% (1.2) | 11.4% (1) | 5.9% (0.7) |
| Seafood | 48.8% (1.6) | 15.3% (1.2) | 17.7% (1.2) | 9.9% (1) | 8.3% (0.9) |
| Food packaging | 48.1% (1.6) | 16.2% (1.2) | 16.9% (1.2) | 11.8% (1) | 7% (0.8) |
| Non-stick cookware | 47% (1.6) | 16.9% (1.2) | 16.8% (1.2) | 12.8% (1.1) | 6.5% (0.8) |
| Personal hygiene products | 46.6% (1.6) | 14.7% (1.1) | 18.2% (1.3) | 12.6% (1) | 8% (0.9) |
| Household products (fabrics, cleaning products, paints and sealants) | 45% (1.6) | 16% (1.2) | 18.7% (1.3) | 12.3% (1) | 7.9% (0.9) |
| Fire extinguising foam | 50.7% (1.6) | 14.7% (1.1) | 15.8% (1.2) | 11.7% (1) | 7.1% (0.8) |
| Fertilizers from wastewater plants | 45.9% (1.6) | 17.1% (1.2) | 17% (1.2) | 11.9% (1) | 8.1% (0.9) |

On average, 47.6% (minimum of 45% and maximum of 51.5%) of respondents were “not at all familiar” with potential sources of PFAS included in the survey (Table ). The probability of response decreased for increasing levels of familiarity across all potential PFAS sources. Only 7.2% (minimum of 5.9% and maximum of 8.3%) of respondents, on average across all sources, responded “extremely familiar.”

## Model Results

Table : Odds ratios and approximate p-values from proportional odds model explaining self-described knowledge of PFAS.

| **Characteristic** | **OR12** | **SE2** | **95% CI2** | **p-value** |
| --- | --- | --- | --- | --- |
| Sex |  |  |  |  |
| Male | — | — | — |  |
| Female | 0.99 | 0.133 | 0.76, 1.29 | >0.9 |
| Other | 0.80 | 0.735 | 0.19, 3.38 | 0.8 |
| Age |  |  |  |  |
| 18:24 | — | — | — |  |
| 25:34 | 1.38 | 0.235 | 0.87, 2.18 | 0.2 |
| 35:44 | 1.63\* | 0.237 | 1.03, 2.59 | 0.039 |
| 45:54 | 1.03 | 0.244 | 0.64, 1.67 | 0.9 |
| 55:64 | 1.09 | 0.255 | 0.66, 1.80 | 0.7 |
| 65+ | 0.94 | 0.259 | 0.57, 1.57 | 0.8 |
| Race/Ethnicity |  |  |  |  |
| White | — | — | — |  |
| Non-white | 0.95 | 0.149 | 0.71, 1.27 | 0.7 |
| Education |  |  |  |  |
| Some high school | — | — | — |  |
| High school graduate or GED | 0.65 | 0.327 | 0.34, 1.23 | 0.2 |
| Associate degree | 0.89 | 0.344 | 0.46, 1.76 | 0.7 |
| Bachelor's degree | 0.98 | 0.341 | 0.50, 1.92 | >0.9 |
| Master's degree | 1.04 | 0.363 | 0.51, 2.12 | >0.9 |
| Doctorate or terminal degree | 1.38 | 0.448 | 0.57, 3.32 | 0.5 |
| Other | 1.56 | 0.474 | 0.61, 3.94 | 0.4 |
| Community PFAS Exposure |  |  |  |  |
| Yes | — | — | — |  |
| No | 0.28\*\*\* | 0.189 | 0.19, 0.41 | <0.001 |
| Not sure | 0.23\*\*\* | 0.203 | 0.16, 0.35 | <0.001 |
| 1\*p<0.05; \*\*p<0.01; \*\*\*p<0.001 | | | | |
| 2OR = Odds Ratio, SE = Standard Error, CI = Confidence Interval | | | | |

We did not find evidence for any association between Sex, Race/Ethnicity, and Education with self-described knowledge about PFAS (Table ). We can’t exclude potential Race/Ethnicity sub-populations effects because Race/Ethnicity was collapsed into “White” and “Non-white” categories. For example, all the respondents identifying as “Native Hawaiian or Other Pacific Islander” (unweighted n = 3) responded that they had never heard of or knew what PFAS were. Some sub-populations might be more likely to answer that they have less knowledge about PFAS. With the sample size used in the current study we were not able to incorporate the sub-populations and develop a model that would converge.

There was not strong evidence for the influence of age on PFAS knowledge among most of the age brackets (Table ). However, there is evidence to support that someone in the 35:44 age bracket will respond with a higher self assessed knowledge level (OR = 1.63, p = 0.039) than someone in the reference bracket (18:24). There is also strong evidence that people aware of PFAS exposure in their communities self-report higher levels of knowledge about PFAS. People aware of PFAS exposure in their communities are 3.57 times (OR = 0.28, p < 0.001) and 4.35 times (OR = 0.23, p < 0.001 ) more likely to respond with a higher self-assessed knowledge level than those responding “No” or “Not sure” to awareness of PFAS contamination in their communities.

|  |
| --- |
| Figure 1: Average marginal predicted probabilities for self assessed knowledge of PFAS. Horizontal lines indicate the 95% confidence intervals of the marginal predicted probabilities. |

Marginal predicted probabilities show that someone that is aware of community PFAS exposure is is much more likely to respond that they confidently know what PFAS (16%) are compared to those that said their community hasn’t been exposed (5%) or don’t know (4%; Figure [Figure 1](#fig-m1margins)). Conversely, someone that is aware of community PFAS exposure was much less likely to respond that they had never heard of PFAS and didn’t know what is was (20%) compared to those that said their community hasn’t been exposed (46%) or don’t know (51%). We infer that respondents that are aware of community PFAS exposure are most likely to respond that they have at least heard of PFAS, but they may or may not be aware of what PFAS’s are. Respondents that are unaware or uncertain of community PFAS exposure are most likley to have never heard of it, and if they have they don’t know what it is.

|  |
| --- |
| Figure 2: Average marginal predicted probabilities to the question, “how familiar are you with the following items as potential sources of PFAS contamination?” Horizontal lines indicate the 95% confidence intervals of the marginal predicted probabilities. |

Awareness of community PFAS exposure also show strong associations with familiarity of potential PFAS sources and intentions to change use of items with potential for PFAS contamination ([Figure 2](#fig-m2margins), [Figure 3](#fig-m3margins)). On average, 46% and 46.5% of those that were unaware of or not sure if their communties were contaminated by PFAS responded they were “not familiar at all” with specific sources of PFAS contamination. This decreased to averages of 6% and 4% for the “extremely familiar” response. On average, people that stated their communities were contaminated by PFAS had a lower probability (18%) of responding that they were “not familiar at all” and higher probability (20%) of being “extremely familiar” with PFAS sources compared to the other two groups.

People indicating no or unsure of PFAS community contamination were on average more likely to say they will never change their use of items (17% and 17%) compared to those aware of PFAS contamination in their community (8%; [Figure 3](#fig-m3margins)). Those aware of community PFAS contamination were also more likely on average to have already changed use of items (22%) relative to the other two groups (11% for the “no” group and 10% for the “unsure” group).

|  |
| --- |
| Figure 3: Average marginal predicted probabilities to the question, “Please rate your intention to change your use of the following items because of their potential for PFAS contamination.” Horizontal lines indicate the 95% confidence intervals of the marginal predicted probabilities. |

# References

Agresti A. 2002. Categorical Data Analysis. 2nd ed. New York: Wiley-Interscience (Wiley series in probability and statistics).

DeBell M, Krosnick JA. 2009. Computing Weights for American National Election Study Survey Data. Ann Arbor, MI, and Palo Alto, CA: American National Election Studies ANES Technical Report series Report No.: nes012427. <http://www.electionstudies.org/resources/papers/nes012427.pdf>.

Hanmer MJ, Ozan Kalkan K. 2013. Behind the curve: clarifying the best approach to calculating predicted probabilities and marginal effects from limited dependent variable models. American Journal of Political Science. 57(1):263–277. doi:[10.1111/j.1540-5907.2012.00602.x](https://doi.org/10.1111/j.1540-5907.2012.00602.x).

Kennedy L, Khanna K, Simpson D, Gelman A, Jia Y, Teitler J. 2022. Using sex and gender in survey adjustment. <http://arxiv.org/abs/2009.14401>.

Lumley T. 2004. Analysis of complex survey samples. J Stat Soft. 9(8). doi:[10.18637/jss.v009.i08](https://doi.org/10.18637/jss.v009.i08).

Pasek J. 2018. anesrake: ANES Raking Implementation. <https://CRAN.R-project.org/package=anesrake>.

R Core Team. 2022. R: A Language and Environment for Statistical Computing. <https://www.R-project.org/>.

Santos J. 2023. svyEffects: Simulation-Based Marginal Effects for Survey-Weighted GLMs. <https://github.com/jb-santos/svyEffects>.