

CASE STUDIES

Planning and Executing Scientifically Sound Community Science in a Public-Facing Institution

Tiffany M. Nuessle*, Patricia A. McNamara† and Nicole L. Garneau*

The Genetics of Taste (GOT) Lab is an active research lab that studies human genetics at the Denver Museum of Nature & Science. In addition to the scientific research it conducts, the now-10-year-old lab is designed to accommodate two forms of public participation: community (citizen) science and crowdsourcing participants (human subjects). These crowdsourced participants provide a diverse sample for collecting taste and genetics data, and typically spend about 30 minutes in the Lab during enrollments. Alternatively, the community scientists invest a much more significant amount of time; they volunteer five hours weekly to train and then to assist in all aspects of the Lab's research, including data collection, DNA extraction, and manuscript preparation. This case study featuring the GOT Lab's pilot project, the Bitter Study, highlights three key topics: the ability to replicate established findings in the taste field using a community science model; the learning experiences of participants; and best practices and recommendations for other institutions that might wish to develop and implement a similar bi-level community-science model. Here we present data to illustrate that both the scientific and educational goals of community science and crowdsourcing are attainable within the same working space.

Keywords: community science; citizen science; museum; human-subject research; genetics

Introduction

When the Denver Museum of Nature & Science (hereafter known as the Museum) upgraded the aging health exhibit, *Hall of Life*, a decision was made to move away from the disease and wellness model of health exhibits to an interactive human biology–focused exhibit that engaged guests in a personally relevant way. The resulting health exhibit, *Expedition Health (EH)*, was guided by one main principle: It's not about THE human body; it's about YOUR human body. Each component of the exhibit was designed to provide a personal interactive element, an educational element, and the chance to see real specimens.

Genetics was to be one of the core topics featured in the new exhibit; however, the design team was aware that genetics is a subject of such complexity that it intimidates many people (Kassem, Girolami, and Sanoudou 2012). The team decided to hire a geneticist and create a public-facing research lab; this lab would not only study human genetics, but it would also be a space where the public could participate in the scientific process in as many ways as feasible. The guest experience would include each of the following educational and scientific elements: an invitation to enter the Lab and to interact with staff and

community scientists; an opportunity to learn about how genetics relates to their everyday lives; and the chance to contribute their own phenotypic and genotypic data to the research as part of a crowdsourcing effort.

Although genes are often understood to be responsible for identifiable characteristics like eye color and hair color, our preliminary audience research suggested that genes' contribution to more intangible qualities like sensory perception is often underestimated (McNamara 2012). Therefore, the Lab would need a sensory theme that the general public was both interested in and passionate about—a story about genetic variability that would resonate with people's everyday lives. People discuss taste preferences passionately, talking about the foods they love and the foods they hate with equal fervor (Rubenstein, 2009). Additionally, there are gaps in the published literature concerning how genetics affects taste variation and nutrition. For all of these reasons, the “genetics of taste” became the main theme of our public research, and the Genetics of Taste (GOT) Lab became a reality in 2008, generously supported by a Science Education Partnership (SEPA) award from the National Institutes of Health (Principal Investigator: Bridget Coughlin, award number 1R25RR025066). The Lab opened its doors to the public in April 2009, and in October 2009, the Museum hired a geneticist (co-author Garneau) to take the lead on both the Lab's scientific research and its educational programming.

* Denver Museum of Nature & Science, US

† Independent Evaluator, US

Corresponding author: Tiffany M. Nuessle (tiffany.nuessle@dmns.org)

The community-based Lab offers two levels of public participation. First, as human subjects (or crowdsourced participants), interested Museum guests give approximately thirty minutes of their time during their museum visit to contribute their genotypic and phenotypic data to the current research study, while learning how genetics relate to their lives through the sense of taste. The second, deeper level of participation gives community members the opportunity to volunteer as community (citizen) scientists who are trained in human-subject research protocols and ethics, to enroll participants in the study, to collect and prepare data, to extract DNA for sequencing, to analyze results, and to help disseminate the information in publications or presentations. Neither level of community participation in our Lab requires a science background, making it a feasible entry point for people without conventional credentials to become involved in scientific research. As far as we have been able to ascertain, this undertaking represents one of the first opportunities for public participation in modern human-genetics research, despite a long history of using community scientists in the fields of ecology and conservation (Miller-Rushing, Primack, and Bonney 2012), and in research departments typical of natural history museums (e.g., paleontology, zoology; Smithsonian Environmental Research Center 2016).

To ensure that our community science model was reliable and successful, the first study was designed to replicate the findings from previously published and well-established work on the topic of supertasting (Bartoshuk, Duffy, and Miller 1994). Scientifically, this would entail showing a relationship between the gene *TAS2R38* and the role of fungiform papillae (FP)—the bumps on the tongue that house the taste buds—in the detection of bitter taste (Bartoshuk, Duffy, and Miller 1994).

Of note, this topic also has a long history of public interest, which is a desirable factor for a public-facing lab. In 1931, Arthur L. Fox was pouring phenylthiocarbamide (PTC) into a bottle when his co-worker, C.R. Noller, complained that the PTC dust floating in the air left a bitter taste on his lips. Fox could detect nothing. They had several others try it and some, like Fox, tasted nothing, while others detected varying degrees of bitter. Thus, they discovered *taste blindness*, a term used to describe one's inability to perceive a taste that others can (Fox 1932). Those who are taste blind to PTC and its chemical cousin, 6-n-propylthiouracil (PROP), are called *non-tasters*, whereas those who can perceive the bitter taste are grouped into a category called *tasters*. Those who find it excruciatingly bitter were given the superlative name of *supertasters* (Bartoshuk, Duffy, and Miller 1994).

Research on taste blindness for PROP and PTC continued for decades, but it was not until 2003 that Kim and colleagues showed that the gene *TAS2R38* was responsible for the varying taste perception (Kim et al. 2003). The authors show that people who are able to perceive the bitterness have at least one copy of the dominant haplotype (PAV), whereas the majority of non-tasters (those taste blind to PTC and PROP) have two copies of the recessive haplotype (AVI). Therefore, because of the strong genetic component and the fact that PROP is well studied

by taste researchers, it made an ideal topic for the Lab's first study. It would provide a fun way to discuss genetics with guests, and concurrently offer our community lab the opportunity to replicate previous work in the field of taste completed in more conventional academic labs. We could offer our guests a new and engaging experience and demonstrate that our community science model is a viable way to collect sound human-subject data from crowdsourced human subjects.

It is important to note that this case study will focus on the Lab's second iteration of the Bitter Study as the data collected during the first iteration did not meet the methodological standards of the taste field; this was because of weaknesses in study design, not the community science model. This first iteration will be mentioned in the Discussion and Recommendations section to enable other institutions to learn from our mistakes. In addition, the Discussion and Recommendations section will draw upon all six of our research studies (conducted from 2009 to 2019) as we have refined our model and learned lessons during each subsequent study.

We developed this report with four main goals in mind:

- 1) to describe the background and details of the community-science enrollment model developed in our GOT Lab;
- 2) to demonstrate that the involvement of community scientists in a genetics research program does not affect research integrity;
- 3) to share findings from a third-party evaluation to document the model's contribution to engaging and educational guest experiences; and most importantly;
- 4) to provide recommendations to encourage and inform future community science work in the arena of human health and genetics, and specifically, the development of similar community-based research labs in informal-science venues.

Experimental Design

Our experimental design included four key elements: the GOT Lab's physical setting inside the Museum, the design of the community science model, the research study in which community scientists enrolled guests, and the evaluation methods used to examine each of these factors.

Genetics of Taste Lab setting

The GOT Lab is housed in the Museum's *Expedition Health* exhibit adjacent to a public-access wet lab, where both adults and children can spend 15 minutes or more engaged in hands-on science activities (e.g., viewing their own cheek cells under a microscope and extracting DNA from wheat germ). A glass wall separates the two spaces to enable Museum guests in the wet lab to see inside the working research lab and hopefully observe parallels between the activities they are doing and the procedures that lab staff are performing. During the Bitter Study, the Lab's staff and volunteers shared responsibility for maintenance of the adjacent wet lab's activity stations.

Design of community science model

Since its inception in 1897, the Museum has been home to a rich volunteer program, from docents in the public exhibits to volunteer community scientists working behind the scenes in the research and collection division (Cochran et al. 2013). Therefore, when the GOT Lab volunteer positions were posted, prospective community scientists applied through the Museum's existing volunteer program. Once they were accepted, they committed to one year of volunteering and were assigned a half-day shift (weekly or every other week) on the basis of their requested schedule and pending availability. The Lab's volunteer shifts aligned with the standard shifts across the museum: 3–4 volunteers per morning (0830–1300) and afternoon (1230–1700) shifts.

Each volunteer community scientist's first six shifts in the GOT Lab revolved around training. And as a human subject-based research lab, ethics was and remains the top training priority. Prior to learning any procedures related to the research study, staff and volunteer community scientists completed the online course "Protecting Human Research Participants" by the National Institutes of Health Office of Extramural Research (<https://phrp.nihtraining.com>, discontinued in 2018¹). Once community scientists finished their ethics training, they transitioned to in-depth training on the scientific context of the study. This training included detailed explanations on how the data would be collected and on how community scientists would become certified to work with human subjects, as well as a review of the study's resource binder. The binder reiterated what was covered during in-person training, and offered troubleshooting advice, sample scripts, and answers to anticipated questions from both community scientists and study participants. Because this research takes place within the Museum and requires guests to spend 30 minutes of their visit, the script for study enrollments was designed to be a fun and engaging experience rather than clinical and impersonal. Previous evaluation of Museum-goers shows that the typical Museum guest expects to both learn and be entertained during a visit (Cochran, Coughlin, and Garneau 2013). Therefore, in addition to the research standards, the Lab implemented guest experience standards. It is important to note that both the research standards and guest experience standards were given equal weight in training.

To become certified to enroll human subjects, community scientists were required to complete ethics training and six shifts during which they practiced data collection before conducting a mock enrollment. Their performance during that mock enrollment was observed by a staff member and assessed for competent achievement of both guest experience and research standards. For example, in the pilot Bitter Study, a research standard for the fungiform papillae (FP) image station required a close-up, in-focus photograph of the tongue. The guest experience standard dictated that community scientists explain what FP are, why the FP are being photographed, and the purpose of blue dye on the tongue (to aid in FP identification). Community scientists were to inform study participants that the blue dye can stain the tongue for up to two hours

(in case a guest wanted to opt out), and were to show the participant the picture of their FP after the image was taken. For the complete Bitter Study standards, please see Community Scientist Enrollment Certification Sheet in the Supplemental Files. After the mock enrollment, staff conferred and gave feedback to the community scientist, either letting them know that they had passed or identifying areas in either research or guest experience standards that required additional practice.

Materials and methods for crowdsourced data collection

Between November 2011 and August 2013, staff and community scientists enrolled 1,347 crowdsourced participants. Although the Lab welcomed volunteers 16 years and older to help conduct data analysis, because this was the first community science project overseen by the Western Institutional Review Board, the age of community scientists conducting enrollments was restricted to 18 years and above. Crowdsourced participants also were adults aged 18 years or older. They answered demographic questions, were trained to use the scale on which they would record the intensity of the PROP-infused wafer's bitterness (general Labeled Magnitude Scale; Bartoshuk, Duffy, and Miller 1994), and used a buccal swab to provide a DNA sample from cells on the inner cheek. Also, an image of the participant's tongue was taken to quantify the number of FP during analysis. Participants volunteered their time and gave written consent. For full scientific methodology, see Garneau et al. 2014 and Nuessle et al. 2015.

Comparison of data collected by staff and community scientists

To assess how the involvement of community scientists affected the quality of data collected, we reviewed both the amount of usable data and the quality of data collected by community scientists compared with staff scientists. Professional research staff ($n = 3$) enrolled 381 human subjects; community scientists ($n = 44$) conducted the remaining 966 enrollments. Data from these enrollments were considered usable if study participants applied the scale correctly to rate the taste intensity of the PROP wafer, and the image of their tongue was clear and taken at the correct angle. Using Fisher's exact test, we compared the proportion of usable data collected by staff and community scientists (SigmaPlot v14). To check for difference in the quality of data collected, taste-intensity scores and overall FP counts were compared using the Mann-Whitney test. Because the method we used to count FP requires two people to count and confer (Nuessle et al. 2015), a one-way ANOVA was run to compare the difference in individual FP counts when the pair counting was staff-staff, staff-community scientist, or community scientist-community scientist.

Evaluation of community scientists and their experience

To evaluate how well the model was achieving its experience goals for both community scientists and guests, the GOT Lab hired external evaluator, Patricia

McNamara (second author), to design and execute a third-party summative evaluation. Community scientists were given the opportunity to participate in small focus-group discussions (n = 12) and complete online surveys (n = 24). These forums gave the Lab's community scientists the opportunity to share feedback about their experiences in the Lab, offer suggestions for improvement, and share their understanding of the Bitter Study design, hypotheses, and anticipated findings. See Citizen Scientist Feedback Survey in the Supplemental Files for copies of surveys.

Evaluation of crowdsourced human subjects' guest experience

Guests who participated in the Bitter Study completed a self-administered survey just after they left the Lab (n = 90). Survey items included rating scales and open-ended questions that encouraged participants to share feedback about their overall experience and their understanding of study-related topics (e.g., the genetics of taste and the scientific process; see GOT Enrollment Survey in the Supplemental Files for a copy of this survey). A subset of the survey sample (n = 27) was also interviewed by telephone approximately four weeks later to gauge their recall of their Lab experience, details of the study itself, and their use of the take-home packet (see Phone-Call Follow-Up Survey in the Supplemental Files for a copy of the phone interview questionnaire). Study enrollees' understanding of study-relevant concepts was compared with that demonstrated by Museum guests who had completed a similar survey but had not enrolled in the Bitter Study (n = 147).

Results

The results focus on feedback from three key groups: the community scientists, the study participants as compared to a baseline sample, and the staff working in the lab.

Community scientist demographics

The community scientists who volunteered in the GOT Lab over the course of the Bitter Study ranged in age from 16 to 79 years; 67% were female. Twenty-four community scientists filled out an online survey. Of those community scientists, approximately 66% earned college or advanced degrees in science, and 60% worked (or currently work) in a science-related field. Ninety percent of the community scientists described themselves as being "very interested in science."

Evaluation of community scientist experience

Of the community scientists who completed the survey, 29% stated they were already familiar with genetics. Of the remaining respondents, 88% agreed that after volunteering in the lab they understood more about genetics; 24% strongly agreed with that statement (see **Figure 1**). Thirty-seven percent of these volunteers stated that they were already familiar with how scientists work. Forty-two percent of the remaining community-scientist respondents reported that their experience increased their understanding of scientific practices, including 8% who strongly agreed that that was the case (see **Figure 1**). As one community scientist stated during a focus group discussion, "I have learned that the process of science is dynamic and [researchers are] always searching for a better way to gather data or analyze it."

Approximately 33% of community scientists reported that they most enjoyed learning something new or doing something of value, while nearly 60% especially valued the opportunity to work in the Lab and to interact with the professional staff and their fellow volunteers.

Quality of scientific data collected by community scientists versus staff

Data collected by community scientists were no more likely to be excluded for the guest failing to understand

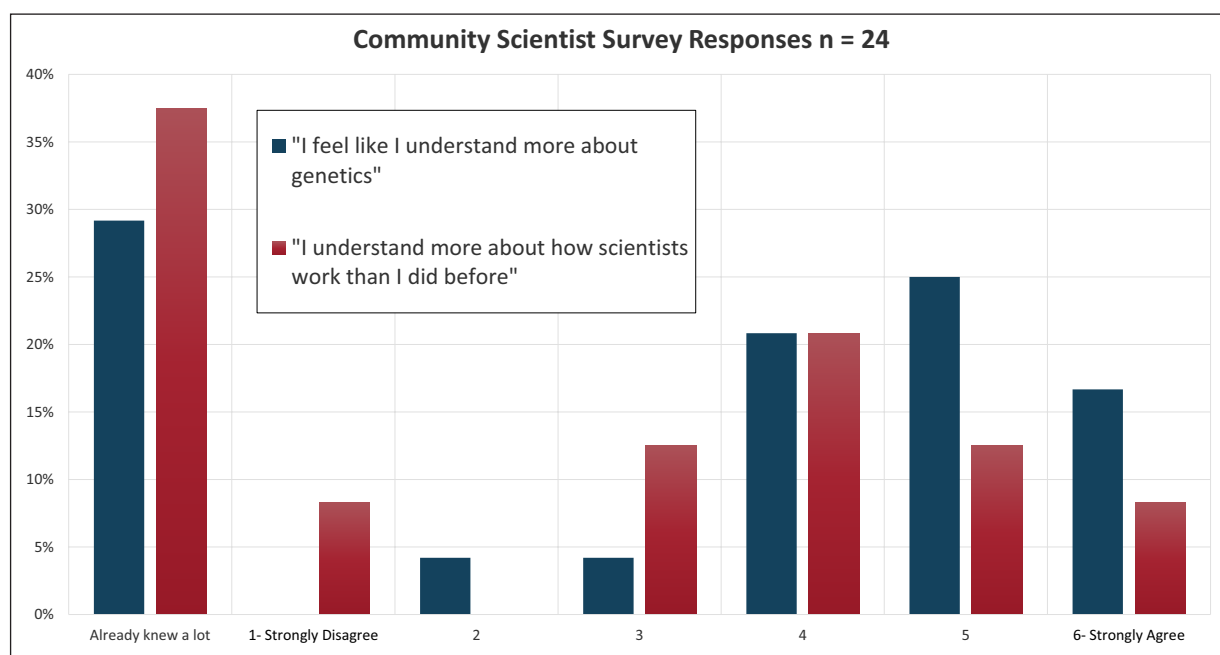


Figure 1: Community scientist survey responses.

the scaling tool than that collected by research staff ($p = 0.5305$), and the reported PROP intensity ratings also did not differ across these two groups ($p = 0.2826$; see **Figure 2**). Even though photo images taken by community scientists were excluded approximately twice as often as those taken by staff ($p < 0.001$; see **Figure 3**), acceptable images collected by staff and by community scientists yielded no difference in FP counts ($p = 0.2961$). Individual FP count variation between scorers showed no significant differences when images were scored by two staff members, one staff member and one community scientist, or two community scientists ($p = 0.746$).

Evaluation of the participant experience

Ninety percent of surveyed participants agreed with the statement “I feel like I participated in a real scientific study” (see **Figure 4**). Those who disagreed sometimes explained that even though the study seemed profes-

sional, it was too fun and laid back to feel *real*. The majority of interviewees described having positive interactions with the researchers (whether community scientists or staff); 86% of survey respondents agreed strongly with the statement “I really enjoyed myself,” and 94% agreed strongly with the statement “I felt very comfortable” (see **Figure 4**). These guests also agreed with both the statements, “I learned many new things about myself” and “I understand much more about genetics” (see **Figure 4**). When attempting to explain how their DNA affects how things taste to them, approximately 45% of study enrollees hypothesized that DNA *does* affect how things taste, compared with 30% of Museum guests completing the baseline survey. Approximately 75% of the survey respondents predicted that they would encourage family and friends to participate, and indeed, all of the phone interviewees stated that they had discussed their study experience with someone else.

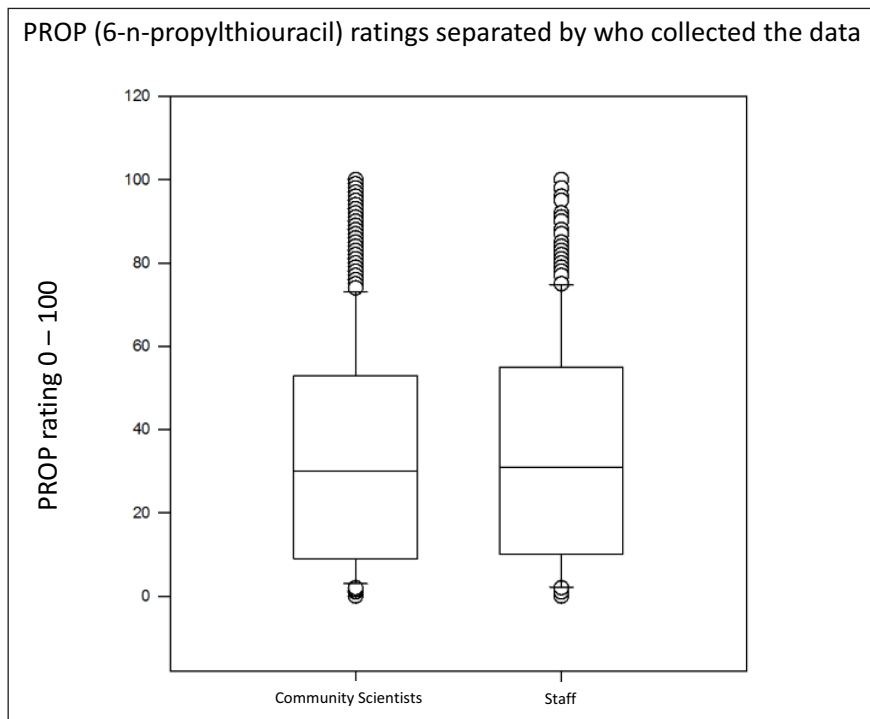


Figure 2: PROP (6-n-propylthiouracil) ratings separated by who collected the data; $p = 0.2826$.

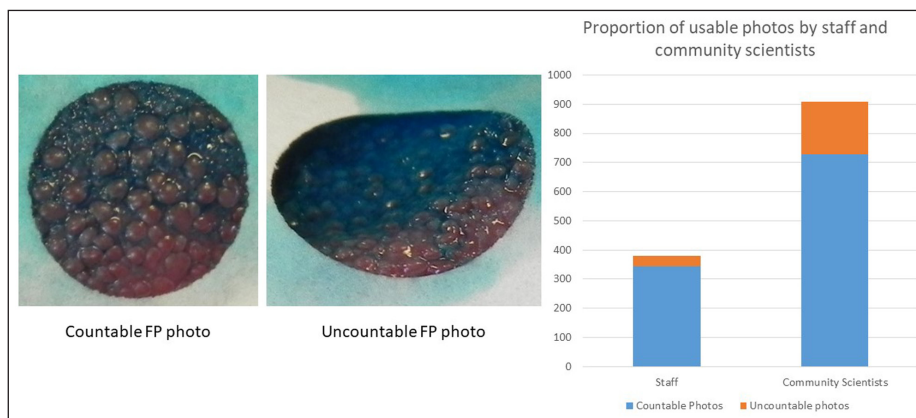


Figure 3: Proportion of usable and unusable photos collected by staff and community scientists; $p < 0.001$. FP, fungiform papillae.

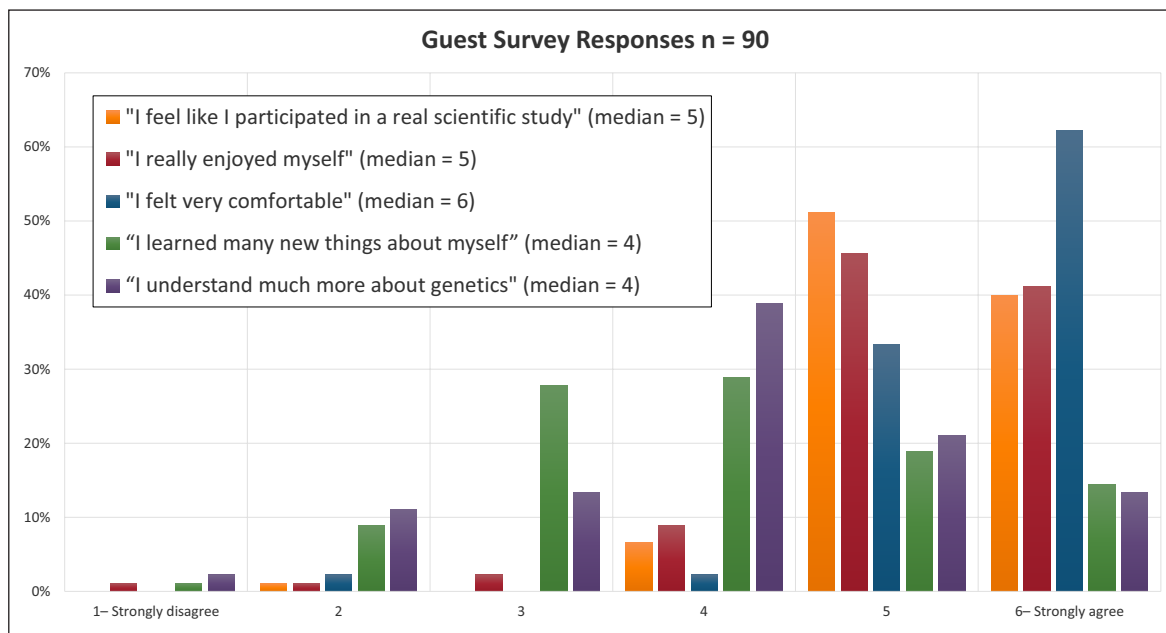


Figure 4: Guest survey responses.

Several of the enrollees' survey responses showed that even though their participation increased their *awareness* of the relationship between one's genetic profile and taste, their *understanding* reflected just partial recall of points made in the enrollment script. For example, instead of relating the taste of bitterness to one of three variations within a single gene, one participant explained, "if you have the three taster genes then you can taste the bitterness" (McNamara 2012b: 15). Similarly, participation in the study did not increase enrollee's understanding of "what it means to study something scientifically."

Feedback from Lab staff – Challenges of the staffing model

When the Lab originally opened, there were four part-time staff members. The budget restricted staff hours, and as a consequence, staff schedules did not overlap on a consistent basis. Surveys and interviews with both staff and community scientists highlighted the challenges of this staffing model. During interviews conducted by the external survey evaluation (or in survey responses), almost all Lab staff members said they were surprised by how long things took and how challenging the training and retraining process was when procedures changed. Many Lab staff indicated they were not used to working with human subjects in a museum setting and found it difficult to equally balance quality guest experience with sound research practice. The evaluation documented that different community scientist shifts sometimes received slightly different directions and were given different priorities. Moreover, in the initial pilot study, the staffing model was constantly evolving—a frustrating reality highlighted in the staff interviews. These professional challenges undoubtedly contributed to high staff turnover; one staff member who had given notice just prior to the evaluation described being the fourth team member to leave in that year alone.

Discussion and Recommendations

Importantly for a community-based lab, our participants' genetic profiles and demographics aligned with those previously reported by established academic labs (Hayes et al. 2008; Kim et al. 2003). Also, our FP findings were consistent with those reported by two other large-scale studies published at about the same time (Feeney and Hayes 2014; Fischer et al. 2013). This gave us confidence that our analysis was both accurate and valuable to the greater taste-science field.

Based on the data presented here, both on the ability of the community scientists to conduct human-genomics research and the qualitative feedback provided by staff and participants, we are confident that museums and other public-facing institutions can use a similar model to involve community scientists in the collection of sound human-subject genetics data. The following discussion offers 10 recommendations for designing and implementing an effective community science program that benefits both the scientific endeavor and the volunteers and museum guests involved in that research. These recommendations will not be limited solely to what we learned during the pilot study. The Lab has just celebrated its tenth anniversary and completed its sixth research study. To provide the best advice we can, we will draw upon the full extent of our experience.

Recommendation 1: Make sure that research staff appreciate that things work differently in a community-based lab. When hiring staff to work in a community lab, it is important to highlight the differences between both the speed of work and the order of priorities in an academic institution versus in a museum setting. As it was originally implemented (and noted above), the initial Lab staffing model led to confusion and inconsistency in the way the Lab ran from day to day. One of the staff's greatest challenges was achieving the appropriate balance among the Lab's competing priorities

(e.g., training community scientists, collecting data from guests, analyzing the data, and restocking and staffing the wet-lab activity area that shares a glass wall with the Lab). Staff often viewed the research they were hired to do as their primary job, and they felt they could not fulfill their research-related responsibilities because they were compelled to handle the urgent needs of the adjacent wet lab. These frustrations led to high turnover and the subsequent re-evaluation of staffing priorities.

In late 2011, a new staffing model went into effect to address these issues. Part-time non-research staff members were hired to handle the day-to-day demands of the exhibit's public wet lab so that Lab researchers could focus on training, data collection, and analysis. In addition to the project's principal investigator, three overlapping research staff members were regularly assigned to the Lab (including only one from the original staff). The Lab staff were specifically recruited to reflect both the Lab's scientific and educational missions—one staff member had solely a scientific background, one had both a science and education background, and the third had a background in education and communication. Divvying up the primary responsibilities ensured that the Lab could achieve its three key goals: maintain scientific integrity, offer proper training, and provide an engaging guest experience. This new team worked together closely to make sure that none of these goals were achieved at the expense of another.

Recommendation 2: Understand and value your community scientists' motivation, expertise, and time. People elect to volunteer in a museum like ours for a variety of reasons (e.g., to pursue professional/educational goals, to socialize with others, to contribute to the advancement of science, to share their love of science with museum guests, and to satisfy other personal goals). As is the case for any setting, it is important to understand what motivates individual volunteers and to try to make their experience satisfying and worthwhile. For example, one community scientist in our Lab planned to apply for graduate school and knew she would stand out if she was co-author on a paper. She proposed to us that she look at the rare haplotypes of the *TAS2R38* gene that were prevalent in our unusually large data set. She partnered with staff to assess how those haplotypes detect PROP (Boxer and Garneau 2015), benefitting the scientific field, the Museum, and her own career. Some community scientists may be interested only in conducting the analysis while others might be much more interested in collecting data from participants. For the Bitter Study, we required that all of the community scientists participate in both aspects. However, both informal feedback and evaluation findings confirmed that our community scientists derived different kinds of satisfaction from their work in the Lab. Some especially enjoyed the behind-the-scenes data processing but were less confident about interacting with the public during enrollments (and vice versa). It might be better to assess volunteer interests at the outset and structure different paths matched to those known interests (e.g., an enrollment track, an analysis track, and a publishing track). Obviously, such a system should be

flexible enough to ensure that as many volunteers as possible have a personally satisfying experience.

In addition, many people are interested in volunteering, but might not be able to commit to the schedule that Lab staff would prefer. Many of our volunteers could not consistently attend evening trainings or additional events because of other commitments, and they often missed shifts due to vacation, school workload, or inclement weather. During the Bitter Study, we required a weekly commitment and participation in all Lab activities. But more recently, we have opened up more flexible avenues of participation. For example, we have introduced a program that encourages teens to volunteer in the Lab on their high school breaks, committing to twenty shifts in a year. Although this prevents them from collecting data, they have been invaluable in conducting analysis and so have been able to contribute to the science when it works with their schedule. This program has been so successful that the Museum's volunteer department has now instituted a similar Museum-wide opportunity.

If study designers and staff maintain an awareness of community scientists' interests and backgrounds, they better position themselves to see new ways to involve community scientists in the study. Our volunteers come from a wide variety of professions and backgrounds and represent a fantastic resource that staff could draw on more effectively. Community scientists with education or communication backgrounds may be skilled in developing simple, age-appropriate explanations of complex topics and can contribute to the development of better enrollment scripts. For the Bitter Study, a few community scientists had backgrounds in photography and were able to offer ideas for capturing better images of participants' tongues. As the study progressed and new volunteers joined the Lab, we also relied more on experienced community scientists to supplement the more formal staff-led training. Novice volunteers were partnered with peers who excelled in particular areas and could share their skills and expertise. This proved to be especially useful during enrollments, when guests sometimes had very "wiggly" tongues or otherwise made it difficult to get a clear photograph of the fungiform papillae. In situations like that, the enroller could reach out to a peer trained in photography for help capturing a high-quality photo.

Finally, having many different people performing procedures quickly brought confusions to the surface so that they could be resolved. Procedures were formalized, streamlined, and described in language that could be easily understood no matter one's scientific background. Working with varied community scientists also helped the staff scientists strengthen their own public communication skills because the community scientists were willing to ask questions and let the staff know when their explanations weren't very clear.

Recommendation 3: Set realistic goals. It is important to temper the enthusiasm of both staff and community scientists to accomplish big things by clarifying the limitations of a museum environment. When initially planning the Bitter Study, we naively believed

that we could conduct enrollments non-stop, enrolling at least 14 people a day and approximately 5,000 per year. This did not take into account the realities of our community science model and its reliance on a volunteer workforce. As noted in Recommendation 2, we hadn't realized how volunteers' unpredictable absences would affect our schedules and hadn't anticipated the high turnover rate. As a result, we underestimated the time that would be required to train new volunteers and to familiarize returning volunteers with procedures that they might have forgotten during Lab absences. Add to this the staff turnover, and other Museum-related tasks competing for both staff and volunteer time, and it soon became obvious that we couldn't meet our original targets. Even so, we did enroll 1,347 museum guests during the course of the Bitter study, far exceeding the typical study enrollments for conventional taste labs.

Recommendation 4: Provide your community scientists with as much background information as possible. Community scientists are eager to share information with the public, but their fear of not being able to answer guests' questions often discourages them from conducting study enrollments (McNamara 2017). It is important to provide ample background information when preparing for a study. We failed to do this during one of our studies and we saw the number of enrollments decrease significantly. During informal conversations with volunteers, we learned that their nervousness about being able to answer guests' questions led them to avoid doing enrollments, which in turn led them to feel even more unprepared when the next opportunity arose. By contrast, in our subsequent Science of Sour Study (comparing taste sensitivity to five different types of acids), we provided volunteers with an overview of each acid (including the foods and beverages in which each is found), with an explanation of the pH scale, with results from previous studies, and with the location of gaps in the literature. This additional information helped community scientists engage with the script, highlight the points that interested them or seemed to interest Museum guests, and increase their confidence about answering guests' questions.

Recommendation 5: Don't cannibalize the science to increase public engagement. Even though the Lab opened in 2009, the Bitter Study data discussed in this paper was actually collected from 2011 to 2013. The study presented here is the second iteration of the Bitter Study. During its first iteration, we inadvertently designed a data collection method that did not follow established protocols in the taste-research field. We had focused so intently on creating an easy-to-understand protocol for guests that we neglected to keep up with current standards in the taste field. Looking back, we can see how this happened—none of the project's original staff had a background in taste science. At an international meeting of taste and smell experts in 2011, we learned that our data collection methods were flawed and we were told by many colleagues at that meeting that our data would never survive peer review. It is important to note that this was a study design flaw rather than a weakness in our community-science model. The community scientists collected the data exactly as they were trained to do. There was an unseen

benefit to this lesson. Our detailed examination of every data-collection method employed in the taste field led us to develop the Denver Papillae Protocol (DPP) to count FP consistently (Nuessle et al. 2015); DPP is now used in many academic taste labs (Cattaneo et al. 2019; Reynolds et al. 2017; Spinelli et al. 2017), and researchers attempting to automate FP counts have used DPP to check the accuracy of their proposed methods (Eldeghaidy et al. 2018; Piochi et al. 2017). The taste-research community ultimately benefited from our experience, even if our original data didn't advance the field's understanding of the genetics of taste.

Recommendation 6: Quality control for both learning and scientific goals. Quality control starts with staff providing clear expectations so that both staff and community scientists can understand the achievement standard and staff can be confident that enrollments will be executed consistently, whether by staff or by community scientists. During the first iteration of the Bitter Study, staff decided that no script would be provided so that the enrollment experience could be tailored to guests' interests. Unfortunately, this decision inadvertently compromised the scientific study. Community scientists and guests would often have a great conversation and only after enrollees left would community scientists realize that they had forgotten to complete a key station (e.g., DNA collection or photographing the participant's tongue), leading to several incomplete data sets. As we prepared to launch the second iteration of the Bitter Study, we created the previously mentioned resource binder that provided community scientists with sample scripts, copies of the research and guest-experience standards, and troubleshooting techniques for each station. The certification process for the new enrollment protocol included mistakes made by the mock enrollee that required the community scientist to demonstrate their ability to troubleshoot such situations or to explain things in multiple ways. When procedures change, it is important to communicate those changes in a variety of ways (in person, via email, and during formal trainings if possible). Formalizing how changes are communicated helps community scientists who are not in the lab regularly remember that a change took place and what it is. When we have introduced changes more organically, a high percentage of community scientists could not recall the changes or remember to implement them. It is also important to reiterate the changes frequently, reminding community scientists at the beginning of every shift about the changes that have recently taken place. Expect to repeat yourself on every shift for six to eight weeks.

Finally, mistakes can and will happen. It is important to have multiple levels of quality control and to plan for how errors will be addressed. We developed a quality-control checklist (see Quality Control Template in the Supplemental Files), and all community scientists and staff were trained on how to use it after enrollments. As often as possible, the quality-control checklist was completed on the same shift as the enrollment so that any errors could be pointed out to the volunteer who conducted the enrollment. Staff should regularly monitor quality-control data. Errors identified on the checklist may need to be addressed in different ways, depending on the nature of the mistake.

In our Lab, we sometimes need to discard data (e.g., if an ID sticker was not attached to the DNA sample). More serious errors have to be addressed immediately so that the review board can be contacted (e.g., if someone who had a pacemaker was put on the body analyzer). Regular monitoring by staff will also ensure that re-training is conducted when necessary. As much as possible, it is important to build accountability among the community scientists and to partner them during enrollments to reduce the likelihood of accidental errors.

Recommendation 7: Challenge all of your assumptions. When the new staffing model was introduced in 2011, we faced a steep learning curve and were thrilled whenever we could rely on previously established methods. It took us a long time to realize that in some cases we assumed there was a good reason behind a protocol or rule when that rationale either did not exist or was no longer valid. For example, because the Bitter Study was the first one undertaken in our new Lab, our institutional review board (IRB) was not comfortable allowing anyone under 18 to conduct enrollments or participate in the study. When we moved on to our next study (the Fatty Acid Study) and were working with a new IRB, we were able to enroll guests as young as 8 (with an accompanying legal guardian). Allowing parents and children to enroll together made this engaging experience more accessible and meaningful for families, an important Museum audience. We had also mistakenly assumed that there was a legal requirement that data collectors interacting with human subjects must be 18 or older and so never considered asking our IRB about lowering that age limit for participating community scientists. For our first five research studies, we did not allow volunteers under 18 to conduct enrollments, until a colleague remarked that she has teenagers collect human-subject data in her research project with full approval of her IRB. We immediately proposed to our IRB that we lower our enroller age limit to 16, which was approved. It was an opportunity we unknowingly denied several volunteers because we didn't even think to ask.

Recommendation 8: Expect additional publishing hurdles. As the inaugural team to carry out this kind of study in a museum-based, community-science Lab, we encountered significant hurdles when trying to publish our findings in professional journals. Many professional scientists are skeptical about the quality of data collected and analyzed by lay people in a nontraditional venue (Golumbic et al. 2017; Irwin 2018). For example, in the process of peer review for Garneau et al. (2015), we faced additional scrutiny because we used community scientists to quantify FP. Even though we followed a much more rigorous procedure than that typically used by taste-science practitioners (Miller and Reedy 1990), and we improved our reliability by requiring that two scorers (rather than one) reach agreement on an FP count for each photo (e.g., Delwiche et al. 2001), we still had to address reviewer concerns about the reliability of the community scientists' FP counts. Reviewers required us to re-examine FP counts done by staff member–community scientist pairs to document the extent to which their individual FP counts varied, and to include those statistics in our paper before it

was accepted for publication. Professional scientists who have used a more subjective method (and whose teams don't include community scientists) typically just cite Miller and Reedy (1990) and present their results without additional comment.

Recommendation 9: Don't pat yourself on the back just yet... the value of objective evaluation. It is important to provide multiple avenues for feedback. Informally, this can be accomplished during conversations on shift or after a challenging enrollment. This builds trust between the professional and community scientists and can set the stage for additional training and changes to methodology. Formally, we recommend working with an external evaluator to design and conduct evaluation to assess how well you meet your project's goals. Because their anonymity was assured, community scientists were able to provide feedback to staff that they may not have felt comfortable sharing more directly. While the overall feedback we received was positive, several community scientists critiqued various aspects of the program or offered specific suggestions for improving both their own experience and that of study enrollees. Inclusive evaluation, led by an external evaluator, likely encouraged less confident volunteers to anonymously offer critiques or recommend improvements to the more experienced professional staff.

Our project's evaluation component wasn't just about identifying strengths and weaknesses in our program. The evaluator also examined why the community scientists chose to volunteer in the Lab, what motivated their continuing commitment, and how we might ensure that we offered satisfying experiences for a variety of volunteers. We realized that key elements of our community science model (professional staff and community scientists working side by side, and teams of volunteers who worked regular shifts together) especially supported the development of these satisfying relationships among both staff and volunteers. Indeed, strong friendships were forged and have continued in the years since the Bitter Study was completed. By understanding how important these relationships are, we can deliberately create additional opportunities for staff and community scientists to get to know each other, to share insights and expertise, and to build the kinds of relationships that will in turn support volunteer satisfaction and retention.

Recommendation 10: Share your experience with other scientific institutions and colleagues. It is important to attend scientific conferences to present the community science model and to demonstrate to colleagues that community science can produce good data. In addition, such participation creates opportunities for Lab staff to share their expertise with a broader professional community and meet potential collaborators. Presentations shouldn't be limited to museum-related organizations; we have shared our model, findings, and experiences at a variety of professional conferences (e.g., Association for Chemoreception Sciences, Experimental Biology, and the Annual Meeting of the Society for the Study of Ingestive Behavior).

Moreover, as other research labs see the workforce available in a community-based lab and the integrity of the science being conducted, productive relationships are

built. For example, because we had a large core of trained volunteers, professional scientists in smaller labs have asked us to perform their DNA extractions (Burgess et al. 2018), and others have reached out to us for partnerships because of our ability to collect large data sets (Tucker et al. 2015; Baker et al. 2018). For three of our research studies, we have partnered with university researchers, leveraging their expertise with our volunteer pool and our ability to attract a wider cross-section of the public.

Conclusion

Implementing a community science model in a museum-based lab can seem daunting. However, one doesn't have to fully develop the project model before getting started. As we learned, it's better to start with clear goals, build on the experience of others, solicit feedback from your key audiences, evaluate progress as you go, and revise as necessary. As this in-depth case study testifies, we understand that thoughtful planning is imperative, but flexibility is paramount to the success of any museum-based lab that intends to include the public in scientific research. The initial set up of a new community science project inevitably involves considerable trial and error, especially when combining two levels of public participation as we did. In a few cases, the consequences of our mistakes were significant. For example, the scientific data collected in the Bitter Study's first iteration (December 2009 to August 2011) was determined to be of no scientific value—best methods in the taste research field were sacrificed to improve the guest experience of museum-goers enrolling as human subjects. Such mistakes prompted us to pause the project for three months while we changed protocols, submitted updates to our IRB, and retrained staff and community scientists. These trials allowed us to document that involving community scientists in data collection doesn't compromise data quality or study integrity (as long as the study uses current best practices and standardized methodology). We hope that the results presented here will increase confidence in the usefulness and value of research models that incorporate public participation, offer encouragement to those institutions envisioning the incorporation of a bi-level public participation model into their own research program, and allow those institutions and staff to avoid many of the pitfalls we encountered.

Data Accessibility Statement

To protect the privacy and confidentiality of the evaluation participants, non-aggregated data is not shared with outside parties. More detailed summaries of the evaluation studies are available from the Denver Museum of Nature & Science (contact Tiffany Nuessle at tiffany.nuessle@dmns.org). Researchers wishing access to edited data files should also contact Ms. Nuessle.

Note

¹ We now use the Collaborative Institutional Training Initiative (<https://www.citiprogram.org/>).

Additional Files

The additional files for this article can be found as follows:

- **Supplemental File 1.** Community Scientist Enrollment Certification Sheet. DOI: <https://doi.org/10.5334/cstp.263.s1>
- **Supplemental File 2.** Citizen Science Feedback Survey. DOI: <https://doi.org/10.5334/cstp.263.s2>
- **Supplemental File 3.** GOT Enrollment Survey. DOI: <https://doi.org/10.5334/cstp.263.s3>
- **Supplemental File 4.** Phone-Call Follow-Up Survey. DOI: <https://doi.org/10.5334/cstp.263.s4>
- **Supplemental File 5.** Quality Control Template. DOI: <https://doi.org/10.5334/cstp.263.s5>

Ethics and Consent

All procedures were approved by the Western Institutional Review Board (Study No. 1109386, Protocol No. 2009 1028), and the study complied with the Declaration of Helsinki for Medical Research involving Human Subjects.

Acknowledgements

The authors wish to thank all of our Bitter Study community scientists and staff who have waded through the challenges of setting up a project of this magnitude with smiles and insight. Listed in alphabetical order: Sue Ellen Alishouse, Elizabeth Aragon, Mike Archer, Sunanda Babu, Michael Bagley, Bob Barber, Alex Barnard, Dave Bauer, A'nette Bertrand, Bradley Bingham, Cristin Bohnen, Erica Borresen, Audrey Bower, Emma Boxer, Dan Brennand, Joanna Bruno, Claire Brunvand, Dick Burrows, Mehtap Canastar, Nicole Carey, Kaitlin Ching, Beth Cohen, Theresa Connelly, Jerome Cooper, Bridget Coughlin, Wendy Covert, Anthony Cross, Eric Crowell, Lindsey Daines, Christopher Del Re, Tracy Doan, Bonnie Doran, Patty Drever, Tim Dubois, Nicole Durant, Katelyn Engel, Jess Ern, Julian Fife, Jennifer Fisher, Diana Granados, Jonathan Grudis, Yedan Guo, Marcus Gutierrez, Lisa Harrington, Bob Hild, Brian Hostetler, Gretchen Hultman, Joyce Hutchens, Sally Huynh, Martin Indiatsi, Cynthia Jensen, Callan Jones, Chalise Jones, Rachel Jones, Wesley Jones, Matthew Joo, Leta Keane, Matthew Kern, Kyleigh King, Torry Knodell, Mary Krawczyk, Hannah Kubica, Flavia Kuta, Dalia Landman, Willem Leenhouts, Justin Li, Tanya Liakhova, Rachel Lindstrom, Ryan Loesch, Stephania Lukjan, Elisabeth Mansfield, Ashley Matthews, Mike Mauser, Erin McCool, Laurie McKee, Ashley Meredith, Dani Meyers, Stephanie Miller, Wendy Miracle, Anjelica Miranda, Becca Morgan, Kimberly Nesbitt, Susan Nicholson-Dykstra, Gina Olberding, Jericho Oviedo, Dushyant Parekh, Gabriel Parker, Peter Philpott, Jane Rech, Brian Reinhart, Tanya Reither, Willy Reusser, Emily Robitschek, Glenn Rodriguez, Paul Rosen, Hannah Schumann, Shana Schwartz, Adam Schwindt, Sara Shanahan, David Shander, Catherine Sheldon, Betsy Shepardson, John Simmons, Terri Simons, Vanessa Slack, Meghan Sloan, Katherine Smith, Stephanie St. George, Keely Sudhoff, Kizra Sullivan, Davida Terrell, Kathleen Tinsworth, Kelsey Thornburg, Joseph Tollas, Rudy Torres, Laureen Trainer, Julie Treece, Weston Truman, Sean Tseng, Kari Vagstad, Jo Van De Wege, Anna Vigran, Stephen Walden, Devin Walecka, Nancy Walsh, Terry Ann Watts, Mary Wehrwein, Jacklyn Widhalm, Tyler Wilson, Anastasia Wise, Diane Woltkamp, and Jennifer Zuniga.

We are also indebted to a cadre of evaluation assistants who surveyed nearly 400 museum guests during the Bitter Study summative evaluation: Anabel Adler, Robyn Baxendale, Cynthia Conner, Adey Dimalanta, Jennifer Fox, Nicoleta Gustafson, Ashley Jaramillo, Johanna Ladis, Patty Lindstrom, Mary Livaidais, Jen McCluskey, Meighan Meeker, Elizabeth Muñoz, Lania Ross, Lisa Roll, Lorie Savinar, Neil Sittler, Fred Staab, Vanessa Wagner, and Joanna Zelazko.

Funding Information

The Bitter Study was supported in part by a Science Education Partnership Award from the National Center for Research Resources, National Institutes of Health (award number 1R25RR025066). The Science of Sour, Savory & Sour Study, Genes & Grains, and the Comunidad Student Scientists high school program were supported in part by a Science Education Partnership Award from the National Institute of General Medical Sciences, National Institutes of Health (Award # R25OD021909).

Competing Interests

The authors have no competing interests to declare.

Author Contributions

TMN and NLG made substantial contributions to the conception and design of the work; PAM served as a core member of the project team, designing the evaluation studies, and supervising evaluative data collection and analysis. PAM compiled detailed evaluation reports from which findings shared in this paper were drawn. TMN led the analysis and interpretation of research data for the work. All authors contributed to drafting the work, and revising it critically for important intellectual content.

References

- Baker, AN, Miranda, AM, Garneau, NL and Hayes, JE.** 2018. Self-reported smoking status, Tas2r38 variants, and propylthiouracil phenotype: an exploratory crowdsourced cohort study. *Chemical Senses*, 43: 617–625. DOI: <https://doi.org/10.1093/chemse/bjy053>
- Bartoshuk, LM, Duffy, VB and Miller, IJ.** 1994. PTC/PROP tasting: anatomy, psychophysics, and sex effects. *Physiology & Behavior*, 56(6): 1165–1171. DOI: [https://doi.org/10.1016/0031-9384\(94\)90361-1](https://doi.org/10.1016/0031-9384(94)90361-1)
- Boxer, EE and Garneau, NL.** 2015. Rare haplotypes of the gene TAS2R38 confer bitter taste sensitivity in humans. *SpringerPlus*, 4: 505. DOI: <https://doi.org/10.1186/s40064-015-1277-z>
- Burgess, B,** et al. 2018. Effects of cd36 genotype on oral perception of oleic acid supplemented safflower oil emulsions in two ethnic groups: a preliminary study. *Journal of Food Science*. DOI: <https://doi.org/10.1111/1750-3841.14115>
- Cattaneo, C,** et al. 2019. New insights into the relationship between taste perception and oral microbiota composition. *Scientific Reports*, 2019 Mar 5; 9(1): 3549. DOI: <https://doi.org/10.1038/s41598-019-40374-3>
- Cochran, C,** et al. 2013. Education: informal science learning for the public. In: Krell, FT (ed.), (Number 4, December 31, 2013). *Denver Museum of Nature & Science Annals*, 105–136. Denver: Denver Museum of Nature & Science.
- Cochran, C, Coughlin, B and Garneau, N.** 2013. Health Sciences: a newcomer and a pioneer. In: Krell, FT (ed.), (Number 4, December 31, 2013). *Denver Museum of Nature & Science Annals*, 363–382. Denver: Denver Museum of Nature & Science.
- Delwiche, JF, Buletic, Z and Breslin, PA.** 2001. Relationship of papillae number to bitter intensity of quinine and PROP within and between individuals. *Physiology & Behavior*, 74: 329–337. DOI: [https://doi.org/10.1016/S0031-9384\(01\)00568-6](https://doi.org/10.1016/S0031-9384(01)00568-6)
- Eldeghaidy, S,** et al. 2018. An automated method to detect and quantify fungiform papillae in the human tongue: validation and relationship to phenotypical difference in taste perception. *Physiology & Behavior*, 184: 226–234. DOI: <https://doi.org/10.1016/j.physbeh.2017.12.003>
- Feeney, EL and Hayes, JE.** 2014. Exploring associations between taste perception, oral anatomy and polymorphisms in the carbonic anhydrase (gustin) gene CA6. *Physiology & Behavior*, 128: 148–154. DOI: <https://doi.org/10.1016/j.physbeh.2014.02.013>
- Fischer, ME,** et al. 2013. Factors related to fungiform papillae density: a Beaver Dam offspring study. *Chemical Senses*, 2013 Oct; 38(8): 669–677. DOI: <https://doi.org/10.1093/chemse/bjt033>
- Fox, AL.** 1932. The Relationship between chemical constitution and taste. *Proceedings of the National Academy of Sciences of the United States of America*, 18: 115–120. DOI: <https://doi.org/10.1073/pnas.18.1.115>
- Garneau, NL,** et al. 2014. Crowdsourcing taste research: genetic and phenotypic predictors of bitter taste perception as a model. *Frontiers in Integrative Neuroscience*, 8(33). DOI: <https://doi.org/10.3389/fnint.2014.00033>
- Golumbic, YN, Orr, D, Baram-Tsabari, A and Fishbain, B.** 2017. Between vision and reality: a study of scientists' views on citizen science. *Citizen Science: Theory and Practice*, 2(1): 6. DOI: <https://doi.org/10.5334/cstp.53>
- Hayes, JE, Bartoshuk, LM, Kidd, JR and Duffy, VB.** 2008. Supertasting and PROP bitterness depends on more than the TAS2R38 gene. *Chemical Senses*, 33: 255–265. DOI: <https://doi.org/10.1093/chemse/bjm084>
- Irwin, A.** 2018. No PhDs needed: how citizen science is transforming research. *Nature*, 562: 480–482. DOI: <https://doi.org/10.1038/d41586-018-07106-5>
- Kassem, HS, Girolami, F and Sanoudou, D.** 2012. Molecular genetics made simple. *Global Cardiology Science and Practice*, 2012(1): 6. Published online 2012 Jul 4. DOI: <https://doi.org/10.5339/gcsp.2012.6>
- Kim, UK, Jorgenson, E, Coon, H, Leppert, M, Risch, N and Drayna, D.** 2003. Positional cloning of the human quantitative trait locus underlying taste sensitivity to

- phenylthiocarbamide. *Science*, 299: 1221–1225. DOI: <https://doi.org/10.1126/science.1080190>
- McNamara, P.** 2012. Genetics of Taste: summative evaluation report. Unpublished manuscript, Denver Museum of Nature & Science.
- McNamara, P.** 2017. More than a taste of citizen science: Front-end evaluation report. Unpublished manuscript, Denver Museum of Nature & Science.
- Miller, IJ, Jr and Reedy, FE, Jr.** 1990. Quantification of fungiform papillae and taste pores in living human subjects. *Chemical Senses*, 15(3): 281–294. DOI: <https://doi.org/10.1093/chemse/15.3.281>
- Miller-Rushing, A, Primack, R and Bonney, R.** 2012. The history of public participation in ecological research. *Frontiers in Ecology and the Environment*, 10: 285–290. DOI: <https://doi.org/10.1890/110278>
- Nuessle, TM,** et al. 2015. Denver papillae protocol for objective analysis of fungiform papillae. *Journal of Visualized Experiments*. DOI: <https://doi.org/10.3791/52860>
- Piochi, M,** et al. 2017. Comparing manual counting to automated image analysis for the assessment of fungiform papillae density on human tongue. *Chemical Senses*, 42(7): 553–561. DOI: <https://doi.org/10.1093/chemse/bjx035>
- Reynolds, S, Burgess, ME, Hymowitz, N, Snyder, DJ and Lane, SJ.** 2017. Fungiform papilla number and olfactory threshold assessment in males with and without barth syndrome. *Chemical Perception*, 10(60). DOI: <https://doi.org/10.1007/s12078-017-9228-4>
- Rubenstein, S.** 2009. Across the land, people are fuming over an herb (no, not that one)-cilantro haters boo ‘fetid barb of green’; a prominent critic recants. *Wall Street Journal, Eastern edition*, [online]. 13 Feb 2009. Available at: <https://www.wsj.com/articles/SB123446387388578461> 13 Feb 2009: A.1. [Accessed 06 February 2019].
- Smithsonian Environmental Research Center.** 2016. *2016 citizen science updates*, [online]. Available at <https://serc.si.edu/2016citizenscienceupdate> [Accessed 6 July 2019].
- Spinelli, S,** et al. 2017. Investigating preferred coffee consumption contexts using open-ended questions. *Food Quality and Preference*, 61: 63–73. DOI: <https://doi.org/10.1016/j.foodqual.2017.05.003>
- Tucker, RM, Nuessle, TM, Garneau, NL, Smutzer, G and Mattes, RD.** 2015. No difference in perceived intensity of linoleic acid in the oral cavity between obese and nonobese individuals. *Chemical Senses*, 1–7. DOI: <https://doi.org/10.1093/chemse/bjv040>

How to cite this article: Nuessle, TM, McNamara, PA and Garneau, NL. 2020. Planning and Executing Scientifically Sound Community Science in a Public-Facing Institution. *Citizen Science: Theory and Practice*, 5(1): 9, pp.1–12. DOI: <https://doi.org/10.5334/cstp.263>

Submitted: 09 July 2019

Accepted: 27 November 2019

Published: 09 March 2020

Copyright: © 2020 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <https://creativecommons.org/licenses/by/4.0/>.



Citizen Science: Theory and Practice is a peer-reviewed open access journal published by Ubiq Press.

OPEN ACCESS