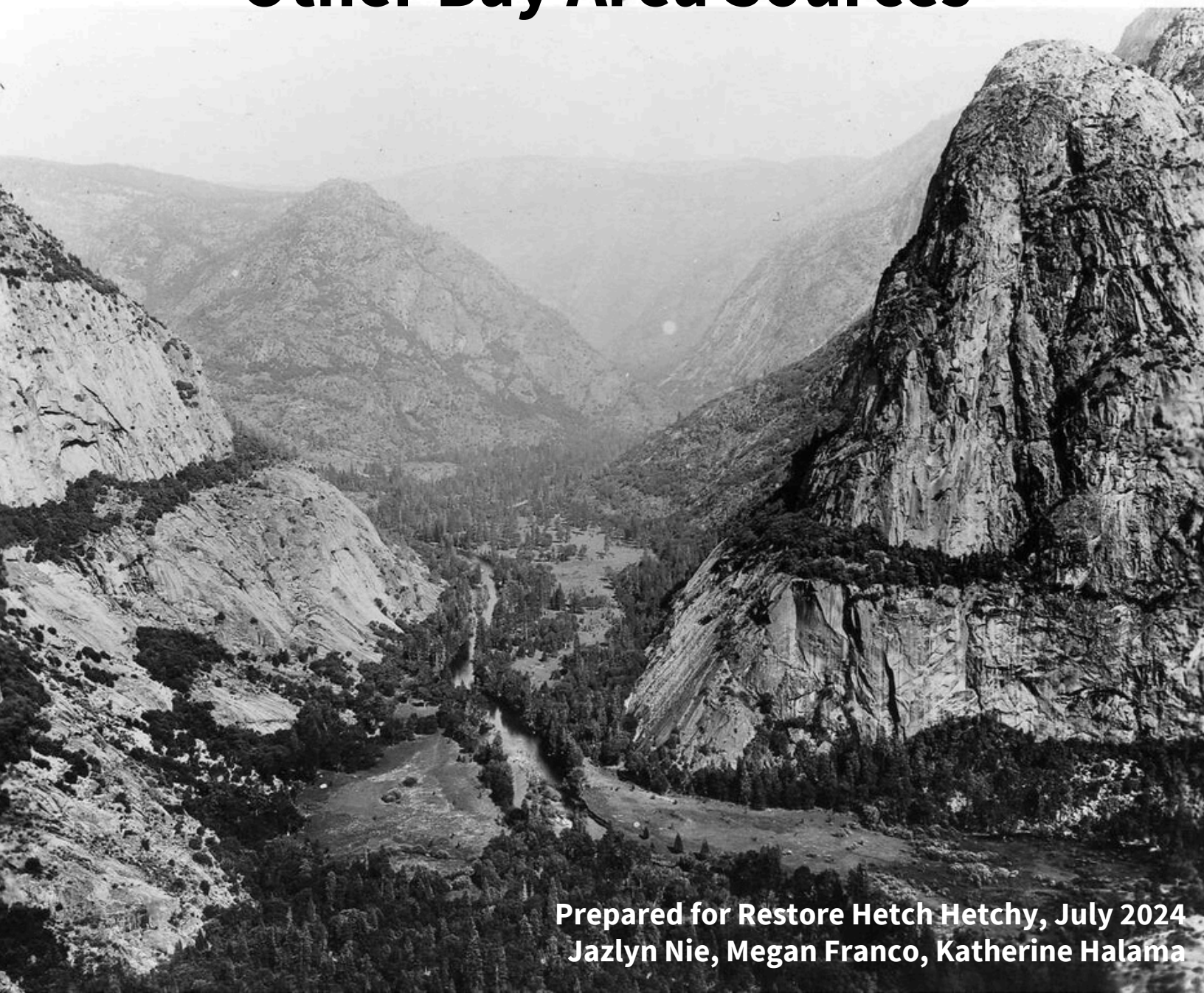
Masters of the Environment
UNIVERSITY OF COLORADO **BOULDER**

Water Taste Test

Comparing Hetch Hetchy Water to Other Bay Area Sources



Prepared for Restore Hetch Hetchy, July 2024
Jazlyn Nie, Megan Franco, Katherine Halama

Water Taste Test Report: Comparing San Francisco Regional Water System Supplies to Other Bay Area Sources

Prepared for Restore Hetch Hetchy, July 2024

Jazlyn Nie, Megan Franco, Katherine Halama

Table of Contents

Abstract

City leaders, media and others in San Francisco have historically claimed that their drinking water has superior taste compared to other water suppliers. San Francisco blends its Tuolumne water supplies with local runoff and groundwater; collective civic pride often uses the phrase "Hetch Hetchy water" as a nickname for all supplies provided by the San Francisco Regional Water System (SFRWS).

To question these claims, we conducted a double-blind water taste test to determine whether Bay Area residents can distinguish the difference between San Francisco's water sourced from Hetch Hetchy and other regional water sources. The test involved tasting three successive water samples, sourced from Marin Municipal Water District (Marin), East Bay Municipal Utility District (EBMUD), and SFRWS. Participants were asked to complete a questionnaire rating the water samples as favorite, second favorite, and least favorite, and whether they would accept all three samples as their daily drinking water. Initial results indicated no clear preference among the three samples although, by small margins, Marin was rated the favorite, EBMUD second and San Francisco last (See Figure 1). It is essential to point out that both more data and a review of the study design are needed before reaching firm conclusions about the results. In this trial run of the taste test, our participant pool was limited and our experience suggested refinements in methodology should be considered before additional testing.

Overall, our study provided Restore Hetch Hetchy with a repeatable methodology so they can collect more data to better understand if Bay Area residents can distinguish between local water sources. We also provided several recommendations for how Restore Hetch Hetchy can conduct future taste testing.

Background

The San Francisco Public Utilities Commission (SFPUC), who supplies the San Francisco Bay Area with water, boasts on its website that its water, most of which is diverted from Hetch Hetchy Reservoir, is "among the purest water in the world." Most of the city's water comes from the Tuolumne River in the Sierra Nevada and is stored in Hetch Hetchy, Cherry, Eleanor, and Don Pedro Reservoirs more than 100 miles from the city. Water from the Tuolumne

is diverted from Hetch Hetchy Reservoir in Yosemite National Park, which was dammed in the early 20th century after the passing of the controversial Raker Act of 1913.[1] Supplies diverted directly from Hetch Hetchy are exempt from state and federal filtration requirements due to its "exceptional quality." However, SFPUC does filter runoff from local watersheds. Additionally, supplies diverted from Hetch Hetchy are temporarily stored in local reservoirs and account for about 30% of total system deliveries.[2] SFPUC treats all its water with ultraviolet light, pH adjustment, fluoridation, and chloramination (chlorine and ammonia).

Many locals claim they can taste the difference between "Hetch Hetchy" water and other tap water. We aimed to question this assertion by developing a double-blind water taste test, to determine if there is a clear taste difference between SFRWS water and other tap water sources in the Bay Area, including water from EBMUD and Marin. We conducted a double-blind taste test so neither the test team nor participants knew which water sample correlated with each location, providing us with the most objectivity and minimizing participant and researcher bias.

These three utilities, like most utilities, do not deliver water provided from a lone source. Rather, supplies are typically blended from different water sources. The exact blend changes from time to time and can differ within a utility's service territory.[3]

- SFRWS – 85% of San Francisco's water comes from the Tuolumne River in the Sierra Nevada. Most of this water is delivered directly to customers, but some is stored in Bay Area reservoirs. Water stored in Bay Area reservoirs is both filtered and treated before being delivered. San Francisco has also added slight amounts of groundwater to make its supplies stretch further. In recent years, San Francisco has shut down all deliveries from

[1] "Water Quality," San Francisco Water Power Sewer, accessed June 5, 2024, https://sfpuc.org/accounts-services/water-quality.

[2] "Groundwater Supply Project."

[3] McDonald et al., "Water on an Urban Planet."

the Tuolumne (and Hetch Hetchy) for two months in the winter, so SFPUC can do critically important maintenance on its Mountain Tunnel. During this time San Francisco customers rely entirely on water from local watersheds and do not receive any "Hetch Hetchy" water.[4]

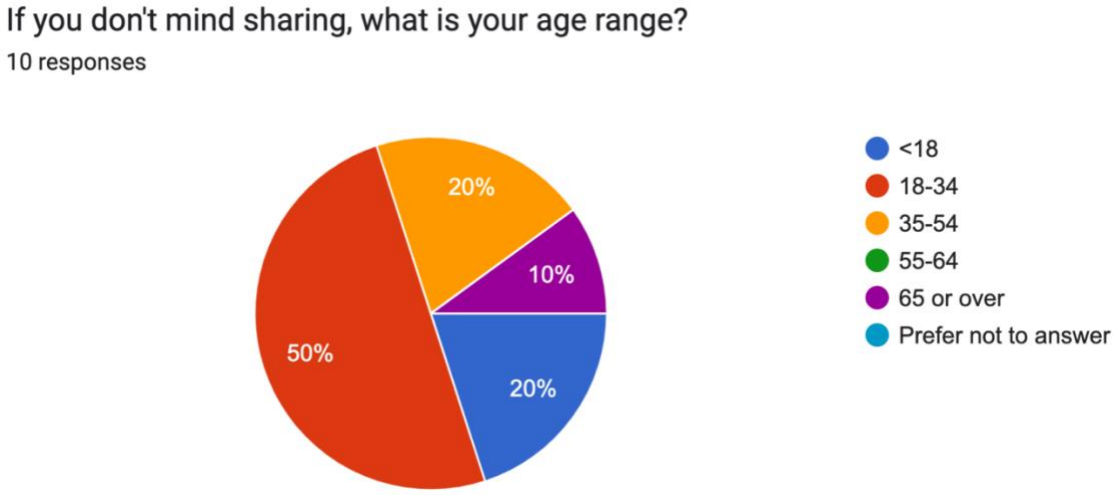
- EBMUD – East Bay cities including Berkley and Oakland rely on water imported from the Mokelumne River in the Sierra Nevada (near the Tuolumne River) for most of their supply. EBMUD also harvests runoff in local watersheds and, during dry years, brings in water from its Freeport project on the Sacramento River.[5]
- Marin - Marin gets no water from the Sierra's, and instead relies on rain to fill its reservoirs within the county. Other supplies are bought from Sonoma County Water Agency, who act as a water wholesaler, and local groundwater.[6]

Results

We recruited 22 participants in our first trial, which took place at the BART Rockridge Station in Oakland for three hours on June 14th, 2024. Spreck Rosekrans, Executive Director of Restore Hetch Hetchy, recruited an additional 6 participants after our initial trial, for a total of 28 participants and survey respondents.

[4] "Water Supply."
[5] "Water Supply: East Bay Municipal Utility District."
[6] "About Our Water System | Marin Water."

Figure 1: Participants' age distribution



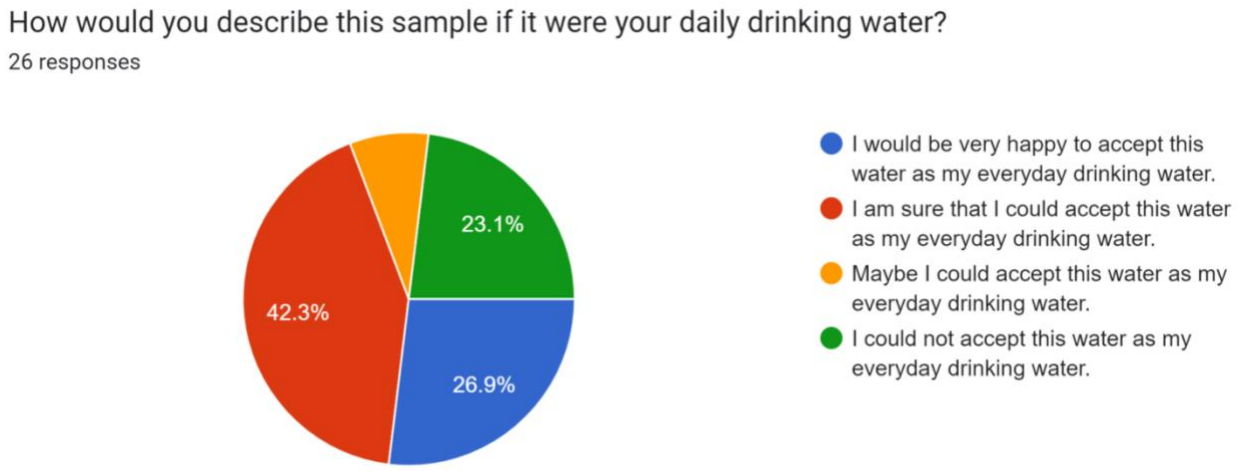
The survey question asking for a participant's age was not required (to see the survey questions, refer to the supplemental Methodology document), so we only received answers from 10 participants. As Figure 1 shows, 50% of respondents were between the ages of 18-34 and 10% were over the age of 65. The other 40% of respondents were between the ages of 35 and 64.

Participants could try each of the samples in any order. However, many (74.1%) of participants tried Sample A first, nearly all (81.5%) participants tried Sample B second, and many (77.8%) tried Sample C last.

For each sample, we asked participants to describe the sample using the following scale:

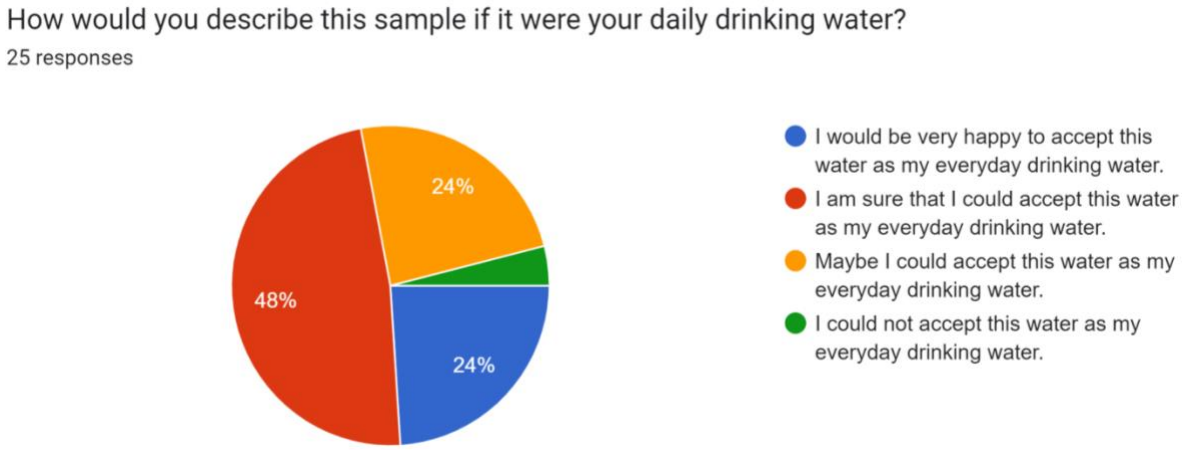
4	I would be happy to accept this water as my everyday drinking water.
3	I am sure that I could accept this water as my everyday drinking water.
2	I could accept this water as my everyday drinking water.
1	I could not accept this water as my everyday drinking water.

Figure 2: Participants' Descriptions of Sample A



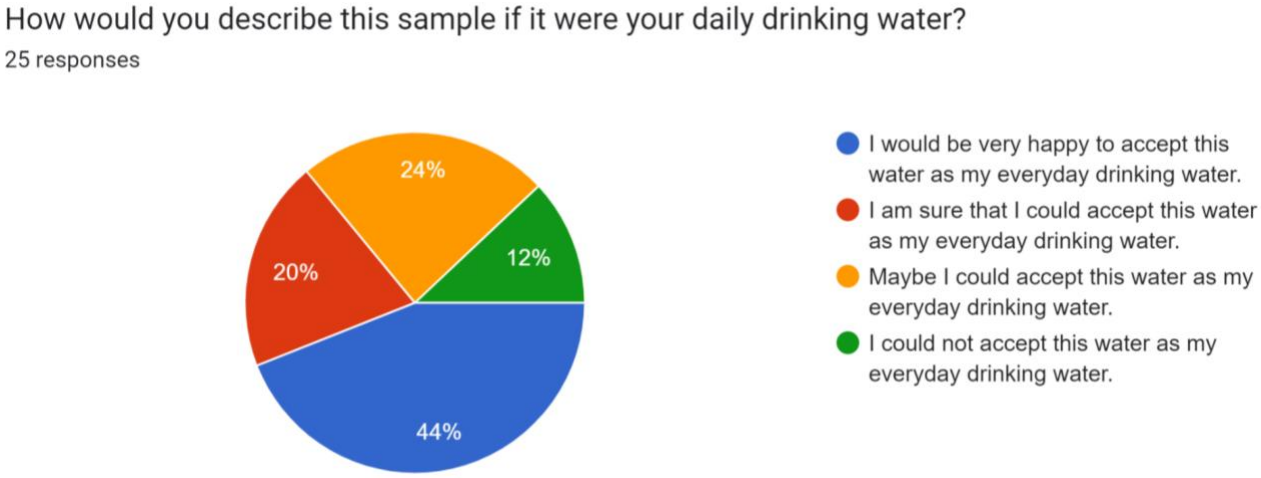
As Figure 2 shows, for Sample A, 26.9% of participants rated the sample a "4", or they "would be very happy to accept this water as [their] everyday drinking water," while 23.1% rated the sample A "1", or they "could not accept this water as [their] everyday drinking water."

Figure 3: Participants' Descriptions of Sample B



As Figure 3 shows, for Sample B, 24% of participants rated the sample a "4", or they "would be very happy to accept this water as [their] everyday drinking water," while 4% rated the sample a "1", or they "could not accept this water as [their] everyday drinking water."

Figure 4: Participants' descriptions of Sample C



As Figure 4 shows, for Sample C, 44% of participants rated the sample a "4", or they "would be very happy to accept this water as [their] everyday drinking water," while 12% rated the sample a "1", or they "could not accept this water as [their] everyday drinking water."

When asked to rate the three samples, Sample A was ranked as “Favorite” by 12 participants, Sample B was ranked as “Favorite” by 9, and Sample C was ranked as “Favorite” by 8. Sample B was the most common “Second Favorite,” with 13 participants choosing Sample B. Sample C was the most common “Least Favorite,” with 11 participants choosing Sample C.

While this question asked the participants to rank each sample with the intention of only one sample being “Favorite,” one being “Second Favorite,” and one being “Least Favorite,” a few participants said they could not distinguish a taste difference between samples and were unsure of how to rank the samples. In these instances, we permitted participants to rank more than one sample as the same. For example, one participant ranked all three samples as “Second Favorite,” since he could not distinguish any taste differences between the samples.

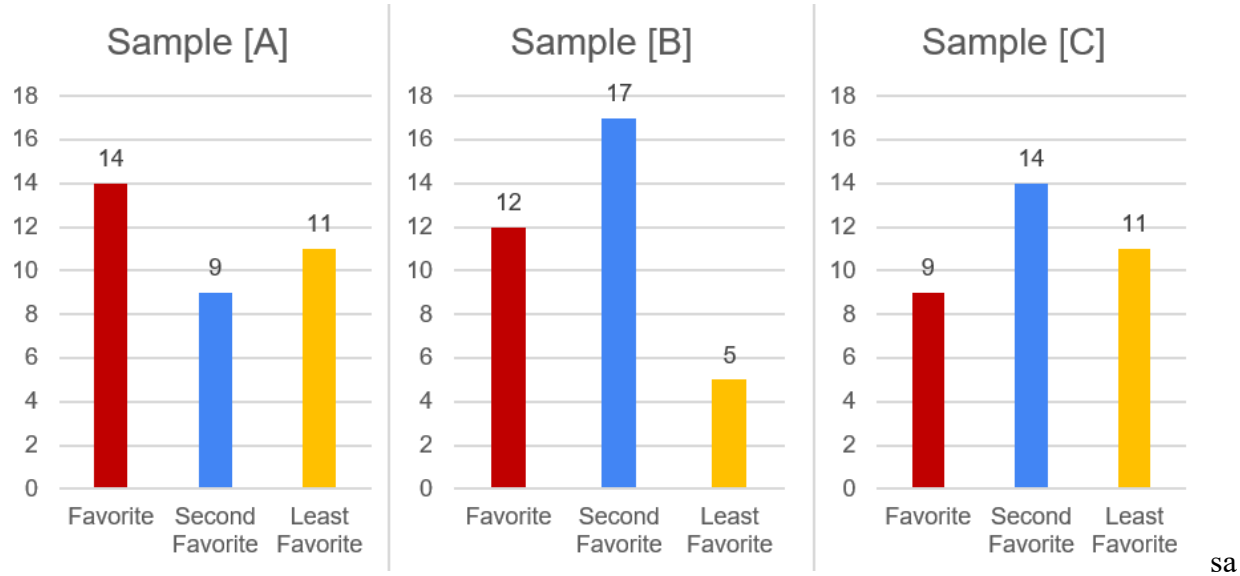
After analyzing the test results, the test team was given the key to the water samples, correlating each sample with its location. Sample A was from Marin, Sample B was from EBMUD, and Sample C was from SFRWS.

To better interpret the results, we assigned numerical values to responses according to Table 1.

Table 1: The assigned values to each response

Response	**Numerical Value**
Favorite	3
Second favorite	2
Least favorite	1

Figure 5: Results of participants' ranked favorites for the three samples



sa

Then, we totaled the numerical values based on responses for each Sample to create a total number of points. As shown in Figure 5, 12 participants rated Sample A as "Favorite" (12 participants x 3), 9 participants rated Sample A as "Second Favorite" (9 participants x 2), and 7 participants rated Sample A as "Least Favorite" (7 participants x 1), which totaled 61 points for Sample A. Applying the same calculations for Samples B and C, results show in Table 2.

Table 2: Total points for each sample

Sample	**Total Points**
A	61
B	59
C	53

Sample A, or Marin, was the "Favorite" with 61 points, Sample B, or EBMUD, was the "second favorite" with 59 points, and Sample C, or SFRWS, was the "Least Favorite" with 53 points.

In addition to conducting the taste test with public participants, we also visited Anistacia Barrak-Barber, Colorado's only certified water sommelier, for a double-blind professional assessment of water smell and taste one day after the public test. With Anistacia's authorization, the tasting process was recorded. Anistacia tasted each of the samples and ranked Sample C as her favorite, Sample A second, and Sample B as her least favorite. For more details of the results of the water sommelier's findings, refer to the Appendix.

Discussion

We hoped to test 50-75 participants, but too few passersby were willing during the time available. For data to be to be statistically significant the participant pool must be over 100 responses.[7] We only recruited 22 participants and therefore could not produce statistically significant findings from our results. More participants and more data are needed to determine if people can distinguish between our water samples.

When considering future testing with at least 100 participants, we recommend analyzing the data with both the chi-square fit for goodness and the test for independence. Fit it for goodness would help us answer our first hypothesis to determine if people can differentiate between water samples. The test for independence would answer our second question and allow us to see if there is a correlation between taste and the location of the water sample. Chi-squared would determine if there is a meaningful relationship between taste and sample location. Due to

[7] Andrade C. Sample Size and its Importance in Research. Indian J Psychol Med. 2020 Jan 6;42(1):102-103. doi: 10.4103/IJPSYM.IJPSYM_504_19. PMID: 31997873; PMCID: PMC6970301.

our limited sample size, we were unable to conduct the chi-square tests. However, we analyzed our data on a smaller scale and were able to pull general trends from our test results.

We discovered that Sample A (Marin) was the preferred water sample, placing Sample C (SFRWS) as the least favored. However, our findings present a paradox: while 44% of participants were happy to accept Sample C for their everyday drinking water, only 26.9% would accept Sample A and 24% would accept Sample B as their daily drinking water source. We concluded that a larger participant pool is necessary to establish more definitive trends in our data.

Different variables have the potential to impact water quality and subsequently water taste. Contamination from pipes and other related infrastructure can potentially expose water samples to unwanted minerals. Water that encounters copper, iron, or galvanized pipes for extended periods of time may develop a metallic or bitter taste, becoming more noticeable when water is heated up. Plastic or rubber pipes can leech certain chemicals into water, giving it an unpleasant taste. Many pipes in the San Francisco Bay Area are made of cast iron and have the potential to influence the taste of tap water. To minimize the influence of contamination from pipes, the taps were flushed for longer than five minutes to avoid collecting water that has sat stagnant in the pipes. Additionally, water temperature can influence the flavor of the water. To maximize the flavor profile, water should be kept at room temperature; colder water numbs the taste buds which mask the taste of chlorine and fluoride, whereas warmer water can increase the taste of sodium and calcium.[8]

[8] Burlingame, Gary A., Andrea M. Dietrich, and Andrew J. Whelton. "Understanding the Basics of Tap Water Taste." *Journal AWWA* 99, no. 5 (May 1, 2007): 100–111. https://doi.org/10.1002/j.1551-8833.2007.tb07930.x.

To gain deeper insight on the chemical composition of our water samples, each sample underwent analysis at SimpleLab, an independent laboratory specializing in drinking water testing. Samples A and B were ranked closely to each other, with many participants claiming that they tasted either similar or identical to one another. The chemical composition of both Samples A and B were found to be similar. Both Samples A and B and B showed similar characteristics, with nearly identical pH levels of 7.86 for Sample A and 7.98 for Sample B. In contrast, Sample C had a higher pH level of 8.23, indicating alkalinity.

Sample A registered the highest on the water hardness scale, with a rating of 69.44 mg/L. Hard water is characterized by elevated levels of dissolved minerals such as calcium and magnesium, which can give water a bitter and salty taste. Samples B and C rated lower on the water hardness scale, with Sample B at 16.61 mg/L and Sample C even lower at 11.68 mg/L, classifying them as soft waters. Soft water, like that delivered by EBMUD and SFRWS, has reduced calcium and magnesium content, typically offering a more neutral flavor profile. The laboratory analysis results are echoed in the sommelier's opinions of the water samples. For additional details of her findings and the discussion of the sommelier's interview, refer to the Appendix.

It is also important to note that the blend of water delivered by water agencies is not constant but varies over time and can vary within a service territory. For this taste test, SFRWS supplies were collected on the ground floor of the SFPUC headquarters. Likewise, EBMUD water was collected at its headquarters. Staff at the Marin headquarters declined to provide samples for testing, so the Marin water was collected at a nearby restaurant.

Conclusion & Areas for Future Research

This study aims to challenge San Francisco's claims that water from Hetch Hetchy reservoir is superior in its taste, by determining if participants can tell the difference between local water sources and by ranking three local water sources from "Favorite" to "Least Favorite." Our study contained several limitations but was a good start in developing a water taste test methodology and interpreting initial results.

The data we collected from our 28 participants in the initial test results show that participants favored Sample A (Marin) and Sample B (EBMUD) over Sample C (SFRWS). While Sample C was ranked as the "Least Favorite" sample most often, many of our participants said they would accept it as their drinking water. Further, most of our participants said they would accept all three samples as their daily drinking water. These results suggest that water taste is subjective and that residents may not be able to tell the difference between local water sources.

However, our study is inconclusive due to limitations in our test design and a lack of sufficient data for the results to be statistically significant. We identified several key barriers in our methodology and provided recommendations for future testing in Table 3 below.

Table 3: Study Limitations and Recommendations

Study Limitation	Recommendation for Future Research
Limited participants due to testing location – We conducted our test at the Rockridge BART station, a busy transit hub, in Oakland as requested by RHH (Restore Hetch Hetchy). During our test, we realized that	In future testing, we recommend RHH to conduct the taste test at a different location where people would be more willing to stop and participate in our test. Conducting the test at a college campus, such as the University of

people were quickly trying to get to a destination, leaving them little time to stop and participate in our water taste test. This made it difficult for us to collect enough data and produce statistically significant results.	California Berkeley campus during the academic year may be a better testing location, as there would be many willing college students who have time to stop and participate in our test. Changing the test location to an area with more willing participants would help RHH achieve statistically significant results.
Inherent participant bias – Before conducting the test, we recognized that taste is inherently subjective and our participants might bring their own unconscious bias to the taste test, potentially influencing results. Specifically, participants might have inadvertently preferred samples that tasted like their daily drinking water, ranked samples based on where they thought each sample was from, or for other unknown reasons.	We could not identify changes in our methodology or test design to reduce inherent participant bias. We recommend RHH closely follow our methodology and recommendations to standardize the test as much as possible. Additionally, RHH could consult water quality experts to improve the test methodology and find additional strategies that account for inherent participant bias.
Bias in sample preference from tasting order of each sample – When analyzing our results, we noticed that nearly all our participants sampled Sample A, then B, and then C. This may have influenced how participants ranked each sample. Specifically, if some participants were not able to distinguish a noticeable difference between samples, they may have ranked samples in order (with A as favorite, B as second favorite, and C as least favorite). Additionally, a water sample could	For future testing, we recommend RHH label the water samples using symbols that do not have an inherent sequencing. For example, sample labels could be "#," "%," and "&" instead of A, B, and C, which have an obvious sequence. RHH could also label the samples with "1", "2", and "3", and use a Research Randomizer, such as random.org, to randomly generate a sequence for each participant's sample testing. This would prevent participants from selecting their own

leave an aftertaste on the palette which may influence the taste of the following sample.	sequence, which may default to "1", "2", and "3" and introduce bias into the test.
Difficulty controlling water sample temperature – We identified temperature as a study limitation before our test and did our best to control the temperature of the water samples. Water quality experts recommended keeping our samples at room temperature, as this is known to be the ideal temperature to distinguish between water sources. Our samples were chilled in the fridge before the test, and then kept in a cooler while we were conducting our test outside. We believe this may have changed the temperature of our water samples throughout the test, and therefore may be influencing how our participants ranked each sample.	We recommend RHH to store water samples at room temperature before conducting future water taste tests, since experts recommend it as an ideal temperature for water testing. During the test, samples should be kept in a temperature-controlled cooler such as a Microyn Portable Refrigerator at room temperature. This will prevent any changes in temperature during the taste test, reduce any subsequent changes in sample taste or flavor, and standardize the temperature of each sample tried by every participant.
Change in sample flavor from sample container or storage time – Before our test, we identified the sample container and storage time as factors that could change the water samples' taste. To control this, we collected and stored each of our samples in glass containers to reduce taste impacts from plastic or other materials, as recommended by water quality experts. However, our samples were collected the day before conducting the taste test, which may have allowed time for the samples to change from their initial flavor.	We recommend that RHH continue to use the same glass containers to collect and store all water samples in future testing. Additionally, we recommend the collection of new samples for each test on the same day. We recommend against using the same water samples for testing that spans multiple days, as water can change flavor over time. To reduce this influence on sample taste as much as possible, RHH should collect the new water samples for each test and conduct the test on the same day.

Impact of water system infrastructure on taste – We identified the influence of water system infrastructure on water sample taste before conducting our test. Our water samples can pass through hundreds of miles of pipes, aqueducts, canals, several reservoirs, and treatment plants before reaching a tap, which may change the flavor of any water sample. To control this, water quality experts recommended "flushing" the tap, or allowing the water to freely pass through the tap, before collecting the sample.	In our test, we instructed RHH to flush the tap for five minutes before collecting each sample to prevent the collection of water that has been stagnant in pipes and may have changed taste. While the materials used in a water system's infrastructure will inherently affect water taste on any sample, we believe flushing the tap is sufficient in limiting the influence of the system's materials on our water samples taste. Therefore, we recommend that RHH continue to do so in future testing.
Water sample taste may be influenced by mineral content, pH, total dissolved solids, and other factors – Findings from our background research clearly stated that water components such as mineral content, pH, and total dissolved solids, among others, contribute significantly to water taste. Prior to conducting our test, we advised RHH to send our water samples to a testing lab to measure a few of these components. While the lab provided us with the chemical composition, pH, and mineral content of each sample, more research is needed to understand how these components contribute to each water sample's taste.	We recommend RHH conduct further research and review existing literature exploring how water components influence water taste. RHH should also continue to send water samples collected in future tests to testing labs to measure various water components. This may provide explanations for why participants preferred the taste of some water samples over others, information which RHH could use in future advocacy work challenging claims that San Francisco's water has superior taste.

In conclusion, our water taste test was a good first step in developing a detailed methodology and considering variables that may affect results. RHH should conduct more water

taste testing to increase the sample size and consider our recommendations to improve the method. Then, RHH may have more statistically significant results that show whether San Francisco Bay residents can distinguish between local water sources.

Appendix A: Water Sommelier

To further explore the differences between these water supplies in terms of smell and taste, and to understand the professional process and metrics used to taste water, our team visited Anistacia Barrak-Barber, Colorado's only water sommelier, and conducted the same double-blind taste test. (For the purposes of this visit, the questions for Anistacia were more specific and detailed than those for the public taste test, but the experimenter tasted the same samples used in the test on June 14th. At the time of the interviews with Anistacia, neither the experimenter nor the subjects knew which water source any of the samples came from).

With Anistacia's authorization, the tasting process was recorded. Below is a brief description of the testing process and her comments on the three samples in terms of smell and taste:

- *"With Sample A you can capture a pretty strong smell of Chlorine…Sample B smells kind of sweet …C has like minimal to no odor…"*

- *"A, you can actually taste the Chlorine, and it leaves your mouth a drying-out feeling…B has something that makes it taste like tainted, a sweet, perfumy taste…C doesn't have outstanding taste, just like from melted ice, leaves your mouth a cooling and moisture feeling…"*

In addition to the description itself, the sommelier also gave her conjectures on which sample belongs to which location. She assumed that Sample A was from EBMUD; Sample B was from Marin, and Sample C was from SFRWS. According to our records, the sommelier's guesses for A & B were inaccurate, while her guess for C was correct, and her description of the flavor is consistent with the lower levels of minerals found in the water from the SFRWS.

The sommelier's tasting results differed from those of the public. The sommelier ranked Sample C (SFRWS) first, Sample A (Marin) second and Sample B (EBMUD) last. The public ranked Sample A (Marin) first, Sample B (EBMUD) second, and Sample C (SFRWS) last. She favors Sample C (SFRWS) while it was ranked last by the public. We hypothesize that the lower mineral levels in the SFRWS sample may make it taste bland to the general public's palate.

These intriguing results challenge San Francisco's claims of the 'fineness' of their water's taste. If only an expert can discern the superior taste, the claim's credibility should be reconsidered. It is the residents of San Francisco and other neighboring cities who drink the water daily.

Bibliography

Burlingame, Gary A., Andrea M. Dietrich, and Andrew J. Whelton. "Understanding the Basics of Tap Water Taste." *Journal AWWA* 99, no. 5 (May 1, 2007): 100–111. https://doi.org/10.1002/j.1551-8833.2007.tb07930.x.

Capehart, Kevin W., and Elena C. Berg. "Fine Water: A Blind Taste Test." *Journal of Wine Economics* 13, no. 1 (February 2018): 20–40. https://doi.org/10.1017/jwe.2017.50.

"City of San Francisco 2023 Annual Water Quality Report." San Francisco Waterpower Sewer, 2023. https://sfpuc.org/sites/default/files/accounts-and-services/water-quality/EN-SF-2023.pdf

Cook, Sherri. Consulting Dr. Sherri Cook to Develop a Water Taste Test Methodology. Interview by Katherine Halama, Jazlyn Nie, and Megan Franco, May 17, 2024.

Decker, Zach. Consulting Dr. Zach Decker to Consider the Scientific Method in a Study. Interview by Katherine Halama, Jazlyn Nie, and Megan Franco, May 13, 2024.

Jhuang, Jing-Rong, Wen-Chung Lee, and Chang-Chuan Chan. "A Randomized, Double-Blind Water Taste Test to Evaluate the Equivalence of Taste between Tap Water and Filtered Water in the Taipei Metropolis." *Scientific Reports* 10, no. 1 (August 7, 2020): 13387. https://doi.org/10.1038/s41598-020-70272-y.

Labaree, Robert V. "Research Guides: Organizing Your Social Sciences Research Paper: 6. The Methodology." Research Guide. Accessed May 19, 2024. https://libguides.usc.edu/writingguide/methodology.

Lunsford, M. Leigh, and Alix D. Dowling Fink. "Water Taste Test Data." *Journal of Statistics Education* 18, no. 1 (2010). https://www.tandfonline.com/doi/epdf/10.1080/10691898.2010.11889484?needAccess=true.

Masters, Sheldon. Consulting Dr. Sheldon Masters to Develop a Water Taste Test Methodology. Interview by Katherine Halama, Jazlyn Nie, and Megan Franco, May 28, 2024.

San Francisco Water Power Sewer. “Water Quality.” Accessed June 5, 2024. https://sfpuc.org/accounts-services/water-quality.