Winter 2019

Protection Through Participation: Crowdsourced Tap Water Quality Monitoring for Enhanced Public Health

Sarah A. Jakositz
University of New Hampshire, Durham

PROTECTION THROUGH PARTICIPATION: CROWDSOURCED TAP WATER QUALITY MONITORING FOR ENHANCED PUBLIC HEALTH

BY

SARAH JAKOSITZ

BS Environmental Engineering, University of New Hampshire, 2018

THESIS

Submitted to the University of New Hampshire

in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

in

Civil and Environmental Engineering

December, 2019

This thesis was examined and approved in partial fulfillment of the requirements for the degree of Master of Science in Civil and Environmental Engineering by:

Thesis Director, Dr. Weiwei Mo, Assistant Professor of Civil and Environmental Engineering

Dr. M. Robin Collins, Professor of Civil and Environmental Engineering

Dr. Bridie McGreavy, Assistant Professor of Environmental Communication, University of Maine

On November 18th, 2018

Approval signatures are on file with the University of New Hampshire Graduate School.

TABLE OF CONTENTS

APPENDICES

ACKNOWLEDGEMENTS

This research would not have been possible without the advisement and contributions of Dr. Weiwei Mo, Dr. Bridie McGreavy, and Scott Greenwood. Thank you also to Dr. Weiwei Mo, Dr. Bridie McGreavy, and Dr. M. Robin Collins for serving as my thesis committee and providing valuable support and feedback. A special thank you also to Lana Pillsbury, Julie Bryce, and Maria Fahnestock, who were significant contributors to this project and the success of its first paper. In addition, this project would not be possible without the contributions of our citizen participants, as well as the help of Dr. Jim Malley from the University of New Hampshire, Mr. Ian Rohrbacher, and Cindy Klevens and Amy Rousseau from the New Hampshire Department of Environmental Services. I would also like to acknowledge the support of the United States National Science Foundation under an EAGER Award (#1743997). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

LIST OF TABLES

LIST OF FIGURES

PART I

PART II

ABSTRACT

Lead contamination in municipal drinking water is a national public health issue and is generally the result of water contact with leaded distribution piping and on-premise plumbing. As a result, the US Environmental Protection Agency's Lead and Copper Rule requires point of use sampling methods at a small fraction of consumer taps on the public water distribution system. While this approach is practical, it leaves large gaps of consumers without direct monitoring and protection. In response, a novel contest-based crowdsourcing study was conducted to engage the public in monitoring their own water quality at their home taps and study factors that shaped participation in drinking water monitoring. Participants were asked to collect samples of their household drinking water through social media postings, kiosks, and community events with the chance to win a cash prize. The project distributed approximately 800 sampling packets and received 147 packets from participants of which 93% had at least partially completed surveys.

Part I of this thesis investigated lead levels, participant recruitment and demographic patterns, and motivations for participation. On average, private wells were found to have higher lead levels than the public water supply, and the higher lead levels were not attributed to older building age. There was also no statistical relevance between the participants' perceived and actual tap water quality. Survey responses indicated that citizens were motivated to participate in the project due to concerns about their own health and/or the health of their families. In contrast, participants reported that they were not motivated by the cash prize. Part II of this thesis investigated the influence of socioeconomic characteristics on participants' environmental literacy, behavior, and social networks. When looking at actions taken in response to water quality issues, income, age, and educational groups had some of the largest, significant, paired

differences. With regards to knowledge, this project showed success in potentially improving citizen's scientific literacy relating to key lead information, and overall provided self-assessed educational benefits to those who participated. This project helps inform future public engagement with water quality monitoring, create new knowledge about the influence of personal motivations for participation, and provide recommendations to help increase awareness of water quality issues. It also demonstrates that the crowdsourcing method could be used to actively engage and inform citizen participants in water quality monitoring efforts, creating a more scientifically literate and active public.

PART I: Protection Through Participation: Crowdsourced Tap Water Quality Monitoring for Enhanced Public Health

CHAPTER I: Introduction

Improvements in drinking water technology have contributed greatly to public health protection in the last century (Cutler and Miller 2005). However, lead is still a recurring and pervasive problem for residents across the United States (US). A survey conducted by the American Water Works Association in 2016 estimated that there are 6.1 million lead service lines (either full or partial) which serve 15 to 22 million people (Cornwell et al. 2016). Homeowners who do not have lead pipes but have lead solder and/or brass fittings in the premise plumbing are also susceptible to lead contamination by corrosion (NRC 2006). Lead consumption through drinking water is estimated to account for more than 20% of American's total lead exposure (USEPA 2018a). This percentage can increase to 85% or more of total exposure for infants who consume mostly formula made with tap water (Roy and Edwards 2018, USEPA 2018a). Lead can have damaging effects on the cardiovascular, nervous, and hematopoietic systems, especially the developing nerve systems of young children, infants, and fetuses (Eubig et al. 2010, Kim et al. 2015, WHO 2018). The amount and rate of lead release is affected by a variety of factors, including stagnation time, flow rate, scale composition, system configuration, and water quality (Doré et al. 2019, Roy and Edwards 2018, Schock 1990). For example, the water crisis in Flint, MI occurred after the city switched its drinking water source to the Flint River without implementing corrosion controls which resulted in accelerated leaching and dangerously high lead levels (Pieper et al. 2017). The crisis affected approximately 100,000 residents and its repair is estimated to cost US$1.5 billion (Craft-Blacksheare 2017, Gostin 2016, Ruble et al. 2018). Unfortunately, Flint, MI is not alone in its struggles, as pollution and aged infrastructure create disparate impacts for vulnerable populations across the US (Bullard 2008, Campbell et al. 2016).

Realization of the severe health implications of lead consumption pushed the US Congress to amend the Safe Drinking Water Act in 1986, which prohibited the use of leaded pipes, solder, and flux in public water systems (USEPA 1989). In 1991, the Lead and Copper Rule (LCR) was established to address lead issues in US public drinking water systems. The LCR controls lead and copper in drinking water by establishing treatment techniques and requiring systems to regularly monitor drinking water at consumer taps. The LCR set an Action Level for lead of 15 ppb, or 0.015 µg/L. If more than 10% of taps selected for monitoring (usually based on a tiered system prioritizing locations with the highest risk of lead contamination) exceed this Action Level, the municipality is required to take additional corrosion control measures as well as recommend precautionary steps to the public. While the Action Level serves as a practical guide for management actions, there is no known safe level of lead exposure (CFR 2018, USEPA 2018b, WHO 2018). Furthermore, more than 13 million US households relying on private wells are not subject to the LCR and hence have less access and protection from established services (Liu et al. 2005, NRC 2006, Pieper et al. 2015a, USEPA 2018c). The LCR is also limited in the high labor and time cost related to its execution and the difficulty in gaining access to private properties at desired times. As a result, sampling is generally conducted for less than 0.1% of the end users serviced and oftentimes, the same tap locations are monitored for each sampling period (AWWARF 2008, NRC 2006, Zhang et al. 2009). Because of these limitations, the Environmental Protection Agency (EPA) is looking to improve the LCR with the goal of minimizing lead exposure, designing clearer and more enforceable requirements, creating stronger consumer education programs, addressing environmental justice, and integrating drinking water with cross-media lead reduction efforts (USEPA 2016). Achieving this goal requires looking beyond the traditional monitoring and data collection strategies.

Over the last two decades, a rapidly growing body of research has demonstrated that crowdsourcing and engaging members of the public in environmental monitoring can increase capacities to address complex problems like the public health crisis from drinking water contamination (Bonney et al. 2014, Conrad and Hilchey 2011, Fox et al. 2016). Crowdsourcing has traditionally been used as a low-cost solution to large-scale tasks that can be addressed by widely distributed and independent citizens (Howe 2006, Jeppesen and Lakhani 2010, Malone et al. 2010). When applied to water quality monitoring, a crowdsourcing scheme is able to utilize the resources and knowledge of citizens to substantially reduce the cost, time, and professional labor needed for sample collection and/or analysis, increase the efficiency of individual monitoring activities, and help better allocate limited public resources (Bonney et al. 2009, Silvertown 2009). Previous projects that utilize citizen science in environmental monitoring are often focused on natural resources and ecosystem services such as wildlife, water resources, soil, and plants rather than engineered systems (Bonney et al. 2009, Bonney et al. 2014, Conrad and Hilchey 2011, Dickinson et al. 2012, Jollymore et al. 2017, Pandya 2012, Shirk et al. 2012, Silvertown 2009, Wiggins and Crowston 2011). Examples of the few efforts which have examined citizen science in the context of public drinking water monitoring include recent studies that involved collaborations between citizens and researchers to understand the severity of the Flint water crisis (Goovaerts 2019, Roy and Edwards 2019). A few studies have taken an empirical approach to studying social outcomes of citizen science, such as motivations, perceptions, and behaviors (Boakes et al. 2016, Raddick et al. 2009, Seymour and Haklay 2017). However, these studies have not focused on the area of public drinking water monitoring. Furthermore, understanding and strengthening/maintaining participation in crowdsourcing projects remains a key challenge. In the business realm, contest-based crowdsourcing is increasingly used to solicit innovative solutions related to computer programming, process/graphic design, pharmaceutical development, etc. (Boudreau and Lakhani 2013, Lakhani 2016, Riedl et al. 2016). Contests are often shown to be an effective way in attracting a broader audience and generating more desirable solutions (King and Lakhani 2013, Lakhani et al. 2013). Furthermore, latest findings from crowdsourcing researchers indicate that crowds, after being solicited through contests, are more likely to self-organize into a larger number of teams that could function more effectively than artificially matched teams (Blasco et al. 2013). The effect of contests in crowdsourced water quality monitoring, however, remain unknown.

In this study, we designed an innovative city-scale contest-based crowdsourced water quality monitoring scheme at the consumer tap to address some of the aforementioned limitations related to the current LCR and investigated its effectiveness. By applying the crowdsourcing scheme, we engaged citizens to collect their own tap water samples through contests and we then tested lead concentration in these samples and conducted analyses to better understand the intersections of participation, program design, and social-environmental outcomes.

CHAPTER II: Methods

2.1 Project Design

This study was carried out in a New Hampshire city which in 2017 had a population of around 30,797 people and 12,953 households (USCB 2018b). The city was selected based upon its diverse water source (65% public, 35% private) and the strong support of the local water utility on this project.

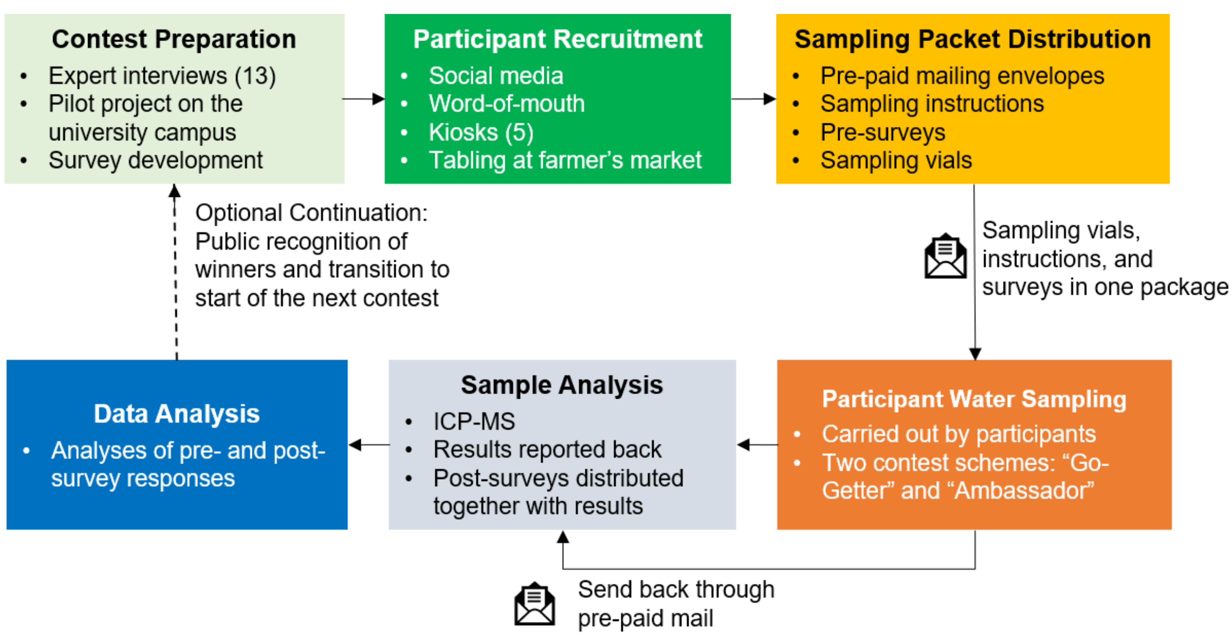


Figure 1. Process flow schematic of the tested sample collection scheme in the current project.

A schematic of the overall project design is provided in Figure 1. A total of 13 expert interviews (with regulators, utility operators, public health experts, etc.) and a pilot project on the University of New Hampshire (UNH) campus were conducted prior to the start of the project for preparation and testing purposes. Suggestions provided by the experts and lessons learned from the pilot project were used to inform the design of the current project. Participant recruitment initiated on August 5, 2018. We applied four different recruitment strategies, including: informational kiosks, social media, word of mouth, and direct interaction at a public event. Five information kiosks (Figure 2(a)), containing sampling packets, were distributed across the city at highly trafficked locations including two local grocery stores, a shopping mall, a community center, and a public library. The kiosks allowed people to take and return sampling packets directly. Facebook was selected as our social media platform based upon the presence of existing groups focused on local issues. We created a dedicated page about the project and launched a campaign specifically targeting residents of the area. During the participant recruitment stage, the page was updated weekly with the latest news about the project. A local Facebook group was also approached and utilized to promulgate the project to approximately 940 of their followers. Word of mouth recruitment approaches were conducted through developing relationships with a local church, with 700 parishioners. Church leadership informed members at weekly mass who distributed and collected samples directly. Lastly, the project team tabled at the local farmer's market on one occasion to evaluate direct interaction with the public. The official sample collection period ended on August 31, 2018. Late samples were accepted through mid-November.

Figure 2. (a) Two-level, cardboard informational kiosks were designed to attract the attention of passersby and provide key information to participate in the project. (b) Sampling packets were housed in pre-paid mailing envelopes to allow for easy return upon completion. (c) Returned samples were analyzed for lead concentration using an inductively coupled plasma mass spectrometer at a University of New Hampshire lab.

Each recruited participant received a sampling packet that contained a pre-survey that requested sampling locations, contact information and social data related to participant perceptions and participation in the project. The packets also included an empty 50-mL sampling vial and instructions on collecting first draw samples without contaminating the sample (Figure S1 of the supporting information). Sample vials were screw top, wide mouth, and fit easily under most kitchen faucets. Participants returned the water sample either via pre-paid mail service or returning it to the sampling/informational kiosks. Upon receipt, triple distilled 70% ultra-high purity nitric acid (400µL) was added directly to the collected samples, for preservation, and samples were stored at 4°C until the time of analysis. Analyses were conducted over two efforts, the first in September 2018 and the second in November 2018, as sample packet were received. Analytical analysis was conducted at the UNH Plasma Geochemistry Lab via inductively coupled plasma mass spectrometer (ICP-MS) in accordance with EPA 200.8 (Brockhoff et al. 1999). External calibration curves using a certified standard (SPEX CertiPrep, Metuchen, NJ) were conducted at the start and conclusion of each run, along with constant calibration verification standards and blanks every 10 samples.

All water quality results were reported to project participants by the end of November 2018 primarily through individual email communications or US mail (when no email was provided). Correspondence included the concentration of lead in their water sample, guidance on interpretation and potential protective action that should be taken base upon their results, and a post-survey. All lead levels less than 1 µg/L were reported as "< 1 µg/L". Protective guidance suggestions were broken into four lead concentration brackets, "<1 µg/L," "1-5 µg/L," "5-15 µg/L," and ">15 µg/L," following resources provided by the New Hampshire Department of Environmental Services (NHDES) resources (NHDES 2016). Retesting was offered for participants who had elevated lead levels (>15 µg/L). Participants were encouraged to contact research staff with questions or concerns regarding their results.

Our social science survey design was informed by the Tailored Design Method (TDM) (Dillman et al. 2014). We used a pre- and post-survey design to describe patterns and identify outcomes from participation in the project. Both the pre- and post-survey were also administered using Qualtrics®. The surveys asked questions about participants' socio-economic characteristics, water quality perception, how participants found out about the project, motivations to participate, intended actions related to water quality information, among others. We also built from the Developing, Validating, and Implementing Situated Evaluation Instruments (DEVISE) Framework using adapted scales for motivations to participate in crowdsourcing (Philips et al. 2017) and for environmental action (Porticella et al. 2017). To test the role of monetary incentives in motivating the public in tap water quality monitoring, two contest schemes were designed and implemented, namely “Go-Getter” and “Ambassador,” each with $200 cash rewards. The “Go-Getter” scheme rewards the participant who collected the most samples from different locations. The “Ambassador” scheme rewards the participant who introduced the highest number of new recruits to the program.

2.2 Data Treatment and Analysis

The project received a total of 149 returned packets either via mail or kiosk drop-off. Two of these packets did not contain a sample. Four additional packets did not include any form of contact information. Hence, a total of 142 packets were analyzed for lead concentrations. For participants who were offered retests, the average of their original and retested sample concentrations was used in the data analysis. A total of 136 pre-survey responses and 42 post-survey responses had more than 50% questions answered, and hence were included in the survey analyses. Out of these responses, 36 pre- and post-surveys were matched for comparison. Survey responses were analyzed using IBM SPSS Version 25. We conducted descriptive and bi-variate statistical tests (t-test and chi-square) to describe patterns of participation, perceptions about water quality, and associations between independent variables. We ran a Principal Components Analysis (PCA) with a Varimax rotation to assess the underlying factors that shaped participant motivations and tested the internal reliability of this scale using Cronbach’s alpha. PCA is a method used to transform a large number of related variables into a smaller number of uncorrelated variables using linear combinations, creating a simpler basis to describe the data (Everitt and Hothorn 2011, Jackson 2005).

We also used the georeferencing technique in ArcMap® to plot sampled locations in this project and compared against the current LCR sampling sites provided by the NHDES based upon their physical addresses (GRANIT; 2019, NHDES 2018). Standard deviation ellipses were then constructed, which contain two standard deviations of locations for each of the LCR and sampled datasets. The ellipses were then used to compare the area coverages of the two datasets.

CHAPTER III: Results and Discussions

3.1 Program Efficacy

The return rate of this project was around 18% (142 analyzed samples out of 800 distributed packets). The sample collection period lasted for a total of 26 days, and the entire project lasted for around 58 days. Figure 3 shows the location of samples collected through this project as well as the samples collected under the latest LCR sampling protocol. This study was able to almost double the number of households tested as compared to the current LCR protocol (77 households). The area covered through this project was around 140.4 km^2 based upon the standard deviation ellipse, which was around 2.3 times of the area covered by the LCR protocol (65.5 km^2). The samples collected through this project were well spread throughout the testbed city. This project has also effectively extended lead monitoring to households relying on private wells. Of the returned samples, around 67% were from families connected with the public drinking water supply and 33% were from households that rely on private wells. This was comparable to the city as a whole in which the municipality served 65% of residences (Interview 2018). When asked if they would be willing to participate in another project like this if given the opportunity, 77% of participants who responded to this question in the post-contest survey selected "Yes" and 19% selected "Maybe." This suggested the project's potential to retain volunteers that were relatively easily accessible for future monitoring/testing activities.

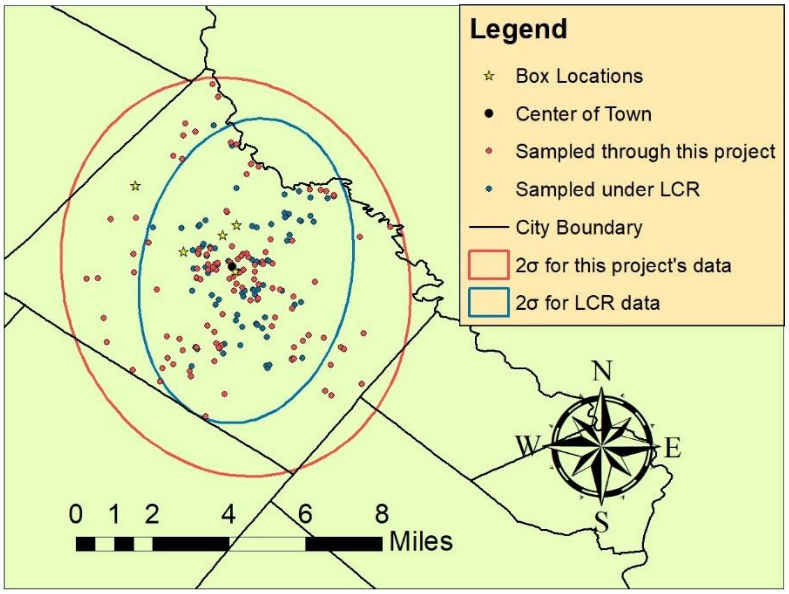


Figure 3. Map of sites in testbed city currently monitored under the LCR (blue) and sites tested during the pilot study (red). Ellipses depict statistical spread of data and contain two standard deviations (95%) of samples.

3.2 Water Sample Analysis

All samples were collected from kitchen faucets except for one sample which was taken from a garden hose and six which were taken from bathroom faucets. The majority of the samples tested (around 68%) had a lead concentration below 1 µg/L (Figure 4(a)). Around 3% of the samples tested had a higher lead concentration than the EPA action level, all of which were from private wells. An additional 7% had a lead concentration between 5 and 15 µg/L, all of which were also from private wells. About 57% of well samples had 1 µg/L or greater lead concentrations, whereas 20% of public supply samples had 1 ppb or greater. All public supply samples were below 5 µg/L, but 30% of private well samples were at or above 5 µg/L. The average lead concentration of private well samples was 7.22 µg/L with a standard deviation of 23.49 (without the retested sample, the average was 3.81 µg/L and the standard deviation was 5.52); while the average lead concentration of public supply samples was 0.613 µg/L with a standard deviation of 0.840. Lead in drinking water was often assumed to be only a problem of public supply, as attention is often centered on the failures and preventive management of large pipeline networks. Nevertheless, our study shows that households that rely on private wells are not necessarily free of lead contaminations. In fact, they could be even more the case. Potential sources of lead in well water include submersible pumps with leaded-brass components, plumping components imported from outside the US where lead is not as strictly regulated, and/or older well packer elements (CDC 2018). Another potential cause for high lead levels could be the use of ion exchange devices to reduce the hardness of sourced groundwater, thus making it more corrosive (NHDES 2009, USGS 2016). Some studies, however, indicate that ion exchange softening does not affect the corrosivity of water (Sorg et al. 1998). In general, low pH water, high dissolved oxygen, high temperature, and high levels of dissolved solids increase corrosion rates (Sadiq et al. 2007).

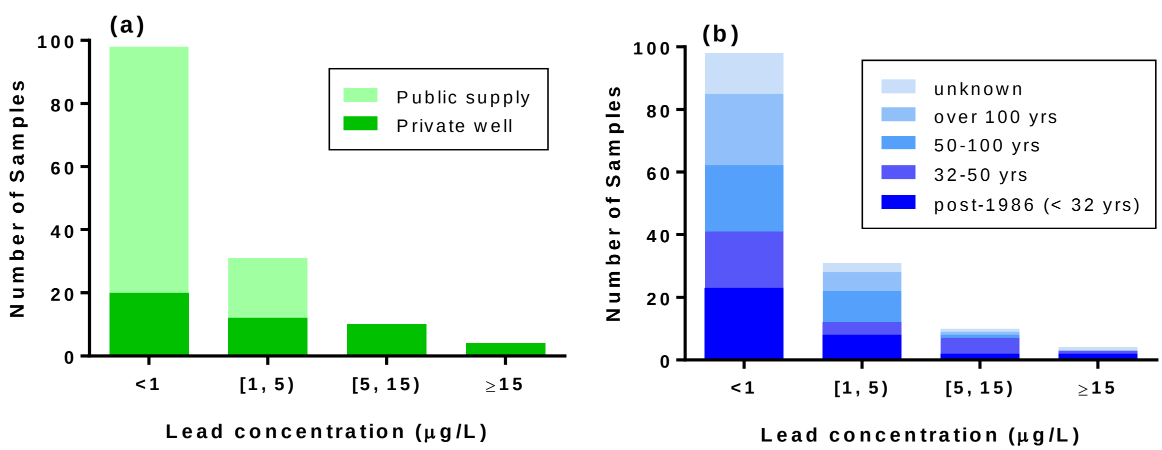


Figure 4. Distributions of (a) water sources (public vs. private) and (b) building age in relevance to their actual lead results. Remodeling was not considered in determining building age.

The higher lead concentration in private wells naturally invites the assumption that the buildings that rely on private wells might be older than the ones that rely on public supply. To obtain the age of the buildings being sampled, we searched the sampling site addresses within the city's public property assessment records database. Building ages that were not available on the database were obtained by running sampling site addresses through real estate search engines (CoStar Group 2018, NAR 2018, Zillow 2018). This analysis assumed that private wells were constructed the same year as their respective buildings. Building ages were eventually obtained for 124 samples and matched with sample analysis results obtained for the respective sample location. Figure 4(b) presents the building age in relation to their actual lead measurements. In fact, the highest lead concentrations were found in some of the newest homes in the region. Since the use of lead-containing solders in potable water systems was banned nationwide in 1986, we particularly investigated the measured lead concentrations for homes that were constructed before and after 1986. Half of the homes that were tested higher than the EPA action level were built after 1986, and one was less than 50 years old. Around 11% of the homes that were built after 1986 (less than 33 years old) have a lead concentration above 5 μg/L and around 21% of the homes that are between 32 and 50 years old have a lead concentration above 5 μg/L. However, this number is 3% for buildings between 50-100 years old and over 100 years old.

3.3 Demographics and Recruitment Patterns

Although this project aimed to encourage communication between participants, specifically with the advertisement of the "Go-Getter" and "Ambassador" contests, only 11% of participants who completed relevant survey questions indicated learning about the project via word-of-mouth and only 8% from social media. This shows the limitations of cash incentives and social media recruitment in a drinking water monitoring/testing project targeting a small, spatially constrained population. Tabling at the farmer's market also saw little success, with many citizens stating "I don't want to know" when offered free water testing. This introduces a challenging barrier of access to those who do not wish to know what is in their water. As this study did not survey those who chose not to participate, it is difficult to draw conclusions to exactly why they were deterred. These barriers would be worth investigating in future studies. Meanwhile, 53% of surveyed participants indicated they learned about the project when they saw a kiosk in person. This implies that persons of interest are more likely to participate if the project materials are directly accessible to them.

Survey results indicate that about half of the participants had an annual household income below the city's median household income (Table 1). This supports this method's ability to reach lower-income consumers, who have often been reported as socially disadvantaged groups facing inequities in water quality and are typically harder for researchers to reach due to project cost and time constraints (Bonevski et al. 2014, VanDerslice 2011). Around 26% of the participants rented their homes while 70% owned them. Persons occupying home rentals often face limited capacity in accessing resources to make changes to their homes, especially in urban areas (Mee et al. 2014). The results also show that our methods were more effective in recruiting more educated population, as the percentage of participants with a Bachelor's degree or higher was 15.5% higher than the city mean and the percentage of participants with less than high school education was 7.8% lower. This indicates education might have a positive effect on people's willingness in participating a program like this. Furthermore, we found that our project had a higher success in recruiting older citizens, with a median age of 55, which is 14 years older than the city-wide median age of 41 (USCB 2018b). This is also a trend seen in other environmentally focused citizen science projects (Merenlender et al. 2016, Trumbull et al. 2000). Around 22% of the respondents had children under the age of six in their home, which was significantly higher than the national average of 11.6% households (family and non-family) with children under six in their home (USCB 2016, 2018a). This indicates that having a young child in the home could be a potential motivator for participating in this project.

Table 1. Demographic data showing gender, education levels, whether the family has children under 6 years old in the house, rent or own, and income

Characteristics	Frequency	%	City Mean %
Gender			
Unknown or other	7	5.2	
Male	39	28.7	48.7 (USCB 2018b)
Female	90	66.2	51.3 (USCB 2018b)
Education			
Missing	3	2.2	
Less than high school	4	2.9	10.7 (USCB 2018b)
High school graduate and/or postsecondary education	78	57.3	67.3
Bachelor's degree or higher	51	37.5	22.0 (USCB 2018b)
Age			
Unknown	12	8.8	
18-24	1	0.7	21.0 (USCB 2018b)
25-34	21	15.4	
35-44	18	13.2	27.7 (USCB 2018b)
45-54	21	15.4	
55-64	22	16.2	12.9 (USCB 2018b)
65+	41	30.1	17.5 (USCB 2018b)

Children under 6 in the House			
Missing	6	4.4	
Yes	30	22.1	
No	100	73.5	
Rent or Own			
Unknown	5	3.7	
Rent	35	25.7	
Both	1	0.7	
Own	95	69.9	
Income			
Unknown	20	14.7	$50,759 in 2016 USD (median household income) (USCB 2018b)
$0 to $24,999	18	13.2	
$25,000 to $49,999	41	30.1	
$50,000 to $74,999	23	16.9	
$75,000 to $99,999	17	12.5	
$100,000 to $124,999	10	7.4	
$125,000 and up	7	5.1	

3.4 Lead Concentration Versus Water Quality Perception and Preventive Actions

Another question we investigated through the study was how people's perception was correlated to their actual water quality (Figure 5). Of the four participants whose lead concentrations were above the EPA action level, three of them considered their water quality to be very good or good. Around 42% of the participants with a lead concentration above 5 µg/L considered their water quality to be excellent or very good. We then recoded lead levels as a continuous data type into a binary category, where scores of ≥1ug/L were recoded as 1 ("Lead") and scores <1 were recoded as 0 ("No Lead"). Similarly, we recoded water quality perceptions into binary categories, where original rankings of "Excellent", "Very good" and "Good" were recoded as 1 ("Good") and "Fair" and "Poor" as 0 ("Bad"). There was no statistically significant relationship between evidence of lead in water and perceptions about water quality (x^2 =.150, df=1, p=0.699) (Table S1 in SI). This is may, in part, be due to lead in drinking water being tasteless, odorless, and colorless, and therefore undetectable to the consumer (CDC 2016). Those who ranked their drinking water as of poor quality likely did so by noticing detectable water quality issues. Lead along with some other drinking water contaminants are often not directly observable by end users, this could potentially hinder people's timely response to prevent potential harms.

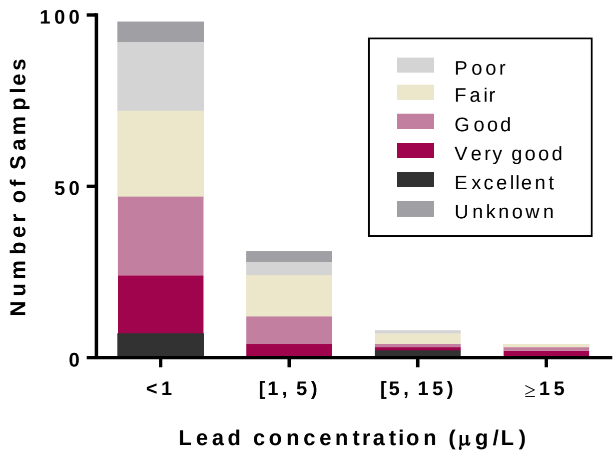


Figure 5. Correlations between lead concentration and water quality perception

We also asked participants whether or not they run their tap to flush their water before using it each day. Another chi-squared test was conducted to test the association between the respondents' water quality perceptions and actions to reduce the potential harm. Results of a Pearson Chi-Square analysis showed no statistically significant relationship between perceptions about drinking water quality (good or bad) and choice to run the tap (yes or no) in the pre-survey responses (x^2=0.108, df=1, p=0.743) (Table S2 in SI). This may be due to a lack of awareness of the benefits of flushing one's tap, which shows the importance of raising the public awareness of the potential drinking water issues and their preventive measures. While flushing has been widely recognized as an effective short-term method for tap water containing high lead levels, insufficient flushing time might result in increased rather than reduced lead exposure (Katner et al. 2018). Hence, it is important to provide the public with clear flushing guidelines, while acknowledging its practical limitations.

3.5 Motivation for Participation

Participation in this project was mostly motivated by wanting to learn about drinking water quality (around 94% of the participants either agreed or strongly agreed in the pre-survey) and concerns about personal health and/or family health (around 95% either agreed or strongly agreed in the pre-survey). A PCA on the pre-survey responses further identified three main motivation factors which we labelled health and identity, extrinsic incentives, and personal satisfaction (Table 2). The health and identity factor corresponded with the highest rated mean motivations, providing further evidence that health and identity related factors were key factors that motivated participation in the contest. This finding has implications for recommendations of how to design communication plans to help encourage people to participate in household drinking water monitoring activities. On the other hand, the contest and the cash prize were found to be an ineffective motivator for people to participate this project. Around 67% of participants either disagreed or strongly disagreed that they were motivated to participate by the cash prize, and 21% were neutral. Additionally, when asked which contest the participant hoped to win, 82% of participants responded "Neither" in the pre-survey (Table S3 in the SI). For those who participated and also responded to the post-survey, 92% (n=39) indicated they were not participating in either of the contest options. A paired t-test comparing responses to pre- and post-survey questions on motivation revealed no significant differences in overall motivations before and after the project (Table S4 in the SI). However, participants do show a decrease in motivation to learn about water quality. We interpreted this result to indicate that the project may have helped satisfy their motivation for learning about water quality.

Table 2. Results from Principal Components Analysis (PCA) of pre-survey data showing rotated factor loadings (Varimax rotation) on two primary motivation factors which accounted for 59% of the variance. Motivation factors included health and identity, extrinsic incentives, and personal satisfaction. The scale demonstrated internal consistency (Cronbach's Alpha = 0.747)

"I'm interested in participating in this drinking water quality contest because…"

N=136	Extrinsic	Intrinsic
Of the recognition or respect I'll get from others.	**0.836**	-0.076
People I look up to think it's good to participate in this contest.	**0.814**	0.015
I want others to think I'm good at doing activities related to environmental health.	**0.727**	0.046
I am required to participate in this contest.	**0.629**	-0.374
Participating in this contest will help me achieve things that are important to me.	**0.621**	0.257
I want to win the cash prize.	**0.621**	-0.171
I care about my family's health.	-0.154	**0.880**
I care about my personal health.	-0.190	**0.873**
I am an environmentally conscious person.	0.189	**0.747**
I want to learn about my drinking water quality.	-0.064	**0.703**
I enjoy doing activities related to environmental health.	0.491	**0.544**

CHAPTER IV: Conclusions

This project has demonstrated that well designed crowdsourcing approaches can identify lead concentrations at the consumer tap while actively engaging and informing the public, which directly addresses a large portion of the EPA's LCR revision goals. An important finding in this study is that lead concentrations were statistically higher at locations served by private wells than the public system, and the lead concentrations are not corresponding to the age of the households. This is consistent with several studies conducted over the past 5 decades in North Carolina, Pennsylvania, and Virginia (Francis et al. 1984, Maas and Patch 1990, Pieper et al. 2015b, Swistock et al. 1993). While lead exposure has decreased in public systems, exposure from private systems remains a large data gap in lead exposure which may pose challenges in achieving the federal goal to eliminate elevated blood lead levels in children by 2020 (DHHS 2012). Perceptions of water quality was found to be neither linked with the actual lead concentration, nor the preventive actions people take to minimize harm (i.e., flush tap), indicating a potential barrier protection of public health. Our participant recruitment has a higher success with female, more educated, and older populations. Though our hands-off recruitment approach was useful in reducing time and resource requirements for researchers, it should be noted that methods such as door-to-door may be more effective in achieving higher return rates (Pieper et al. 2018). Furthermore, personal contact with the kiosks was found to be the most effective approach to recruitment. Participants are mostly motivated by health and identity factors. These findings have implications on the future design of communication strategies to improve communication efficacy and engage the under-represented groups.

We envision the design of the crowdsourced scheme developed in this study could be expanded to other contaminants of concern. Outcomes from this project have demonstrated the effectiveness of such a scheme in terms of the amount of households sampled and the area covered. However, the usefulness and broader adoption of such a scheme in monitoring water quality also depends on the availability of easy-to-understand/use sampling and analysis methods as well as low cost contaminant analysis techniques. In our project, we utilized a different method than the standard lead testing method, because the standard method requires collection of one liter of water sample, which would have been a barrier for transporting and testing those samples in one centralized location. In our design, we were not able to allow participants to directly analyze their samples, due to the lack of low-cost but accurate lead testing techniques that can be freely distributed to the participants. A direction of potential future research is to develop such low-cost measuring techniques, sensors, or surrogate indicators for household water quality monitoring. We expect the availability of such techniques would greatly enhance our capability in providing continuous or random water quality checks for public health protection.

LIST OF REFERENCES

AWWARF (2008) American Water Works Association Research Foundation; Contribution of service line and plumbing fixtures to lead and copper rule compliance issues, Denver, CO.

Blasco, A., Boudreau, K.J., Lakhani, K.R., Menietti, M. and Riedl, C. (2013) Do Crowds have the Wisdom to Self-Organize?

Boakes, E.H., Gliozzo, G., Seymour, V., Harvey, M., Smith, C., Roy, D.B. and Haklay, M. (2016) Patterns of contribution to citizen science biodiversity projects increase understanding of volunteers' recording behaviour. Scientific reports 6, 33051.

Bonevski, B., Randell, M., Paul, C., Chapman, K., Twyman, L., Bryant, J., Brozek, I. and Hughes, C. (2014) Reaching the hard-to-reach: a systematic review of strategies for improving health and medical research with socially disadvantaged groups. BMC Medical Research Methodology 14(1), 42.

Bonney, R., Cooper, C.B., Dickinson, J., Kelling, S., Phillips, T., Rosenberg, K.V. and Shirk, J. (2009) Citizen science: a developing tool for expanding science knowledge and scientific literacy. BioScience 59(11), 977-984.

Bonney, R., Shirk, J.L., Phillips, T.B., Wiggins, A., Ballard, H.L., Miller-Rushing, A.J. and Parrish, J.K. (2014) Next steps for citizen science. Science 343(6178), 1436-1437.

Boudreau, K.J. and Lakhani, K.R. (2013) Using the crowd as an innovation partner. Harvard business review 91(4), 60-69.

Brockhoff, C., Creed, J., Martin, T., Martin, E. and Long, S. (1999) EPA Method 200.8, Revision 5.5: Determination of trace metals in waters and wastes by inductively coupled plasma-mass spectrometry, EPA-821R-99-017.

Bullard, R.D. (2008) Dumping in Dixie: Race, class, and environmental quality, Westview Press.

Campbell, C., Greenberg, R., Mankikar, D. and Ross, R. (2016) A case study of environmental injustice: The failure in Flint. International journal of environmental research and public health 13(10), 951.

CDC (2016) Centers for Disease Control and Prevention; Sources of Lead, https://www.cdc.gov/nceh/lead/tips/water.htm, U.S. Department of Health & Human Services

CDC (2018) Centers for Disease Control and Prevention; Lead and Drinking Water from Private Wells, https://www.cdc.gov/healthywater/drinking/private/wells/disease/lead.html, U.S. Department of Health & Human Services, Centers for Disease Control and Prevention.

CFR (2018) Code of Federal Regulations; Protection of the Enviornment. Government, U.S.F. (ed).

Conrad, C.C. and Hilchey, K.G. (2011) A review of citizen science and community-based environmental monitoring: issues and opportunities. Environmental monitoring and assessment 176(1-4), 273-291.

Cornwell, D.A., Brown, R.A. and Via, S.H. (2016) National survey of lead service line occurrence. Journal-American Water Works Association 108(4), E182-E191.

CoStar Group (2018) Apartments.com, https://www.apartments.com/, CoStar Group Inc.

Craft-Blacksheare, M.G. (2017) Lessons learned from the crisis in flint, Michigan regarding the effects of contaminated water on maternal and child health. Journal of Obstetric, Gynecologic & Neonatal Nursing 46(2), 258-266.

Cutler, D. and Miller, G. (2005) The role of public health improvements in health advances: the twentieth-century United States. Demography 42(1), 1-22.

DHHS (2012) US Department of Health and Human Services; Office of Disease Prevention and Health Promotion, 2020: Improving the health of Americans, https://www.healthypeople.gov/.

Dickinson, J.L., Shirk, J., Bonter, D., Bonney, R., Crain, R.L., Martin, J., Phillips, T. and Purcell, K. (2012) The current state of citizen science as a tool for ecological research and public engagement. Frontiers in Ecology and the Environment 10(6), 291-297.

Dillman, D.A., Smyth, J.D. and Christian, L.M. (2014) Internet, phone, mail, and mixed-mode surveys: the tailored design method, John Wiley & Sons.

Doré, E., Deshommes, E., Laroche, L., Nour, S. and Prévost, M. (2019) Lead and copper release from full and partially replaced harvested lead service lines: Impact of stagnation time prior to sampling and water quality. Water Research 150, 380-391.

Eubig, P.A., Aguiar, A. and Schantz, S.L. (2010) Lead and PCBs as risk factors for attention deficit/hyperactivity disorder. Environmental health perspectives 118(12), 1654-1667.

Everitt, B. and Hothorn, T. (2011) An introduction to applied multivariate analysis with R, Springer Science & Business Media.

Fox, M.A., Nachman, K.E., Anderson, B., Lam, J. and Resnick, B. (2016) Meeting the public health challenge of protecting private wells: Proceedings and recommendations from an expert panel workshop. Science of the total environment 554, 113-118.

Francis, J.D., Brower, B., Graham, W.F., Larson, O. and McCaull, J.L. (1984) National statistical assessment of rural water conditions. Executive summary, Cornell Univ., Ithaca, NY (United States). Agricultural Experiment Station.

Goovaerts, P. (2019) Geostatistical prediction of water lead levels in Flint, Michigan: A multivariate approach. Science of the total environment 647, 1294-1304.

Gostin, L.O. (2016) Politics and public health: the Flint drinking water crisis. Hastings Center Report 46(4), 5-6.

GRANIT; (2019) New Hampshire Geographically Referenced Analysis and Information Transfer System (NH GRANIT), http://www.granit.unh.edu/, Earth Systems Research Center, Institute for the Study of Earth, Oceans and Space, University of New Hampshire.

Howe, J. (2006) The rise of crowdsourcing. Wired magazine 14(6), 1-4.

Interview (2018) personal communication with local drinking water utility operator. Mo, W. (ed).

Jackson, J.E. (2005) A user's guide to principal components, John Wiley & Sons.

Jeppesen, L.B. and Lakhani, K.R. (2010) Marginality and problem-solving effectiveness in broadcast search. Organization science 21(5), 1016-1033.

Jollymore, A., Haines, M.J., Satterfield, T. and Johnson, M.S. (2017) Citizen science for water quality monitoring: Data implications of citizen perspectives. Journal of environmental management 200, 456-467.

Katner, A., Pieper, K., Brown, K., Lin, H.-Y., Parks, J., Wang, X., Hu, C.-Y., Masters, S., Mielke, H. and Edwards, M. (2018) Effectiveness of Prevailing Flush Guidelines to Prevent Exposure to Lead in Tap Water. International journal of environmental research and public health 15(7), 1537.

Kim, H.-C., Jang, T.-W., Chae, H.-J., Choi, W.-J., Ha, M.-N., Ye, B.-J., Kim, B.-G., Jeon, M.-J., Kim, S.-Y. and Hong, Y.-S. (2015) Evaluation and management of lead exposure. Annals of occupational and environmental medicine 27(1), 30.

King, A. and Lakhani, K.R. (2013) Using open innovation to identify the best ideas. MIT Sloan Management Review 55(1), 41.

Lakhani, K.R. (2016) Managing Communities and Contests to Innovate with Crowds. Revolutionizing Innovation: Users, Communities, and Open Innovation, 109.

Lakhani, K.R., Boudreau, K.J., Loh, P.-R., Backstrom, L., Baldwin, C., Lonstein, E., Lydon, M., MacCormack, A., Arnaout, R.A. and Guinan, E.C. (2013) Prize-based contests can provide solutions to computational biology problems. Nature biotechnology 31(2), 108-111.

Liu, A., Ming, J. and Ankumah, R.O. (2005) Nitrate contamination in private wells in rural Alabama, United States. Science of the total environment 346(1), 112-120.

Maas, R.P. and Patch, S.C. (1990) Lead Contamination of North Carolina Domestic Tapwater: Prevalence, Risk Factors and Control Measures, UNC-Asheville Environmental Quality Institute.

Malone, T.W., Laubacher, R. and Dellarocas, C. (2010) The collective intelligence genome. MIT Sloan Management Review 51(3), 21.

Mee, K.J., Instone, L., Williams, M., Palmer, J. and Vaughan, N. (2014) Renting over troubled waters: An urban political ecology of rental housing. Geographical Research 52(4), 365-376.

Merenlender, A.M., Crall, A.W., Drill, S., Prysby, M. and Ballard, H. (2016) Evaluating environmental education, citizen science, and stewardship through naturalist programs. Conservation Biology 30(6), 1255-1265.

NAR (2018) National Association of Realtors; Realtor.com, https://www.realtor.com/, National Association of Realtors.

NHDES (2009) New Hampshire Department of Environmental Services; Ion Exchange Treatment of Drinking Water, https://www.des.nh.gov/organization/commissioner/pip/factsheets/dwgb/documents/dwgb-2-12.pdf.

NHDES (2016) New Hampshire Department of Health and Human Services; Lead and Drinking Water, https://www.dhhs.nh.gov/dphs/bchs/clpp/parents.htm, New Hampshire Department of Health and Human Services.

NHDES (2018) New Hampshire Department of Environmental Services; LCR DBP Sampling Sites, Cynthia Klevens.

NRC (2006) National Research Council; Drinking water distribution systems: assessing and reducing risks, National Academies Press.

Pandya, R.E. (2012) A framework for engaging diverse communities in citizen science in the US. Frontiers in Ecology and the Environment 10(6), 314-317.

Philips, T., Porticella, N. and Bonney, R. (2017) Motivation to Participate in Citizen Science Scale. Technical Brief Series. Cornell Lab of Ornithology, Ithica NY.

Pieper, K.J., Krometis, L.-A., Gallagher, D., Benham, B. and Edwards, M. (2015a) Profiling private water systems to identify patterns of waterborne lead exposure. Environmental science & technology 49(21), 12697-12704.

Pieper, K.J., Krometis, L.-A.H., Gallagher, D.L., Benham, B.L. and Edwards, M. (2015b) Incidence of waterborne lead in private drinking water systems in Virginia. Journal of water and health 13(3), 897-908.

Pieper, K.J., Martin, R., Tang, M., Walters, L., Parks, J., Roy, S., Devine, C. and Edwards, M.A. (2018) Evaluating water lead levels during the Flint water crisis. Environmental science & technology 52(15), 8124-8132.

Pieper, K.J., Tang, M. and Edwards, M.A. (2017) Flint water crisis caused by interrupted corrosion control: Investigating “ground zero” home. Environmental science & technology 51(4), 2007-2014.

Porticella, N., Philips, T. and Bonney, R. (2017) Motivation for Environmental Action Scale (Custom). Technical Brief Series. Cornell Lab of Ornithology, Ithaca NY.

Raddick, M.J., Bracey, G., Gay, P.L., Lintott, C.J., Murray, P., Schawinski, K., Szalay, A.S. and Vandenberg, J. (2009) Galaxy zoo: Exploring the motivations of citizen science volunteers. arXiv preprint arXiv:0909.2925.

Riedl, C., Zanibbi, R., Hearst, M.A., Zhu, S., Menietti, M., Crusan, J., Metelsky, I. and Lakhani, K.R. (2016) Detecting figures and part labels in patents: competition-based development of graphics recognition algorithms. International Journal on Document Analysis and Recognition (IJDAR) 19(2), 155-172.

Roy, S. and Edwards, M. (2019) Citizen Science During the Flint, Michigan Federal Water Emergency: Ethical Dilemmas and Lessons Learned.

Roy, S. and Edwards, M.A. (2018) Preventing another Lead (Pb) in Drinking Water Crisis: Lessons from the Washington DC and Flint MI contamination events. Current Opinion in Environmental Science & Health.

Ruble, K., Carah, J., Ellis, A. and Childress, S. (2018) Flint Water Crisis Deaths Likely Surpass Official Toll, PBS, PBS.org.

Sadiq, R., Kleiner, Y. and Rajani, B. (2007) Water quality failures in distribution networks—risk analysis using fuzzy logic and evidential reasoning. Risk Analysis: An International Journal 27(5), 1381-1394.

Schock, M.R. (1990) Causes of temporal variability of lead in domestic plumbing systems. Environmental monitoring and assessment 15(1), 59-82.

Seymour, V. and Haklay, M. (2017) Exploring engagement characteristics and behaviours of environmental volunteers. Citizen Science: Theory and Practice 2(1), 5.

Shirk, J., Ballard, H., Wilderman, C., Phillips, T., Wiggins, A., Jordan, R., McCallie, E., Minarchek, M., Lewenstein, B. and Krasny, M. (2012) Public participation in scientific research: a framework for deliberate design. Ecology and society 17(2).

Silvertown, J. (2009) A new dawn for citizen science. Trends in ecology & evolution 24(9), 467-471.

Sorg, T.J., Schock, M.R. and Lytle, D. (1998) Leaching of Metals From Household Plumbing Materials: Impact of Home Water Softeners.

Swistock, B.R., Sharpe, W.E. and Robillard, P.D. (1993) A survey of lead, nitrate and radon contamination of private individual water systems in Pennsylvania. Journal of Environmental Health, 6-12.

Trumbull, D.J., Bonney, R., Bascom, D. and Cabral, A. (2000) Thinking scientifically during participation in a citizen-science project. Science education 84(2), 265-275.

USCB (2016) United States Census Bureau; America’s Families and Living Arrangements: 2016, https://www.census.gov/data/tables/2016/demo/families/cps-2016.html, United States Census Bureau.

USCB (2018a) United States Census Bureau; Historical Households Tables, https://www.census.gov/data/tables/time-series/demo/families/households.html, United States Census Bureau.

USCB (2018b) United States Census Bureau; State & County QuickFacts, http://quickfacts.census.gov/qfd/states/12000.html

USEPA (1989) U.S. Environmental Protection Agency; Lead Ban: Preventing the Use of Lead in Public Water Systems and Plumbing Used for Drinking Water. Water, O.o. (ed).

USEPA (2016) U.S. Environmental Protection Agency; Lead and Copper Rule Revisions, https://www.epa.gov/sites/production/files/2016-10/documents/508_lcr_revisions_white_paper_final_10.26.16.pdf. Water, O.o. (ed), EPA.gov.

USEPA (2018a) U.S. Environmental Protection Agency; Basic Information about Lead in Drinking Water, https://www.epa.gov/ground-water-and-drinking-water/basic-information-about-lead-drinking-water#health.

USEPA (2018b) U.S. Environmental Protection Agency; Lead and Copper Rule, https://www.epa.gov/dwreginfo/lead-and-copper-rule.

USEPA (2018c) U.S. Environmental Protection Agency; Private Drinking Water Wells, https://www.epa.gov/privatewells. Agency, U.S.E.P. (ed).

USGS (2016) U.S. Geological Survey; Ground Water and the Rural Homeowner, https://pubs.usgs.gov/gip/gw_ruralhomeowner/.

VanDerslice, J. (2011) Drinking water infrastructure and environmental disparities: evidence and methodological considerations. American journal of public health 101(S1), S109-S114.

WHO (2018) World Health Organization; Lead poisoning and health, www.WHO.int.

Wiggins, A. and Crowston, K. (2011) From conservation to crowdsourcing: A typology of citizen science, pp. 1-10, IEEE.

Zhang, Y., Griffin, A., Rahman, M., Camper, A., Baribeau, H. and Edwards, M. (2009) Lead contamination of potable water due to nitrification. Environmental science & technology 43(6), 1890-1895.

Zillow (2018) Real Estate, https://www.zillow.com/.

PART II: Crowdsourced Tap Water Quality Monitoring: Understanding the Influence of Socioeconomics on Participants' Environmental Literacy, Behavior, and Social Networks

CHAPTER I: Introduction

Over the last decade, crowdsourcing and citizen science have been increasingly applied in environmental monitoring. Crowdsourcing is a method to combat large-scale problems at a low cost, utilizing the collective efforts of independent, widely distributed citizens (Howe 2006, Jeppesen and Lakhani 2010, Malone et al. 2010). By utilizing the resources and knowledge of citizens, the cost, time, and labor needs associated with sample collection and analysis are reduced. This allows for increased efficiency in monitoring activities which allows for improved allocation of public resources (Bonney et al. 2009, Silvertown 2009). Researchers have also found that crowdsourcing can serve as a venue to improve environmental literacy and engage citizens in environmental decision-making (Bonney et al. 2009, Den Broeder et al. 2016). For example, Brossard et al. (2005) evaluated the Birdhouse Network, an informal science education project of the Cornell Laboratory of Ornithology, and found that the project had a positive impact on participants' knowledge of bird biology (Brossard et al. 2005). Similarly, Cronje et al. (2011) saw significant science literacy gains in citizens after participating in invasive species monitoring training (Cronje et al. 2011). Nerbonne and Nelson (2004) saw greater knowledge and involvement in civic processes for citizens involved in volunteer macroinvertebrate monitoring groups (Nerbonne and Nelson 2004). However, our knowledge about the extent of such changes in environmental literacy and behaviors as well as how these changes are influenced by participants' socioeconomic characteristics remains limited.

Of the previous crowdsourced environmental monitoring studies, many have focused on natural resources and ecosystems (Bonney et al. 2009, Bonney et al. 2014, Conrad and Hilchey 2011, Dickinson et al. 2012, Jollymore et al. 2017, Pandya 2012, Shirk et al. 2012, Silvertown 2009, Snik et al. 2014, Trumbull et al. 2000, Wiggins and Crowston 2011), while the monitoring of drinking water quality is a relatively untapped field. In the US, issues such as aging infrastructure and the increased detection of legacy and emerging contaminants require the continued development of drinking water treatment and monitoring solutions to better protect public health (US Environmental Protection Agency Office of Water 2015). Lead, for example, remains a recurring problem for communities. Around 20% of American's lead exposure comes from contaminated drinking water (U.S. Environmental Protection Agency 2018). This is partly because of the limited public resources available for continuous and widely-spread water quality monitoring at the consumer taps. A few recent crowdsourced drinking water monitoring studies occurred in response to the Flint water crisis (Goovaerts 2019, Jakositz et al. 2019, Roy and Edwards 2019). These studies investigated the effectiveness of the crowdsourced approach in monitoring water quality at the consumer taps. However, they did not study the influence of participation on the behavior and knowledge changes of participants.

On the other hand, there have been multiple studies looking at participants' change in attitudes, behavior, and/or knowledge through a citizen science program. For example, Crall et. al (2013) examined how invasive species training affects participants' knowledge about science and behavior towards the environment, and modest changes were found in both areas (Crall et al. 2013). Seymour and Haklay (2017) studied participation patterns in environmental volunteering and citizen science projects. Jordan et. al (2011) investigated changes in participants' knowledge and behavior towards invasive plants. They found that participants' knowledge of relevant issues increased and participants reported changes in their behavior with respect to invasive plants, specifically in the area of communicating with other people about invasive plants (Jordan et al. 2011). Most literature explores participant changes after engaging in outreach opportunities in which educators and/or researchers are directly interacting with participants (Adelman et al. 2000, Anderson et al. 2000, Bogner 1999, Brossard et al. 2005, Cockerill 2010, Crall et al. 2013). However, none of these studies have examined a hands-off approach in a drinking water quality monitoring setting in which researchers or related bodies do not have contact with participants. While many projects have recorded participants' demographic information to evaluate recruitment success (Bruyere and Rappe 2007, Jordan et al. 2011, Weston et al. 2003, Wright et al. 2015), to our understanding, none have analyzed the effect of demographic information on citizen's behavior, social networks, and/or knowledge related to the subject. Understanding participant demographics helps researchers target specific audiences or adjust their methods in order to reach a wider spectrum of participants and achieve the maximum social impact (West and Pateman 2016).

In light of the limitation of the previous studies, our study applied crowdsourcing to drinking water monitoring at the consumer taps to understand how socioeconomic characteristics affect participants' knowledge about drinking water quality and lead contamination, willingness to take preventative actions to improve health protection, and frequency of communication about water quality issues with those around them. It differs from current research as we investigated the effects of a hands-off approach to public engagement in which researchers had no direct interactions with participants. This project aims to provide an enhanced understanding regarding whether and how the crowdsourcing method could be used to actively engage and inform citizen participants in water quality monitoring efforts, creating a more scientifically literate and active public.

CHAPTER II: Methods

2.1 Project Overview and Survey Data Treatment

This project took place in a New Hampshire city during August-November 2018. Five informational kiosks distributed at highly trafficked locations invited participants to bring home a sampling packet which contained a 50 mL sample vial, sampling instructions, a sample information sheet, and a pre-survey, all housed in a pre-paid mailing envelope. Citizen participants were instructed to bring the packet home, collect a sample of their home's tap water according to instructions, and complete the sample information sheet and pre-survey. The participants could either return their completed packet to a kiosk or mail it to researchers at the University of New Hampshire (UNH) using the pre-paid mailing envelope. Researchers at UNH then preserved and analyzed returned samples for lead concentrations using an Inductively Coupled Plasma Mass Spectrometer (ICP-MS). The lead results and an online post-survey made using Qualtrics® were sent to participants via email. A more detailed description of the project design can be found in Jakositz et al. (2019) (Jakositz et al. 2019).

Out of approximately 800 packets distributed, the project received 149 returned packets with 136 packets containing pre-surveys that had more than 50% of questions answered. All participants who submitted a packet were asked to complete the post-survey, and 42 post-surveys were submitted with more than 50% questions answered. Of the 42 post-surveys, 36 were matched with their participant's respective pre-survey for comparison. Descriptive and bivariate statistics were conducted using IBM SPSS Version 25 and JMP Pro 14.

2.2 Multivariate analysis to determine usability of socioeconomic data in regression analysis

A multivariate analysis was performed in JMP to measure correlations between socioeconomic variables to determine whether they were fit to be used in analysis. The socioeconomic data collected in the pre-survey was normalized to allow for the comparison of values with different units (Motulsky and Christopoulos 2004). Age was normalized by dividing each age by that of the oldest participant, 88 years old, resulting in a normalized value between zero and one. Income was normalized by dividing the average income in the selected range by the largest numerical income selection, $200,000, for a normalized value between zero and one. Education was normalized by assigning each of ten options a value of one through ten, with option “A,” “No schooling completed,” receiving a value of one up to option “J,” “PhD and higher,” receiving a value of ten. These values were then each divided by ten to get a normalized value between zero and one. A higher normalized value corresponded to a more highly educated individual. Gender was recoded using a value of 0.25 for “Male”, 0.5 for “Female”, 0.75 for “Other,” and 1 for “Prefer not to answer.” Whether a participant rented or owned their home was recoded as 0 for “Rent,” 0.5 for “Own,” and 1 for “Other.” Lastly, whether or not the participant had any children under the age of six living in their home was recoded as 0 for “Yes” and 1 for “No.”

2.3 Analyzing the influence of participation on environmental literacy

To analyze the influence of the project on the participants’ environmental literacy, five true-or-false questions were included in both pre- and post-surveys. The design and the narrative of these questions were identical in both surveys. Table 1 provides the questions and response options that were presented to participants. Particularly, two questions related to the health impacts of lead, one question related to how most lead contamination enters drinking water, one question related to methods of removing lead from drinking water, and one question related to federal regulations regarding lead in drinking water. For each of these questions, participants were given three choices: true, false, or don't know. They were also advised not to guess the answer.

Table 1. Question included in pre- and post-surveys to understand participants' knowledge about lead-related water quality issues.
Please indicate whether you think the following statements are TRUE or FALSE by placing an X in the appropriate column. Please do not guess. If you are unsure about a statement, please check the "Don't Know" option.

	True	False	Don't Know
1. Lead consumption can cause damage to the brain, red blood cells, and kidneys.			
2. Young children and pregnant women are especially susceptible to the effects of lead.			
3. Lead enters drinking water through corrosion of pipelines and household fixtures.			
4. Boiling water is an effective method of removing lead.			
5. The U.S. Environmental Protection Agency (EPA) does not regulate lead in public drinking water.			

A correctness score was calculated for each question based upon the percentage of participants that answered the question correctly. Equation 1 presents the calculation of the correctness score.

Equation 1: $Correctness\ Score = \frac{Number\ of\ participants\ who\ answered\ correctly}{Total\ number\ of\ participants\ who\ submitted\ an\ answer}$

Responses for participants who completed both the pre- and post- surveys were matched and compared to look for changes potentially influenced by participation in this project. Paired Student's t-tests were then run to detect potential differences in responses to true-or-false questions of different socioeconomic groups.

After conducting the multivariate analysis to assess correlations between socioeconomic data, a stepwise regression analysis was then performed to determine how socioeconomic characteristics might affect a person's knowledge about lead-related drinking water concerns. Each participant's knowledge about lead-related drinking water concerns was characterized by an overall knowledge score. The overall knowledge score was calculated using Equation 2. A correct response to a given question earned the participant one point, and an incorrect response or a "Don't Know" response earned the participant a 0 for the given question. The overall knowledge score was calculated by averaging the scores from all answered questions. Questions without a response were not included in the overall knowledge score calculation. Participants were scored based on their responses to five True/False questions about lead in drinking water in the pre- and post-surveys.

Equation 2: $Knowledge\ Score = \frac{Number\ of\ Questions\ Correct}{Total\ Number\ of\ Questions\ (5)}$

Socioeconomic information was used as inputs for the stepwise regression. Age, income, and education were normalized as described in Section 2.2 for the multivariate analysis. Gender, whether the participant rented or owned their home, and whether or not the participant had any children under the age of six living in their home were coded as categorical values. JMP Pro 14 was used to run a stepwise regression analysis using Minimum BIC as the stopping rule, a forward direction, and combine rules.

2.4 Analyzing participants' responses towards potential tap water quality red flags

In the pre-survey, participants were asked, "How often do you contact your drinking water utility?" to determine the frequency that participants communicated with their utility. This question was compared with the participants' responses to the question, "How would you rate the quality of your drinking water at home?" using a chi-square analysis in which variables were re-coded into dichotomous variables to understand the relationship between perception about drinking water quality and intention to contact the utility.

To understand the participant's potential responses towards potential lead issues in tap water, a relevant question was included the pre- and post- surveys. The question was structured differently in the pre- and post-surveys. In the pre-survey, we asked participants, "Imagine your water quality results indicate that you have lead in your drinking water. How likely are you to take the following actions in response to learning about lead in your water?" In the post-survey, we asked participants, "Now that you have your water quality results, how likely are you to take the following actions in response to the lead in your drinking water?". By comparing the matched responses between the pre- and post-surveys, we were able to examine the potential changes of participants' actions in response to their knowledge of the actual water quality. Table 2 provides the format of the question as displayed in the pre-survey. We provided seven potential responses to drinking water quality issues based upon their prominence in literature. Participants were allowed to add additional types of responses. The surveys asked the participants to rate each of the responses on a Likert scale from highly unlikely, unlikely, neutral, likely, to highly likely.

Table 2. Pre- (a.) and post-survey (b.) questions asking participants about their likelihood to take various actions related to combating potential lead levels in their drinking water.

a. *Imagine your water quality results indicate that you have lead in your drinking water. How likely are you to take the following actions in response to learning about lead in your drinking water?*

b. *Now that you have your water quality results, how likely are you to take the following actions in response to the lead in your drinking water?*

	Highly Unlikely	Unlikely	Neutral	Likely	Highly Likely
1. Do nothing					
2. Install a filter					
3. Move to a new house					
4. Conduct additional research					
5. Tell my neighbors about the problem					
6. Flush my tap water before using it each day					
7. Contact my local water utility or environmental agency					
8. Other (*Please specify*):					

In the pre-survey, paired t-tests were used to detect potential differences in the responses of the various socioeconomic groups. For each question-socioeconomic characteristic pair, the X Factors were the recoded numeric responses to the action question (Table 3), and the Y Responses were the socioeconomic information. The post-survey responses could not be analyzed in the same way as instead of all participants assuming their water contains lead as in the pre-survey, participants in the post-survey were asked to consider their individual lead results. To analyze the likelihood of participants to take various actions according to their individual lead results, participants were broken into groups with lead levels less than 1 ppb, between 1 ppb and 10 ppb, and greater than 10 ppb. Paired t-tests were used to compare the action responses of the difference groups. Pre- and post-survey responses of participants who were found to have lead in their water were also compared to determine whether or not the participant's post-survey responses reflected the actions they said they would in the pre-survey.

Table 3. Recoded numeric values used in analysis of the action question.

Action Response	**Score**
Highly Likely	5
Likely	4
Neutral	3
Unlikely	2
Highly Unlikely	1

Similar to Section 2.2, a stepwise regression was conducted to determine how socioeconomic characteristics might affect a person's likelihood to take given actions regarding lead-related drinking water concerns. For each participant, an overall score was calculated to determine their willingness to take action in the pre-survey. Participant responses to the action response question were recoded according to Table 4. An overall score for their likelihood to take action was calculated using Equation 3. Those with higher scores were considered more likely to take action in response to having lead in their drinking water.

Equation 3: $Action\ Score = Average(Scored\ responses)$

Participants who did not respond to any of the potential actions were not given an action score. Participants who responded to some actions but not others were scored based on the responses they provided and non-responses were ignored in the averaging calculation.

As the post-survey question regarding participants' actions involved asking the participant to consider their personal lead result in their response, post-survey action responses were not included in this analysis.

Table 4. Recoded responses to pre-survey question: Imagine your water quality results indicate that you have lead in your drinking water. How likely are you to take the following actions in response to learning about lead in your drinking water?

Action Response	**Score**
Highly Likely	1
Likely	0.5
Neutral	0
Unlikely	-0.5
Highly Unlikely	-1

Table 5. Question included in pre- and post-surveys to understand participants' likelihood to communicate with their social network about water quality.
How often do you talk to the following groups of people about drinking water quality? Please place an X in the appropriate column to indicate how often you talk with each group about drinking water quality.

	Never	Rarely	Sometimes	Often	Every Time
1. Family					
2. Friends					
3. Neighbors					
4. Colleagues					
5. Strangers					
6. *Other (Please specify):*					

Paired t-tests were run on both pre- and post-survey responses to detect potential differences in the social network responses of various socioeconomic groups. For each question-demographic pair, the X Factors were the recoded numeric responses to social network questions (Table 6), and the Y Responses were the socioeconomic information.

Table 6. Recoded numeric values used in analysis of the social network question.

Social Network Response	**Score**
Every time	5
Often	4
Sometimes	3
Rarely	2
Never	1

A stepwise regression analysis was conducted to determine how socioeconomic characteristics might affect a person's likelihood to talk to different groups of people about their water quality. For each participant, an overall score was calculated to determine their willingness to interact with and/or grow their social network in the pre- and post-surveys. Responses were recoded according to Table 7. An overall score for their willingness to talk to various groups of people was calculated according to Equation 4. Those with higher scores were considered more likely to talk to others about drinking water quality.

Equation 4: $Social\ Network\ Score = Average(Score\ responses)$

Table 7. Recoded responses to pre- and post-survey questions: How often do you talk to the following groups of people about drinking water quality?

Social Network Response	Score
Every time	1
Often	0.5
Sometimes	0
Rarely	-0.5
Never	-1

Participants who did not provide responses for any of the social network groups were not given a social network score. Participants who responded to some social network groups but not others were given a score based on their provided responses.

CHAPTER III: Results and Discussions

In the pre-survey, 66% of participants were female, and the median age of participants was 55 which is 14 years older than the city-wide median age of 41. 22% of participants reported having children under the age of six in their home, and 70% owned their home. The median income of participants in this study was below the median income in the testbed city, and there was a greater percentage of participants with higher education than are in the testbed city. Similarly, in the post-survey, 68% of participants were female and the median age was 52. 21% of participants reported having children under the age of six in their home, and 71% owned their home. Like the pre-survey, the median income of participants was below the median income of the testbed city, and there was a greater percentage of participants with higher education than are in the testbed city.

In the pre-survey, 77.3% of participants indicated that they knew where their water came from and were able to indicate the source as the city's water treatment plant (45.6%), a community shared well (7.4%), or a private well (24.3%). 5.1% of participants indicated that they did not know where their water came from. The high missing response rate of 16.2% may indicate that a higher percentage of people may not know where their water comes from than just those who reported.

3.1 Multivariate analysis to determine usability of socioeconomic data in regression analysis

The multivariate analysis indicated that the socioeconomic variables were not significantly correlated and could therefore be used for further analysis (Table 8).

Table 8. Multivariate analysis correlations of normalized socioeconomic variables

	Gender Normalized	Age Normalized	Education Normalized	Children Normalized	Rent Normalized	Income Normalized
Gender Normalized	1	-0.24727	0.014388	-0.12747	-0.07673	-0.12642
Age Normalized	-0.24727	1	-0.04853	0.389516	0.273066	-0.08281
Education Normalized	0.014388	-0.04853	1	0.107996	0.123393	0.434787
Children Normalized	-0.12747	0.389516	0.107996	1	0.350444	0.136565
Rent Normalized	-0.07673	0.273066	0.123393	0.350444	1	0.348136
Income Normalized	-0.12642	-0.08281	0.434787	0.136565	0.348136	1

3.2 Analyzing the influence of participation on environmental literacy

Overall, respondents had high content knowledge about important health-related issues with lead in drinking water related to the physical impacts and risks to reproductive health prior to participating in the project (Figure 1). However, 51% of respondents did not know if the EPA regulates lead in public drinking water and 14% falsely indicated that the EPA does not regulate lead in public drinking water. Around 38% of respondents did not know if boiling water was an effective method of removing lead from drinking water, and 10% falsely indicated that boiling does remove lead. These two points represent opportunities for sharing relevant information about lead regulation and protection measures. Comparing the pre- and post-survey results, knowledge about the physical effects of lead consumption was high and consistent in both the pre- and post-surveys. There was less uncertainty in the post-survey about how lead enters drinking water through pipe corrosion and about the role of the EPA in regulating drinking water though these changes were not significant. There was a significant change in knowledge about the efficacy of boiling water as a strategy, as more respondents in the post-survey identified that this was not effective for removing lead after the project ($X^2 = 19.422$, $p<0.001$). These results indicate that a project such as ours may contribute to changes in content knowledge about drinking water quality and the effectiveness of specific strategies to flush, filter, and remove lead.

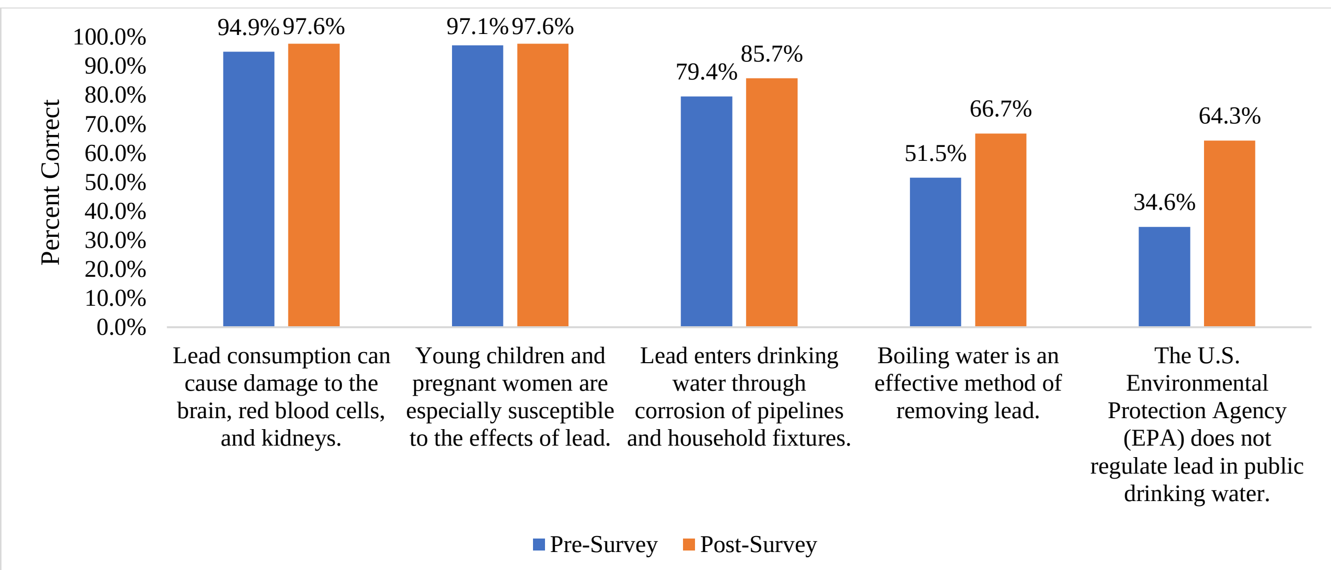


Figure 1. Relative frequency in survey responses that assessed content knowledge before and after the project.

T-tests comparing the overall knowledge scores of pre- and post-survey participants to each of the demographic subgroups yielded no significant results. This suggests that demographics did not have an effect on the participant's knowledge about lead issues before or after participating in the project. The results from the t-tests have been provided in the supporting information. Performing a stepwise regression analysis in an attempt to develop a predictive model relating demographic information to participant responses to knowledge questions yielded no significant results. This suggests that overall no demographic had a leading influence on a participant's knowledge about lead water quality issues.

When asked whether this project improved participants' understanding about lead in drinking water, 81.0% of participants answered "Yes" and 16.7% answered "No". This indicates that this project was successful in educating a majority of participants on the concerns about lead in drinking water and potential ways to mitigate them. This suggests that methods like those employed in this project may be useful in achieving public education recommendations set forth by the National Drinking Water Advisory Council (U.S. Environmental Protection Agency 2016).

3.3 Analyzing participants' responses towards potential tap water quality red flags

In the pre-survey, participants were asked, "How often do you contact your drinking water utility?" 98.4% of all participants, both well water and public supply, indicated that they never contact their water utility, and 90.3% of participants on the public water supply indicated that they never contact their public water utility. Only ten respondents, six of them on public water, indicated that they ever contact their public water utility (Table 9). This demonstrates gaps in communication between citizens and drinking water utilities, as nearly all citizens indicate that they never contact their public water utility, even when this is their main water supply. This question was compared with the participants' responses to the pre-survey question, "How would you rate the quality of your drinking water at home?" and there was no statistical significance (chi-square, re-coded into dichotomous variables) between perception about drinking water quality and intention to contact the utility. This suggests that even if people think their water is unsafe, this does not mean they will contact their utility.

Table 9. Self-reported frequency of contacting public water utility for all respondents and a subset who reported that they are on a public water supply.

	All respondents		**Public water supply**	
Contact	**N**	**%**	**N**	**%**
Never	124	98.4%	56	90.3%
Every couple of years	7	5.6%	5	8.1%
Every year	2	1.6%	1	1.6%
Multiple times per year	1	0.8%	0	0
Missing	2	1.6%	N/A	N/A
Total	**136**	100%	**62**	100%

When asked about what they would do if they learned they had lead in their drinking water in the pre-survey, participants indicated that they would be most likely to flush their tap water and install a filter (Table 10). Participants indicated that they would be least likely to do nothing or move to a new home. In the post- survey, participants said they would be most likely to flush their tap water and do nothing and would be least likely to tell their neighbors and move. Though the result in the pre-survey indicating that they would also contact their local utility was high, the fact that this was the third most likely and that almost a third of respondents indicated that they were either neutral (12%) or not likely (18.4%) to contact the utility or environmental agency is noteworthy, as it further demonstrates gaps in communication between citizens and drinking water utilities.

Table 10. Pre- and post-survey responses to a question that asked participants to imagine how they would act if they learned they had lead in their drinking water.

	Actions	N	Highly likely	Likely	Neutral	Unlikely	Highly unlikely
Pre-Survey	Flush my tap water before using it each day	136	58.1%	30.9%	6.6%	2.2%	0.7%
	Install a filter	136	47.8%	36.8%	9.6%	2.2%	1.5%
	Tell my neighbors about the problem	136	43.4%	33.1%	8.1%	2.9%	9.6%
	Contact my local water utility or environmental agency	136	41.9%	25.0%	11.8%	6.6%	11.8%
	Conduct additional research	136	30.9%	40.4%	10.3%	7.4%	7.4%
	Do nothing	136	3.7%	7.4%	2.2%	14.0%	69.9%
	Move to a new house	136	3.7%	2.9%	14.7%	16.2%	60.3%
Post-Survey	Flush my tap water before using it each day	36	44.4%	27.8%	8.3%	0.0%	11.1%
	Install a filter	36	11.1%	30.6%	11.1%	22.2%	19.4%
	Tell my neighbors about the problem	36	11.1%	27.8%	19.4%	11.1%	22.2%
	Conduct additional research	36	11.1%	27.8%	16.7%	19.4%	16.7%
	Do nothing	36	27.8%	16.7%	5.6%	16.7%	19.4%
	Move to a new house	36	0.0%	2.8%	8.3%	13.9%	61.1%

3.4 Differences in responses to action questions between socioeconomic groups

The t-test results of the pre-survey responses show that those with trade/technical/vocational training were more likely to do “Research” than those with other educational backgrounds, while those with a Ph.D. and higher were overall less likely to do “Research” and more likely to “Do Nothing” than those with less education. This is counterintuitive as it suggests than those with higher education may be less likely to take productive measures toward improving their water quality.

It has to be noted that there were very few participants who claimed to have a Ph.D. or higher (pre-survey N = 4, post-survey N = 3) which may suggest that these results may not be indicative of a larger population. Of the significant pairs in the pre-survey and post-survey t-test analyses of actions, social network, and knowledge questions relating to demographic subgroups (SI Tables S2 – S6), many had responses with low N values. Though the p-values indicate that the pairs are significantly different, those with low p-values may not be representative of the population and therefore may not be notable results. Pairs where either group had an N value of one were excluded from the results. Those with N values of two or greater were kept in reporting tables, however it should be noted that those with relatively smaller effect sizes may not be as informative as those with larger effect sizes (De Winter 2013). For the action and social network questions, a small effect size is considered to be anything less than 0.5, indicating that there is no evidence of a difference in response levels for that question.

Additionally, there was an overall positive correlation between education and the likelihood for the participants to “Tell Neighbors” about their drinking water concerns. This might be because those with higher education may see more value in sharing water quality information with those directly around them who may also be affected by the issue.

Performing a stepwise regression analysis in an attempt to develop a predictive model relating demographic information to participant responses to action questions in the pre-survey yielded no significant results. This suggests that overall no demographic had a leading influence on a participant's likelihood to take action should they find they have lead in their water.

3.5 T-tests assessing post-survey responses to action questions grouped by lead result

Paired t-tests were run comparing demographics of participants with reported lead levels of < 1 ppb, 1-10 ppb, and ≥ 10 ppb (SI Table S2). More than half of participants in the post-survey had lead concentrations < 1ppb (Table 11). The results suggested that participants with 1-10ppb were more likely to "Do Nothing" than those with ≥ 10ppb, with an average scoring difference of 2.03 and a significant p-value of 0.0383. Though not significant, there was also a notable average difference of 1.47 between those with < 1ppb and those with ≥ 10ppb, suggesting that those with < 1ppb were more likely to "Do Nothing" than those with ≥ 10ppb.

Table 11. Number of post-survey participants in each of three lead concentration groups.

Individual Lead Result	N	%
< 1ppb	20	55.6%
1-10 ppb & < 10ppb	12	33.3%
≥ 10ppb	4	11.1%
Missing	0	0.0%
Total	**36**	100.0%

The only other significant pair in this analysis regarded likelihood to do "Research" after finding out about their lead result. Those with ≥ 10ppb were more likely to do "Research" than those with < 1ppb with an average difference of 1.47 and a significant p-value of 0.0468. Though not significant, it is notable that those with ≥ 10ppb were also more likely to do "Research" than those with ≥ 1ppb & < 10ppb with an average difference of 1.43. This suggests that those with the highest levels of lead were the most likely to do research upon receiving their results. Though the type or level of research was not asked in this question, this suggests a positive outcome that those with potentially dangerous levels of lead are most likely to educate themselves more about the issue.

Though not significant in terms of p-value, another notable pair was the comparison of ≥ 10ppb to < 1ppb when looking at the "Tell Neighbors" action. Those with ≥ 10ppb were more likely to tell their neighbors about their lead levels than those with < 1ppb, with an average difference of 1.03. This suggests that participants with potentially dangerous levels of lead in their water were more likely to warn others who may be unknowingly ingesting similar water. Similarly, those with ≥ 10ppb were more likely to "Flush" their tap than those with ≥ 1ppb & < 10ppb, with an average difference of 0.917. This suggests that even though their results indicated that they have lead in their water, participants are less likely to adopt the habit of flushing their taps unless the lead is at a higher level.

In comparing the pre- and post-survey responses of participants with ≥ 10ppb, three out of four participants' overall action scores decreased, meaning three out of four of them were less likely to take action in response to their lead results. Most notably, the average score for likelihood to install a "Filter" decreased from 4.5 (Highly Likely/Likely) in the pre-survey to 3 (Neutral) in the post-survey. Likelihood to "Tell Neighbors" about their lead results decreased on average from 5 (Highly Likely) in the pre-survey, to 3.75 (Neutral/Likely) in the post-survey. These results may indicate a difference between intended actions and actual actions with regards to water quality. However, it is important to note that two of the four post-survey participants with lead levels greater than 10ppb were below the Lead and Copper Rule's Action Level of 15ppb. These two participants both were among the three that saw decreases in likelihood to take action in response to their actual water quality. This may suggest that these participants may not be as concerned with their lead results as they are below the Action Level, a potential danger of assigning a limit that may not necessarily be indicative of exposure health implications.

3.6 Analyzing the influence of participation on the participants' social networks

In the pre- and post-surveys, participants were asked, "How often do you talk to the following groups of people about drinking water quality?" Participants in the post-survey were more likely to communicate with all groups than those in the pre-survey. In both the pre- and post-surveys, the groups that participants most frequently communicated with regarding water quality were family and friends. This makes sense as these are the groups that participants likely have the most contact with. Also, previous research found that the strongest motivation for participation in this project was to protect the health of oneself and one's family (Jakositz et al. 2019), which may also explain why family was the group most frequently communicated with. In both surveys, participants were least likely to communicate with strangers and colleagues. This may suggest that most conversation about drinking water quality takes place in or around the home.

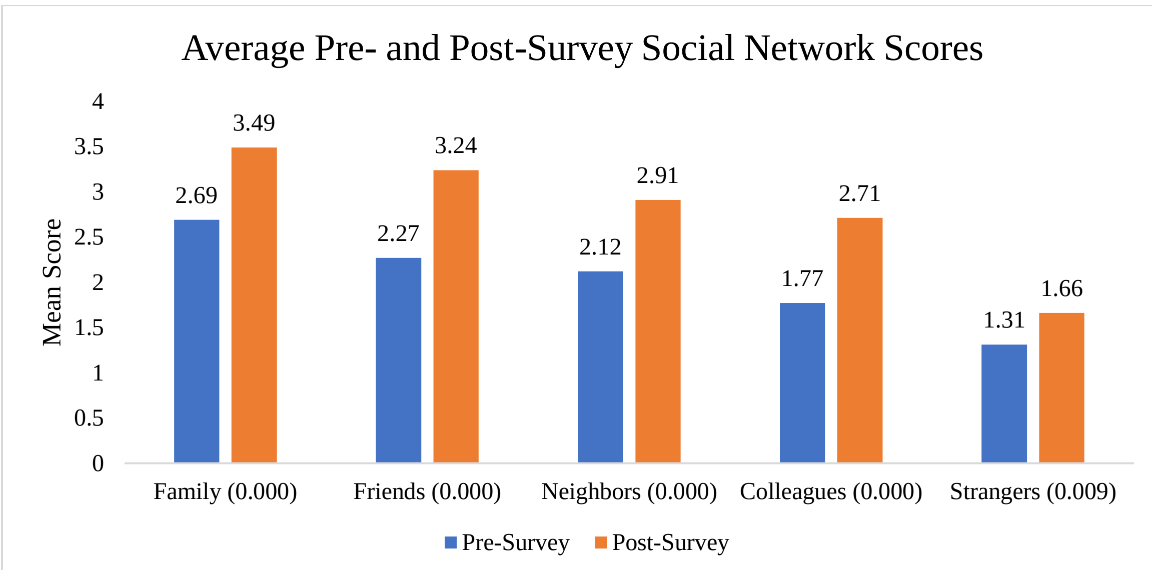


Figure 2. Frequency analysis of responses to pre- and post-survey question, "How often do you talk to the following groups of people about drinking water quality?" Values in parenthesis indicate are p-values indicating significant differences between pre- and post-survey responses.

In the t-test results comparing matched pre- and post-survey responses to the communication question, there were statistically significant positive changes in participants' likelihood to communicate with all groups (Figure 2). The greatest average change was seen in the likelihood of participants to communicate with friends and colleagues about drinking water quality, with average differences of 0.97 and 0.94, respectively. This means that most participants increased their self-reported likelihood by one level, with most going from a 2, "Often" to a 3, “Sometimes.” The least change was seen in participants’ likelihood to talk to strangers, with an average difference of 0.34.

In the pre-survey, the most notable differences between demographic subgroup pairs regarding the social network question occurred between rent/own/other groups (SI Table S5). Those who indicated their living situation as “Other” were more likely to talk to their colleagues than those who rented or owned their homes, with average differences of 2.7 and 2.3, respectively. Those who owned their homes were less likely to talk to strangers about water quality issues than those who indicated “Other” or rented their home, with average differences of 1.7 and 0.5, respectively.

Another notable difference in the pre-survey analysis was seen with income, with those making $0 to $24,999 per year more likely to communicate with strangers than those making more, with significant differences ranging from 0.7 to 1.0. This suggests that those making more money may be less likely to reach outside of their regular social groups to communicate about drinking water quality than those with incomes on the low end of the spectrum.

Looking at the post-survey results, all significant pairs but one involved income (SI Table S6). Those with $125,000 to $149,999 were more likely to communicate with family, friends, and colleagues than those in groups with household incomes between $25,000 and $124,999. Average differences ranged from 1.5 to 3.5, with the 3.5 comparing $125,000 to $149,999 to those with $100,000 to $124,999 and their likelihood to communicate with colleagues.

As mentioned previously in Section 3.4, it is important to note that certain socioeconomic characteristics had small N values and so results from this analysis may not be indicative of a larger population.

Performing a stepwise regression analysis in an attempt to develop a predictive model relating demographic information to participant responses to social network questions for both the pre- and post-surveys separately yielded no significant results. This suggests that overall no demographic had a leading influence on a participant's likelihood to communicate about drinking water quality.

CHAPTER IV: Conclusions

This project demonstrated that a well-designed crowdsourcing approach can be used to engage participants in water quality activities that both benefit the citizen and the researchers. Comparing pre- and post-surveys showed statistically significant increases in participants' likelihood to communicate about drinking water with family, friends, neighbors, colleagues, and strangers. By assessing communication frequency by socioeconomic groups, this study found that different demographics may be more likely to communicate with different social network groups. This finding is important as it may impact how a project might choose to attempt to spread information. When looking at actions taken in response to water quality issues, researchers found that income, age, and educational groups had some of the largest, significant, paired differences. With participants least likely to do nothing or move to a new house, it was notable that most participants would be willing to take preventative actions to protect their health against water quality concerns. With regards to knowledge, this project demonstrated success in improving citizen's scientific literacy relating to key lead information, and overall provided educational benefit to those who participated.

LIST OF REFERENCES

Adelman, L.M., Falk, J.H. and James, S. (2000) Impact of National Aquarium in Baltimore on visitors' conservation attitudes, behavior, and knowledge. Curator: The Museum Journal 43(1), 33-61.

Anderson, D., Lucas, K.B., Ginns, I.S. and Dierking, L.D. (2000) Development of knowledge about electricity and magnetism during a visit to a science museum and related post-visit activities. Science education 84(5), 658-679.

Bogner, F.X. (1999) Empirical evaluation of an educational conservation programme introduced in Swiss secondary schools. International Journal of Science Education 21(11), 1169-1185.

Bonney, R., Cooper, C.B., Dickinson, J., Kelling, S., Phillips, T., Rosenberg, K.V. and Shirk, J. (2009) Citizen science: a developing tool for expanding science knowledge and scientific literacy. BioScience 59(11), 977-984.

Bonney, R., Shirk, J.L., Phillips, T.B., Wiggins, A., Ballard, H.L., Miller-Rushing, A.J. and Parrish, J.K. (2014) Next steps for citizen science. Science 343(6178), 1436-1437.

Brossard, D., Lewenstein, B. and Bonney, R. (2005) Scientific knowledge and attitude change: The impact of a citizen science project. International Journal of Science Education 27(9), 1099-1121.

Bruyere, B. and Rappe, S. (2007) Identifying the motivations of environmental volunteers. Journal of Environmental Planning and Management 50(4), 503-516.

Cockerill, K. (2010) Communicating How Water Works: Results From a Community Water Education Program. The Journal of Environmental Education 41(3), 151-164.

Conrad, C.C. and Hilchey, K.G. (2011) A review of citizen science and community-based environmental monitoring: issues and opportunities. Environmental monitoring and assessment 176(1-4), 273-291.

Crall, A.W., Jordan, R., Holfelder, K., Newman, G.J., Graham, J. and Waller, D.M. (2013) The impacts of an invasive species citizen science training program on participant attitudes, behavior, and science literacy. Public Understanding of Science 22(6), 745-764.

Cronje, R., Rohlinger, S., Crall, A. and Newman, G. (2011) Does participation in citizen science improve scientific literacy? A study to compare assessment methods. Applied Environmental Education & Communication 10(3), 135-145.

De Winter, J.C. (2013) Using the Student's t-test with extremely small sample sizes. Practical Assessment, Research & Evaluation 18(10).

Den Broeder, L., Devilee, J., Van Oers, H., Schuit, A.J. and Wagemakers, A. (2016) Citizen Science for public health. Health promotion international 33(3), 505-514.

Dickinson, J.L., Shirk, J., Bonter, D., Bonney, R., Crain, R.L., Martin, J., Phillips, T. and Purcell, K. (2012) The current state of citizen science as a tool for ecological research and public engagement. Frontiers in Ecology and the Environment 10(6), 291-297.

Goovaerts, P. (2019) Geostatistical prediction of water lead levels in Flint, Michigan: A multivariate approach. Science of the total environment 647, 1294-1304.

Howe, J. (2006) The rise of crowdsourcing. Wired magazine 14(6), 1-4.

Jakositz, S., Pillsbury, L., Greenwood, S., Fahnestock, M., McGreavy, B., Bryce, J. and Mo, W. (2019) Protection through participation: Crowdsourced tap water quality monitoring for enhanced public health. Water Research, 115209.

Jeppesen, L.B. and Lakhani, K.R. (2010) Marginality and problem-solving effectiveness in broadcast search. Organization science 21(5), 1016-1033.

Jollymore, A., Haines, M.J., Satterfield, T. and Johnson, M.S. (2017) Citizen science for water quality monitoring: Data implications of citizen perspectives. Journal of environmental management 200, 456-467.

Jordan, R.C., Gray, S.A., Howe, D.V., Brooks, W.R. and Ehrenfeld, J.G. (2011) Knowledge gain and behavioral change in citizen-science programs. Conservation Biology 25(6), 1148-1154.

Malone, T.W., Laubacher, R. and Dellarocas, C. (2010) The collective intelligence genome. MIT Sloan Management Review 51(3), 21.

Motulsky, H. and Christopoulos, A. (2004) Fitting models to biological data using linear and nonlinear regression: a practical guide to curve fitting, Oxford University Press.

Nerbonne, J.F. and Nelson, K.C. (2004) Volunteer macroinvertebrate monitoring in the United States: resource mobilization and comparative state structures. Society and Natural Resources 17(9), 817-839.

Pandya, R.E. (2012) A framework for engaging diverse communities in citizen science in the US. Frontiers in Ecology and the Environment 10(6), 314-317.

Roy, S. and Edwards, M. (2019) Citizen Science During the Flint, Michigan Federal Water Emergency: Ethical Dilemmas and Lessons Learned.

Shirk, J., Ballard, H., Wilderman, C., Phillips, T., Wiggins, A., Jordan, R., McCallie, E., Minarchek, M., Lewenstein, B. and Krasny, M. (2012) Public participation in scientific research: a framework for deliberate design. Ecology and society 17(2).

Silvertown, J. (2009) A new dawn for citizen science. Trends in ecology & evolution 24(9), 467-471.

Snik, F., Rietjens, J.H., Apituley, A., Volten, H., Mijling, B., Di Noia, A., Heikamp, S., Heinsbroek, R.C., Hasekamp, O.P. and Smit, J.M. (2014) Mapping atmospheric aerosols with a citizen science network of smartphone spectropolarimeters. Geophysical Research Letters 41(20), 7351-7358.

Trumbull, D.J., Bonney, R., Bascom, D. and Cabral, A. (2000) Thinking scientifically during participation in a citizen-science project. Science education 84(2), 265-275.

U.S. Environmental Protection Agency (2016) Lead and Copper Rule Revisions. Water, O.o. (ed), U.S. Environmental Protection Agency, EPA.gov.

U.S. Environmental Protection Agency (2018) Basic Information about Lead in Drinking Water, U.S. Environmental Protection Agency.

US Environmental Protection Agency Office of Water (2015) 21st Century Science Challenges for EPA's National Water Program: An Update to the National Water Program Research Strategy.

West, S.E. and Pateman, R.M. (2016) Recruiting and retaining participants in citizen science: What can be learned from the volunteering literature? Citizen Science: Theory and Practice.

Weston, M., Fendley, M., Jewell, R., Satchell, M. and Tzaros, C. (2003) Volunteers in bird conservation: insights from the Australian Threatened Bird Network. Ecological Management & Restoration 4(3), 205-211.

Wiggins, A. and Crowston, K. (2011) From conservation to crowdsourcing: A typology of citizen science, pp. 1-10, IEEE.

Wright, D.R., Underhill, L.G., Keene, M. and Knight, A.T. (2015) Understanding the motivations and satisfactions of volunteers to improve the effectiveness of citizen science programs. Society & Natural Resources 28(9), 1013-1029.

APPENDICES

APPENDIX A: UNH INSTITUTIONAL REVIEW BOARD (IRB) APPROVAL LETTER

University *of* New Hampshire

Research Integrity Services, Service Building
51 College Road, Durham, NH 03824-3585
Fax: 603-862-3564

05-Jul-2017

Mo, Weiwei
Civil and Environmental Engineering
Gregg Hall
Durham, NH 03824-3521

IRB #: 6735
Study: EAGER:PPER: Development of a Contest-Based Crowdsourcing Scheme for Public Water Quality Monitoring
Approval Date: 03-Jul-2017

The Institutional Review Board for the Protection of Human Subjects in Research (IRB) has reviewed and approved the protocol for your study as Exempt as described in Title 45, Code of Federal Regulations (CFR), Part 46, Subsection 101(b). Approval is granted to conduct your study as described in your protocol.

Researchers who conduct studies involving human subjects have responsibilities as outlined in the document, *Responsibilities of Directors of Research Studies Involving Human Subjects.* This document is available at http://unh.edu/research/irb-application-resources. Please read this document carefully before commencing your work involving human subjects.

Upon completion of your study, please complete the enclosed Exempt Study Final Report form and return it to this office along with a report of your findings.

If you have questions or concerns about your study or this approval, please feel free to contact me at 603-862-2003 or Julie.simpson@unh.edu. Please refer to the IRB # above in all correspondence related to this study. The IRB wishes you success with your research.

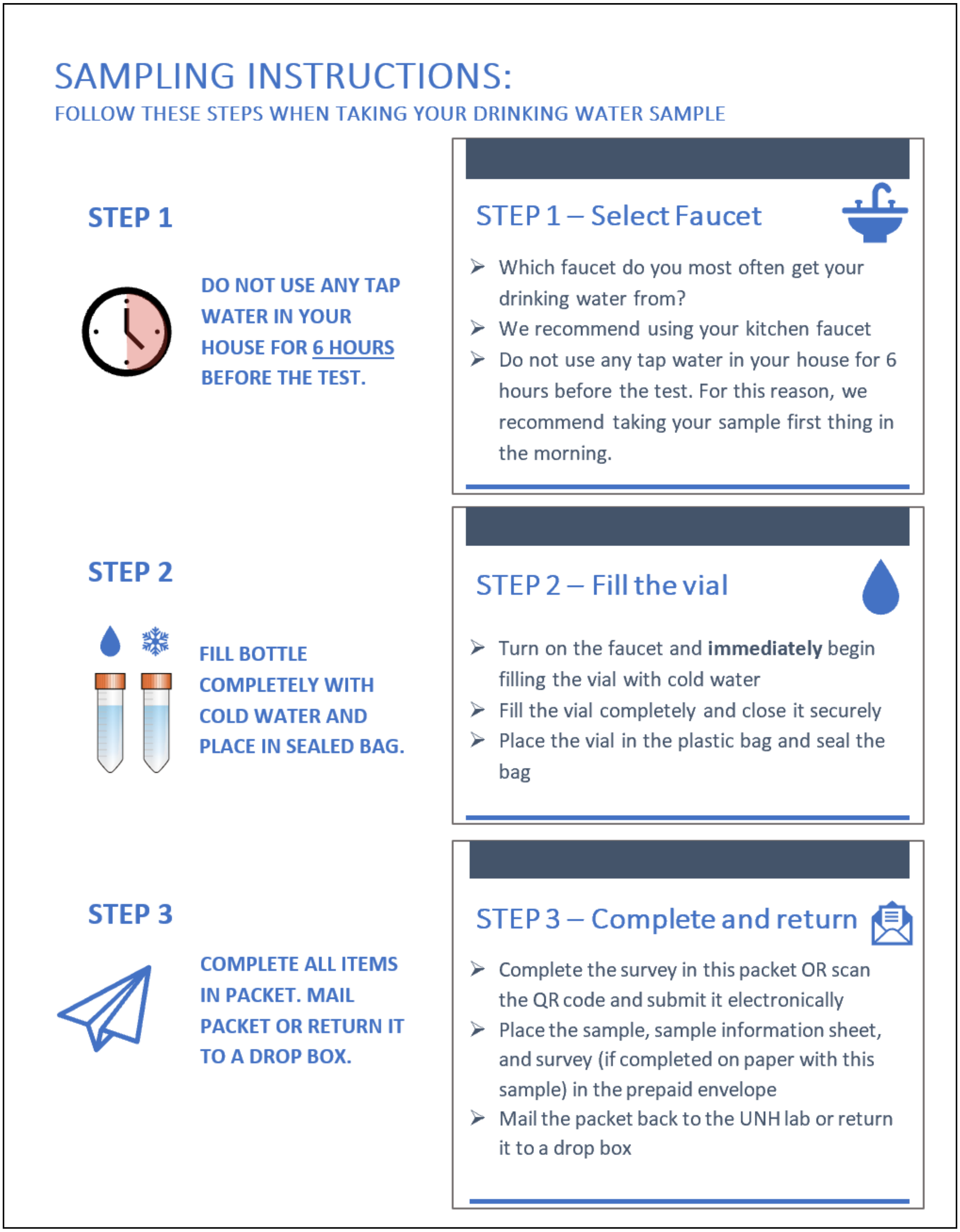
For the IRB,

Julie F. Simpson
Director

cc: File

APPENDIX B: PART I SUPPORTING INFORMATION

Contents:



SAMPLING INSTRUCTIONS:

FOLLOW THESE STEPS WHEN TAKING YOUR DRINKING WATER SAMPLE

STEP 1

DO NOT USE ANY TAP WATER IN YOUR HOUSE FOR 6 HOURS BEFORE THE TEST.

STEP 1 – Select Faucet

- Which faucet do you most often get your drinking water from?
- We recommend using your kitchen faucet
- Do not use any tap water in your house for 6 hours before the test. For this reason, we recommend taking your sample first thing in the morning.

STEP 2

FILL BOTTLE COMPLETELY WITH COLD WATER AND PLACE IN SEALED BAG.

STEP 2 – Fill the vial

- Turn on the faucet and **immediately** begin filling the vial with cold water
- Fill the vial completely and close it securely
- Place the vial in the plastic bag and seal the bag

STEP 3

COMPLETE ALL ITEMS IN PACKET. MAIL PACKET OR RETURN IT TO A DROP BOX.

STEP 3 – Complete and return

- Complete the survey in this packet OR scan the QR code and submit it electronically
- Place the sample, sample information sheet, and survey (if completed on paper with this sample) in the prepaid envelope
- Mail the packet back to the UNH lab or return it to a drop box

Figure S1. Instructions provided to participants in the sampling packet regarding how to take the tap water sample and return their completed packet to researchers.

Table S1. Results of a Pearson Chi-Square analysis showing no statistically significant relationship between presence of lead in drinking water and perceptions about drinking water quality.

		Drinking water lead concentration (≥ 1 ug/L = contains lead)		**Total**
		No Lead	Lead	
Perception	Bad	44	20	64
	Good	46	18	64
	Total	90	64	128
		Value	**df**	**Asymptotic Significance (2-sided)**
Pearson Chi-Square		.150[a]	1	0.699

Table S2. Results of a Pearson Chi-Square analysis showing no statistically significant relationship between perceptions about drinking water quality and choice to run the tap in the pre-survey responses.

		Drinking water quality perception		**Total**
		Bad	Good	
Run tap	Yes	27	40	51
	No	40	40	80
Total		67	64	131
		Value	**df**	**Asymptotic Significance (2-sided)**
Pearson Chi-Square		.108[a]	1	0.743

Table S3. Results of descriptive analysis identifying participant interest in the two contest schemes as per the pre-survey.

Which contest scheme do you hope to win?	**Number of Participants**	**Percent of Total Respondents**
Ambassador - highest number of participants recruited into the contest	12	8.8%
Go-Getter - highest number of samples from unique addresses collected by an individual	9	6.6%
I am **not interested** in winning a contest	101	74.3%
Unknown	14	10.3%

Table S4. Response to pre and post-survey question that asked participants to rate how much they agreed or disagreed with statements to assess motivations to participate in drinking water quality, (1=Strongly Disagree, 2-Disagree, 3=Neutral, 4=Agree, 5=Strongly Agree). The motivation scale is based on DEVISE framework. Mean motivations were mostly consistent across pre and post-survey responses. Paired t-test revealed no significant differences in overall motivations before and after the project, and a significant decrease (n=33, pre-mean 4.73, post-mean 4.3, t-3.677, df 32, $p<0.001$) in the motivation to learn about drinking water quality which may indicate that the project satisfied the interest in learning about drinking water quality.

"I'm interested in participating in this drinking water quality contest because…"

	Pre-survey				Post-survey			
	N	**Missing**	**Mean**	**SD**	**N**	**Missing**	**Mean**	**SD**
I want to learn about my drinking water quality	129	7	4.71	0.687	41	1	4.39	0.67
I care about my family's health	128	8	4.67	0.795	41	1	4.76	0.43
I care about my personal health	130	6	4.62	0.800	41	1	4.63	0.49
I am an environmentally conscious person	130	6	3.98	1.004	41	1	4.05	0.74
I enjoy doing activities related to environmental health	127	9	3.73	1.109	40	2	3.90	0.84
Participating in this contest will help me achieve things that are important to me	127	9	2.76	1.477	40	2	2.55	1.28
People I look up to think it's good to participate in this contest	127	9	2.24	1.288	40	2	2.10	1.10
Because I want to win the cash prize	126	10	1.95	1.258	39	3	1.90	0.97
I want others to think I'm good at doing activities related to environmental health	127	9	1.91	1.224	40	2	1.93	1.12
Of the recognition or respect I'll get from others	128	8	1.75	1.157	39	3	1.79	0.92

I am required to participate in this contest	126	10	1.40	0.931	40	2	1.63	0.81

Pre-Contest Survey

1. How did you learn about this contest? (Select all that apply)

a) Social media (i.e. Facebook)
b) Project flyer
c) Drop box station
d) Friend/family
e) Colleague
f) Other: ____

2. How would you rate the quality of your drinking water at home?

a) Excellent
b) Very good
c) Good
d) Fair
e) Poor

3. This project allows you to participate in one of two contests:

Contest #1: "**The Ambassador**," rewards the person who recruits the most participants

Contest #2: "**The Go-Getter**," rewards the person who collects the most samples from different addresses that have not yet been sampled for this study.

This survey is available online! Scan the QR code below or visit **https://ceps.unh.edu/rochester-water-testing**

If you completed this survey online or with a previous sample, please indicate so on the ***Sample Information***

Which contest scheme do you hope to win?

a) Ambassador – highest number of participants recruited into the contest
b) Go-Getter – highest number of samples from unique addresses collected by an individual
c) I am not interested in winning a contest

4. **In general, why are you interested in participating in this drinking water quality contest?**
Please indicate how much you **DISAGREE** or **AGREE** with each of the following statements by placing an **X** in the appropriate column. Please respond with how you really feel, rather than how you think "most people" feel.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1. Because I want to learn about my drinking water quality					
2. Because I want to win the cash prize					
3. Because of the recognition or respect I'll get from others					
4. Because I am an environmentally conscious person					
5. Because I care about my personal health					
6. Because I care about my family's health					
7. Because people I look up to think it's good to participate in this contest					
8. Because I am required to participate in this contest					
9. Because participating in this contest will help me achieve things that are important to me					
10. Because I enjoy doing activities related to environmental health					
11. Because I want others to think I'm good at doing activities related to environmental health					
12. Other (*Please specify*):					

5. Please indicate whether you think the following statements are TRUE or FALSE by placing an **X** in the appropriate column. Please do not guess. If you are unsure about a statement, please check the "Don't Know" option.

	True	False	Don't Know
1. Lead consumption can cause damage to the brain, red blood cells, and kidneys.			
2. Young children and pregnant women are especially susceptible to the effects of lead.			
3. Lead enters drinking water through corrosion of pipelines and household fixtures.			
4. Boiling water is an effective method of removing lead.			
5. The U.S. Environmental Protection Agency (EPA) does not regulate lead in public drinking water.			

6. Do you know where the tap water in your home comes from?

a) Yes
b) No

7. If you answered "Yes" to the above question, which of the options best completes the following sentence: *My water comes from…*

a) The Rochester water treatment plant
b) A community shared well
c) My own private well
d) I don't know
e) Other: ____

8. Do you run your tap to flush your water before using it each day?

a) Yes
b) No

9. How often do you contact your drinking water utility?

a) Never
b) Every couple of years
c) Every year
d) Multiple times per year

10. If you have ever contacted your drinking water utility, what was the purpose of the interaction(s)? (*Select all that apply*)

a) To ask a question about my bill
b) To ask a question about drinking water quality
c) To submit a complaint
d) To submit a compliment
e) To schedule a tour and visit the facility
f) To get to know utility staff
g) Other: ____

11. Imagine your water quality results indicate that you have lead in your drinking water. How likely are you to take the following actions in response to learning about lead in your drinking water? Please place an **X** in the appropriate column to indicate your likeliness to engage in these activities if you learn you have lead in your drinking water.

	Highly Unlikely	Unlikely	Neutral	Likely	Highly Likely
1. Do nothing					
2. Install a filter					
3. Move to a new house					
4. Conduct additional research					
5. Tell my neighbors about the problem					
6. Flush my tap water before using it each day					
7. Contact my local water utility or environmental agency					
8. Other (*Please specify*):					

12. **How often do you talk to the following groups of people about drinking water quality?** Please place an **X** in the appropriate column to indicate how often you talk with each group about drinking water quality.

	Never	Rarely	Some-times	Often	Every Time
1. Family					
2. Friends					
3. Neighbors					
4. Colleagues					
5. Strangers					
6. *Other (Please specify):*					

13. **What is your gender?**
   a) Male
   b) Female
   c) Other
   d) Prefer not to answer

14. **Please write your age: _________**

15. **What is your highest education?**
   a) No schooling completed
   b) Primary school to 8th grade
   c) Some high school, no diploma
   d) High school graduate, diploma or equivalent (e.g. GED)
   e) Some college credit, no degree
   f) Trade/technical/vocational training
   g) Associate degree
   h) Bachelor's degree
   i) Master's degree
   j) Ph.D. and higher

16. **Do you have any children under 6 years old living in your household?**
   a) Yes
   b) No

17. Do you rent or own your home?
a) Rent
b) Own
c) Other: ____

18. What is your approximate average household income?

a) $0 to $24,999	f) $125,000 to $149,999
b) $25,000 to $49,999	g) $150,000 to $174,999
c) $50,000 to $74,999	h) $175,000 to $199,999
d) $75,000 to $99,999	i) $200,000 and up
e) $100,000 to $124,999	

Post-Contest Survey (Survey has been slightly edited to respect the anonymity of the testbed city)

Q2. **Where did you get your sampling packet(s) from? (*Select all that apply*)**

- ☐ Walmart
- ☐ Hannaford
- ☐ Public Library
- ☐ Community Center
- ☐ Lilac Mall
- ☐ St. Mary's Church
- ☐ Farmers' Market
- ☐ Mail
- ☐ Friend or family member
- ☐ Other (Please specify): ____

Q3. **Did participating in this contest improve your understanding about lead in drinking water?**

- ○ Yes (1)
- ○ No (2)

Q4. This project allows you to participate in one of two contests: Contest #1: "**The Ambassador**," rewards the person who recruits the most participants. Contest #2: "**The Go-Getter**," rewards the person who collects the most samples from different addresses that have not yet been sampled for this study. **Which contest scheme do you hope to win?**

- ◯ Ambassador - highest number of participants recruited into the contest (1)
- ◯ Go-Getter - highest number of samples from unique addresses collected by an individual (2)
- ◯ I am not interested in winning a contest. (3)

Q5. **If you participated in the "Ambassador" contest scheme, how did you recruit participants?**

____

Q6. **If you participated in the "Go-Getter" contest scheme, how did you access sampled locations?**

____

Q7. **Would you be willing to participate in another project like this if given the opportunity?**

- ◯ Yes (1)
- ◯ Maybe (2)
- ◯ No (3)

If "No" was selected in Q7, Q8 was displayed.

Q8 **Why wouldn't you be interested in participating in another project like this?**

____

Q9. **Why would you be interested in participating in another drinking water quality project like this? Please indicate how much you <u>DISAGREE</u> or <u>AGREE</u> with each of the following statements by selecting the appropriate column. Please respond as you really feel, rather than how you think "most people" feel.**

	Strongly Disagree (1)	Disagree (2)	Neutral (3)	Agree (4)	Strongly Agree (5)
Because I want to learn about drinking water quality (1)	○	○	○	○	○
Because I want to win the cash prize (2)	○	○	○	○	○
Because of the recognition or respect I'd get from others (3)	○	○	○	○	○
Because I am an environmentally conscious person (4)	○	○	○	○	○
Because I care about my personal health (5)	○	○	○	○	○
Because I care about my family's health (6)	○	○	○	○	○
Because people I look up to would think it's good to participate in the contest (7)	○	○	○	○	○
Because I would be required to participate in the contest (9)	○	○	○	○	○

Because participating in the contest would help me achieve things that are important to me (10)	○	○	○	○	○
Because I enjoy doing activities related to environmental health (11)	○	○	○	○	○
Because I want others to think I'm good at doing activities related to environmental health (12)	○	○	○	○	○
Other (Please specify): (13)	○	○	○	○	○

Q10. **Please indicate whether you think the following statements are TRUE or FALSE by selecting the appropriate column.** Please do not guess. If you are unsure about a statement, please check the “Don’t Know” option.

	TRUE (1)	FALSE (2)	DON'T KNOW (3)
Lead consumption can cause damage to the brain, red blood cells, and kidneys. (2)	○	○	○
Young children and pregnant women are especially susceptible to the effects of lead. (3)	○	○	○
Lead enters drinking water through corrosion of pipelines and household fixtures. (6)	○	○	○
Boiling water is an effective method of removing lead. (7)	○	○	○
The U.S. Environmental Protection Agency (EPA) does not regulate lead in drinking water. (8)	○	○	○

Q11. **Now that you have your water quality results, how likely are you to take the following actions in response to the lead in your drinking water?** Please select the appropriate column to indicate your likeliness to engage in these activities.

	Highly Unlikely (1)	Unlikely (2)	Neutral (3)	Likely (4)	Highly Likely (5)
Do nothing (1)	○	○	○	○	○
Install a filter (2)	○	○	○	○	○
Move to a new house (3)	○	○	○	○	○
Conduct additional research (4)	○	○	○	○	○
Tell my neighbors about the problem (5)	○	○	○	○	○
Flush my tap water before using it each day (6)	○	○	○	○	○
Other (Please specify): (7)	○	○	○	○	○

Q12. **After participating in this project, how likely are you to talk to the following groups of people about drinking water quality? Please select the appropriate column.**

	Never (1)	Rarely (2)	Sometimes (3)	Often (4)	Every Time (5)
Family (1)	○	○	○	○	○
Friends (2)	○	○	○	○	○
Neighbors (3)	○	○	○	○	○
Colleagues (4)	○	○	○	○	○
Strangers (5)	○	○	○	○	○
Other (Please specify): (6)	○	○	○	○	○

Q13 **Please provide any additional feedback and comments about our project.**

____

____

____

____

____

Q14 **Please provide your NAME and EMAIL so that we can confirm your participation in this survey.**
Be sure to use the same name and email that you provided on the Sample Information Sheet.
Please note that participation in this survey is required to be eligible to win the contests.

Q15. Please fill in your **NAME:**

____

Q16. Please fill in your **EMAIL:**

____

APPENDIX C: PART II SUPPORTING INFORMATION

Contents:

Table S1. Frequency analysis of responses to pre- and post-survey question, “How often do you talk to the following groups of people about drinking water quality?”

Communication Group	Response	Pre-Survey		Post-Survey	
		Frequency	Percent	Frequency	Percent
Family	Every Time	7	5.1%	4	11.1%
	Often	26	19.1%	14	38.9%
	Sometimes	48	35.3%	12	33.3%
	Rarely	28	20.6%	5	13.9%
	Never	24	17.6%	0	0.0%
	Missing	3	2.2%	1	2.8%
	Total	136	100.0%	36	100.0%
Friends	Every Time	3	2.2%	4	11.1%
	Often	21	15.4%	8	22.2%
	Sometimes	44	32.4%	17	47.2%
	Rarely	29	21.3%	4	11.1%
	Never	34	25.0%	1	2.8%
	Missing	5	3.7%	2	5.6%
	Total	136	100.0%	36	100.0%
Neighbors	Every Time	5	3.7%	3	8.3%
	Often	11	8.1%	6	16.7%
	Sometimes	27	19.9%	14	38.9%
	Rarely	39	28.7%	9	25.0%
	Never	48	35.3%	2	5.6%
	Missing	6	4.4%	2	5.6%
	Total	136	100.0%	36	100.0%
Colleagues	Every Time	3	2.2%	3	8.3%
	Often	13	9.6%	7	19.4%
	Sometimes	19	14.0%	12	33.3%
	Rarely	32	23.5%	3	8.3%
	Never	62	45.6%	8	22.2%
	Missing	7	5.1%	3	8.3%
	Total	136	100.0%	36	100.0%
Strangers	Every Time	2	1.5%	0	0.0%
	Often	6	4.4%	0	0.0%
	Sometimes	7	5.1%	6	16.7%
	Rarely	21	15.4%	11	30.6%
	Never	92	67.6%	17	47.2%
	Missing	8	5.9%	2	5.6%
	Total	136	100.0%	36	100.0%

Table S2. Paired t-test results comparing post-survey responses to the action question to lead results. P-values in red are significant at α = 0.05.

	Level A	Level B	Difference	Std Err Dif	Lower CL	Upper CL	p-Value
DO NOTHING	> 1 & < 10	> 10	2.03	0.93	0.12	3.94	0.04
	< 1	> 10	1.47	0.86	-0.28	3.23	0.10
	> 1 & < 10	< 1	0.56	0.63	-0.74	1.85	0.39
FILTER	< 1	> 1 & < 10	0.67	0.52	-0.39	1.72	0.21
	> 10	> 1 & < 10	0.50	0.80	-1.13	2.13	0.54
	< 1	> 10	0.17	0.77	-1.40	1.73	0.83
MOVE	> 10	< 1	0.71	0.51	-0.33	1.74	0.17
	> 10	> 1 & < 10	0.45	0.53	-0.62	1.53	0.39
	> 1 & < 10	< 1	0.25	0.31	-0.39	0.89	0.43
RESEARCH	> 10	< 1	1.47	0.71	0.02	2.92	0.05
	> 10	> 1 & < 10	1.43	0.75	-1.00	2.96	0.07
	> 1 & < 10	< 1	0.04	0.49	-0.96	1.04	0.94
TELL NEIGHBORS	> 10	< 1	1.03	0.77	-0.55	2.60	0.19
	> 10	> 1 & < 10	0.75	0.81	-0.91	2.41	0.36
	> 1 & < 10	< 1	0.28	0.53	-0.81	1.37	0.61
FLUSH	> 10	> 1 & < 10	0.92	0.76	-0.64	2.48	0.24
	> 10	< 1	0.75	0.73	-0.75	2.25	0.32
	< 1	> 1 & < 10	0.17	0.50	-0.85	1.18	0.74

Table S3. Significantly different pairs resulting from t-tests comparing pre-survey responses of various levels of demographics to the action question (α = 0.05).

Category	Question	Level A	Level B	N - Level A	N - Level B	Difference	Std Err Dif	p-Value
AGE	Tell Neighbors	> 44 & < 55	> 34 & < 45	19	18	0.92	0.38	0.02
AGE	Flush	> 54 & < 65	65+	22	41	0.41	0.20	0.04
EDUCATION	Do Nothing	PH.D. AND HIGHER	SOME HIGH SCHOOL, NO DIPLOMA	4	3	1.50	0.71	0.04
EDUCATION	Do Nothing	PH.D. AND HIGHER	TRADE/TECHNICAL/VOCATIONAL TRAINING	4	3	1.50	0.71	0.04
EDUCATION	Do Nothing	PH.D. AND HIGHER	ASSOCIATE DEGREE	4	22	1.41	0.51	0.01
EDUCATION	Do Nothing	PH.D. AND HIGHER	MASTER'S DEGREE	4	16	1.25	0.52	0.02
EDUCATION	Do Nothing	HIGH SCHOOL GRADUATE, DIPLOMA OR EQUIVALENT (e.g. GED)	ASSOCIATE DEGREE	23	22	0.60	0.28	0.03
EDUCATION	Move	SOME COLLEGE CREDIT, NO DEGREE	BACHELOR'S DEGREE	27	29	0.69	0.29	0.02
EDUCATION	Research	TRADE/TECHNICAL/VOCATIONAL TRAINING	PH.D. AND HIGHER	3	4	2.50	0.88	0.01

EDUCATION	Research	TRADE/TECHNICAL/VOCATIONAL TRAINING	SOME HIGH SCHOOL, NO DIPLOMA	3	3	2.33	0.94	0.01
EDUCATION	Research	MASTER'S DEGREE	PH.D. AND HIGHER	15	4	1.50	0.65	0.02
EDUCATION	Research	BACHELOR'S DEGREE	PH.D. AND HIGHER	30	4	1.47	0.62	0.02
EDUCATION	Research	ASSOCIATE DEGREE	PH.D. AND HIGHER	24	4	1.38	0.62	0.03
EDUCATION	Research	HIGH SCHOOL GRADUATE, DIPLOMA OR EQUIVALENT (e.g. GED)	PH.D. AND HIGHER	23	4	1.37	0.63	0.03
EDUCATION	Tell Neighbors	PH.D. AND HIGHER	SOME HIGH SCHOOL, NO DIPLOMA	4	3	2.42	0.92	0.01
EDUCATION	Tell Neighbors	ASSOCIATE DEGREE	SOME HIGH SCHOOL, NO DIPLOMA	23	3	2.06	0.74	0.01
EDUCATION	Tell Neighbors	MASTER'S DEGREE	SOME HIGH SCHOOL, NO DIPLOMA	16	3	1.85	0.76	0.02
EDUCATION	Tell Neighbors	SOME COLLEGE CREDIT, NO DEGREE	SOME HIGH SCHOOL, NO DIPLOMA	27	3	1.78	0.73	0.02
EDUCATION	Tell Neighbors	BACHELOR'S DEGREE	SOME HIGH SCHOOL, NO DIPLOMA	29	3	1.53	0.73	0.04
RENT/OWN	Filter	RENT	OTHER	35	2	1.34	0.61	0.03

RENT/OWN	Filter	OWN	OTHER	92	2	1.32	0.60	0.03

Table S4. Significantly different pairs resulting from t-tests comparing post-survey responses of various levels of demographics to the action question (α = 0.05).

Category	Question	Level A	Level B	N - Level A	N - Level B	Difference	Std Err Dif	p-Value
AGE	Do Nothing	65+	> 54 & < 65	7	4	2.79	0.85	0.00
AGE	Do Nothing	> 24 & < 35	> 54 & < 65	6	4	2.33	0.88	0.02
AGE	Filter	> 34 & < 45	65+	2	7	2.64	1.00	0.01
AGE	Filter	> 34 & < 45	> 24 & < 35	2	6	2.17	1.01	0.04
AGE	Filter	> 54 & < 65	65+	5	7	1.54	0.73	0.05
AGE	Research	> 54 & < 65	65+	5	7	2.14	0.65	0.00
AGE	Research	> 24 & < 35	65+	6	7	1.48	0.62	0.03
AGE	Research	> 54 & < 65	> 44 & < 55	5	7	1.43	0.65	0.04
AGE	Flush	> 54 & < 65	65+	5	7	2.03	0.71	0.01
AGE	Flush	> 24 & < 35	65+	5	7	1.63	0.71	0.03
AGE	Flush	> 44 & < 55	65+	7	7	1.57	0.65	0.02
INCOME	Do Nothing	$50,000 to $74,999	$100,000 to $124,999	5	2	2.90	1.20	0.03
INCOME	Do Nothing	$50,000 to $74,999	$125,000 to $149,999	5	2	2.90	1.20	0.03
INCOME	Do Nothing	$50,000 to $74,999	$25,000 to $49,999	5	8	2.03	0.82	0.02
INCOME	Move	$125,000 to $149,999	$50,000 to $74,999	2	5	1.50	0.66	0.04
INCOME	Move	$25,000 to $49,999	$50,000 to $74,999	8	5	1.00	0.45	0.04
INCOME	Research	$125,000 to $149,999	$50,000 to $74,999	2	5	2.90	0.99	0.01
INCOME	Research	$25,000 to $49,999	$50,000 to $74,999	9	5	1.84	0.66	0.01
INCOME	Research	$75,000 to $99,999	$50,000 to $74,999	7	5	1.83	0.69	0.02
INCOME	Tell Neighbor	$125,000 to $149,999	$50,000 to $74,999	2	5	3.10	0.79	0.00
INCOME	Tell Neighbor	$125,000 to $149,999	$100,000 to $124,999	2	2	3.00	0.94	0.00
INCOME	Tell Neighbor	$25,000 to $49,999	$50,000 to $74,999	9	5	2.27	0.52	0.00
INCOME	Tell Neighbor	$25,000 to $49,999	$100,000 to $124,999	9	2	2.17	0.73	0.01
INCOME	Tell Neighbor	$75,000 to $99,999	$50,000 to $74,999	7	5	1.89	0.55	0.00

INCOME	Tell Neighbor	$75,000 to $99,999	$100,000 to $124,999	7	2	1.79	0.75	0.03
INCOME	Flush	$100,000 to $124,999	$50,000 to $74,999	2	5	2.20	0.80	0.01
INCOME	Flush	$75,000 to $99,999	$50,000 to $74,999	7	5	1.77	0.56	0.00
INCOME	Flush	$125,000 to $149,999	$50,000 to $74,999	2	5	1.70	0.80	0.05
INCOME	Flush	$25,000 to $49,999	$50,000 to $74,999	10	5	1.50	0.52	0.01
EDUCATION	Do Nothing	PH.D. AND HIGHER	SOME COLLEGE CREDIT, NO DEGREE	3	2	3.67	1.39	0.01
EDUCATION	Do Nothing	BACHELOR’S DEGREE	SOME COLLEGE CREDIT, NO DEGREE	11	2	2.55	1.17	0.04
EDUCATION	Do Nothing	PH.D. AND HIGHER	ASSOCIATE DEGREE	3	4	2.42	1.16	0.05
EDUCATION	Filter	ASSOCIATE DEGREE	SOME COLLEGE CREDIT, NO DEGREE	4	2	2.75	1.17	0.03
EDUCATION	Filter	MASTER’S DEGREE	SOME COLLEGE CREDIT, NO DEGREE	3	2	2.67	1.24	0.04
EDUCATION	Filter	PH.D. AND HIGHER	SOME COLLEGE CREDIT, NO DEGREE	3	2	2.67	1.24	0.04
EDUCATION	Research	SOME COLLEGE CREDIT, NO DEGREE	BACHELOR’S DEGREE	2	12	2.58	0.88	0.01
EDUCATION	Research	PH.D. AND HIGHER	BACHELOR’S DEGREE	3	12	2.42	0.74	0.00
EDUCATION	Research	ASSOCIATE DEGREE	BACHELOR’S DEGREE	4	12	1.58	0.66	0.03

Table S5. Significantly different pairs resulting from t-tests comparing pre-survey responses of various levels of demographics to the social network question (α = 0.05).

Category	Question	Level A	Level B	N - Level A	N - Level B	Difference	Std Err Dif	p-value
EDUCATION	Family	BACHELOR'S DEGREE	HIGH SCHOOL GRADUATE, DIPLOMA OR EQUIVALENT (e.g. GED)	30	23	0.61	0.31	0.05
EDUCATION	Strangers	HIGH SCHOOL GRADUATE, DIPLOMA OR EQUIVALENT (e.g. GED)	MASTER'S DEGREE	23	15	0.67	0.30	0.03
INCOME	Friends	$50,000 to $74,999	$25,000 to $49,999	23	38	0.65	0.29	0.03
INCOME	Neighbors	$75,000 to $99,999	$25,000 to $49,999	16	40	0.80	0.33	0.02
INCOME	Colleagues	$0 to $24,999	$50,000 to $74,999	17	21	0.80	0.37	0.03
INCOME	Strangers	$0 to $24,999	$50,000 to $74,999	16	21	1.01	0.30	0.00
INCOME	Strangers	$0 to $24,999	$75,000 to $99,999	16	17	0.96	0.31	0.00
INCOME	Strangers	$0 to $24,999	$100,000 to $124,999	16	10	0.95	0.36	0.01
INCOME	Strangers	$0 to $24,999	$25,000 to $49,999	16	40	0.73	0.27	0.01
RENT/OWN	Colleagues	OTHER	OWN	2	94	2.69	0.76	0.00
RENT/OWN	Colleagues	OTHER	RENT	2	31	2.27	0.78	0.00
RENT/OWN	Strangers	OTHER	OWN	2	94	1.66	0.63	0.01
RENT/OWN	Strangers	RENT	OWN	30	94	0.49	0.18	0.01

Table S6. Significantly different pairs resulting from t-tests comparing post-survey responses of various levels of demographics to the social network question (α = 0.05).

Category	Question	Level A	Level B	N - Level A	N - Level B	Difference	Std Err Dif	p-value
INCOME	Family	$125,000 to $149,999	$100,000 to $124,999	2	2	2.00	0.89	0.03
INCOME	Family	$125,000 to $149,999	$25,000 to $49,999	2	10	1.70	0.69	0.02
INCOME	Family	$125,000 to $149,999	$75,000 to $99,999	2	8	1.50	0.70	0.04
INCOME	Friends	$125,000 to $149,999	$25,000 to $49,999	2	10	1.78	0.77	0.03
INCOME	Friends	$125,000 to $149,999	$75,000 to $99,999	2	8	1.75	0.78	0.04
INCOME	Colleagues	$125,000 to $149,999	$100,000 to $124,999	2	2	3.50	1.25	0.01
INCOME	Colleagues	$50,000 to $74,999	$100,000 to $124,999	5	2	2.40	1.04	0.03
EDUCATION	Strangers	SOME COLLEGE CREDIT, NO DEGREE	MASTER'S DEGREE	2	3	1.50	0.67	0.03