How to Poach Eggs, and Whether You Should Reconsider Dining at The French Laundry

Nicholas K. Branigan

December 9, 2021

1 Abstract

I performed an experiment to determine how the visual quality of a poached egg is related to the poaching water's temperature, depth, and vinegar content, as well as whether the egg is strained before poaching, and whether the poaching water is disturbed with a vortex before the addition of the egg. Evidence from the experiment suggests that vinegar can be avoided and vortex should be avoided. The former has no positive effect and the latter has a pronounced negative one. Conclusions for the other factors are more complicated. My results defy guidance from sources as exalted as Thomas Keller of The French Laundry fame. Furthermore, the experiment reveals that the technique used to add the egg to the water, which sources on poaching eggs neglect to mention, could be important to producing a beautiful poached egg. If I were to re-run the experiment, I would formally include this technique as a factor, and, if I had a lot of time on my hands, get many more replicates. The number of observations (32) prevents analysis of how the factors interact with the high and low cut-points and interferes with the fitting of models including all main effects and first order interactions because of complete or quasi-complete separability issues.

2 Introduction

Poaching is a classical method of preparing eggs. Poached eggs have a starring role in eggs benedict, and they can elevate dishes ranging from salads and soups to sandwiches and pastas. Unfortunately, poaching eggs beautifully is difficult. The task is often left to professionals. But, for the brave, there is no shortage of advice on how to poach eggs at home. Famous chefs, cooking publications, and even TikTok celebrities have formed strong opinions about how to best poach eggs and would like to share their findings with you. However, in my experience, cooking practices can be based on very weak evidence. Occasionally, people will try to test cooking ideas, and it is not uncommon for them to find that even methods embraced by the world's most-renowned culinary professionals have little or no effect [3].

In this experiment, I put to the test five factors in the poaching of eggs: the temperature of the water used, the depth of the water, whether the egg is strained before poaching, whether vinegar is added to the water, and whether a vortex is created in the water immediately before the egg is added. I believe that these are the most prominent recommendations from more reputable sources. I was most confident that straining would have a positive impact, since it is advocated these days by a large number of sources, including those that appeal to scientific explanations and, sometimes, even experiments. Still, I was not especially confident that any factor would have a positive effect. I have found justifications for the methods I tried offered by Thomas Keller and J. Kenji López-Alt. Keller is the chef and owner of The French Laundry in Yountville, CA and Per Se in New York City. Both restaurants have been awarded three Michelin Stars each year since the inaugural editions of The Michelin Guide for the San Francisco Bay Area and Wine Country (2007) and New York City (2006) [8] [9]. He is the first and, so far, only American chef to have two of their restaurants receive this most coveted award [14]. J. Kenji López-Alt is a chef and food writer. He is famous for putting cooking practices to the test through experiments. His New York Times best selling book is titled "The Food Lab: Better Home Cooking Through Science" [5]. Following are the justifications.

> "Poaching at a 180°F sub-simmer means fewer disturbances in the water, and cleaner, more tender cooking." [10]

"The best method for poaching eggs is in a deep pot of water. As the weight of the yolk pulls the egg through the water, the white encircles the yolk and sets. The deeper the water, the farther the egg travels before it stops, and the more the poached egg will resemble its original shape." [7]

"Straining eggs in a fine mesh strainer removes excess wispy, loose whites, creating tight, perfectly-shaped eggs." [10]

"Adding distilled vinegar to the boiling water is essential to help set the egg white." [6]

"The swirling water [of the vortex] will help the egg white envelop the yolk evenly as the egg white proteins set, creating a nice natural shape." [6]

3 Experiment

3.1 Measuring poached egg quality

My measurement of the quality of poached eggs is subjective. My goal in this experiment was to find the method that results in the best poached eggs *for me*. To this end, I developed a rating scale based entirely on the appearance of the eggs. I reasoned that the different cooking methods would result in eggs that would taste essentially the same in a blind tasting. The goal, then, became to produce a visually pleasing egg. Broadly, I decomposed the visual characteristics of a poached egg into texture and shape. My ideal texture is smooth and without wrinkles and wisps. My ideal shape is a subtle tear drop. For my goal, see FIGURE 1. After examining a lot of photos of poached eggs from the internet, as well as from my experiment, I arrived at the following scale, where texture * shape denotes the combined quality of texture and shape (i.e., roughly speaking, bad texture and good shape should get the same score as okay texture and okay shape).

1. Texture * shape is appalling. Yolk and whites may be completely separated. This is a complete embarrassment.
2. Texture * shape is bad. Yolk and whites are not separated. Ugly but clear progress over 1. I might eat one of these.
3. Texture * shape is decent. Looks unmistakably like a poached egg. Not a disappointment, but both texture and shape leave something to be desired. I would eat one of these.
4. Texture * shape is good. This is a nice poached egg.
5. Texture * shape is really nice. Differences with 4 are subtle, and, I think, subject specific. 5 is sometimes just intangibly more visually pleasing to me than 4.

3.2 Design

The levels I chose for my factors were temperature at 185 degrees or 200 degrees Fahrenheit, water at 2.75 inches or 5.5 inches of depth (corresponding to 3.4 liters and 6.8 liters of water), no straining or straining, no vinegar in the water or approximately 1 part of vinegar per 100 parts of water, and no vortex or vortex. Under no straining, I cracked the egg directly into a ramekin from which I placed it directly into the water. Under straining, I cracked the egg into the strainer, gently swirled 10 times, let the egg sit for 2 minutes, gently swirled 10 more times, then delicately transferred the egg from the strainer into the ramekin. I achieved the vortex each time by making deep and large circles (with the spoon touching the bottom as well as the side of the pot during the circle motion) 10 times at about 1 hertz.

I used a 2^{5-1} fractional factorial design with Resolution V whose defining relation is $I = ABCDE$. I did two replicates, treating each replicate as a block. I reasoned that I had time for 32 trials, and that the replicates would be worth gaining at the cost of aliasing with third, fourth, and fifth order interactions, which I suspected were zero. I randomized the order of the runs within each replicate block. I purchased 4 cartons of Alexandre Family Farms

eggs (12 eggs per carton).[1] Each carton had the same pack date (Nov. 15), which suggests that the freshness of all of the eggs was very similar. Since the cartons shared a pack date, I assumed that there would not be differences across the cartons and did not randomly select eggs from the cartons for the runs. Instead I went carton by carton for simplicity. I poached the eggs one at a time, and used new water for each trial, cleaning the pot, strainer, and ramekin between trials. I poached all eggs from essentially fridge temperature. (Those being strained were at room temperature for roughly an additional 2 minutes than those not being strained.) Once an egg was done poaching, I transferred it to a plate (the same for all eggs), and tried to make it as presentable as possible by wiping excess water and sometimes trimming the edges with scissors. I took every photo with the same phone from about the same angle (from above) under about the same lighting conditions. I measured the water temperature using a food thermometer, and I began trials (added the egg) only when the temperature read within 0.5 degrees Fahrenheit of the desired temperature. The photos I took included trial information written on the plates. In order to rate the egg photos blind, I cropped this information out of the photos and then randomized the order of the cropped photos and rated these de-identified egg pictures. I ran the two blocks back to back on the same day. For the exact procedure I used to run the experiment, see the appendix.

3.3 Calibration

Before running the experiment, I needed to determine the right level of vinegar to use, the correct cooking time at each temperature, and how to set the gas burners on my stove-top to maintain the temperatures. To determine the vinegar level, I used Thomas Keller's recommendation that I should look for a "slightly acidic taste" [6]. This turned out to be around 1 part distilled vinegar per 100 parts water. I adjusted the cook time to achieve my ideal doneness. I was looking for fully set whites and a runny yolk. I achieved this at 3.5 minutes at 200 degrees and 4.5 minutes at 185. I also wanted to ensure that the temperature once at 185 or 200 (in practice, within 0.5 degrees of these marks) would not fluctuate much during the poaching session. I settled on setting the burner a quarter of the way from low to medium for 185 and halfway between low and medium for 200. In my calibration testing this resulted in temperature fluctuations less than 3.5 degrees over 5 minutes.

4 Analysis

4.1 Methods

I analyzed the data using logistic regression with varying cut-points on the rating. Let rating $\geq j$ denote the indicator that is 1 when rating $\geq j$ and 0 otherwise. I offer analyses on rating ≥ 3 and rating ≥ 4. Unfortunately, there were too few observations where the rating was 1 or 5 to offer the others. These models faced quasi-complete or complete separability problems. And the issue of separability showed up for other analyses too. Including too many interactions in the rating ≥ 3 or rating ≥ 4 models made them too separable. The approach I took then to examine interactions was to add the interactions one at a time to a model and see if they were significant ($\alpha = 0.05$), or close. By this procedure, I identified 3 interactions of possible interest in the rating ≥ 4 model (water*vinegar, temp*water, and water*vortex) and zero in the rating ≥ 3 model. Unfortunately, adding all three to the rating ≥ 4 model raised separability problems. In fact, adding water*vinegar or temp*water to a rating ≥ 4 model containing all main effects and water*vortex led to separability problems.

In Table 1, I present the model for rating ≥ 3 (model 1), and in Table 2 the model for rating ≥ 4 with only main effects (model 2). For rating ≥ 4, I have the model including water*vinegar and temp*water (model 3) in Table 3, and the the model with water*vortex in Table 4.

I considered an ordinal logistic regression as well. However, Table 1 and Table 2 provide evidence against the assumption of proportional odds that an ordinal logistic model makes. Specifically, consistency with proportional odds would require that the coefficients in the binary logistic regression models with different cut-points between the binary categories would be quite close (with the exception of the intercept) [11]. In fact, we see that coefficients (e.g., temp and strain) change dramatically as we change the cut points. Since the full model with all interactions

[1] Confined Animal Feeding Operations (CAFO), better known as factory farms, constitute the world over "a crime of stupefying proportions" [2] [13]. And among CAFOs, egg "farms" are perhaps the very worst [12]. I take intense ethical issue with brutalizing non-human animals, nonstop from birth to death, for the sake of gustatory pleasure in humans. The eggs I have used bear the Certified Humane label for pasture-raised eggs. The Animal Welfare Institute gives this label its top rating for requiring the "highest animal care standards" with compliance "verified by a third-party auditing program" [1].

	Estimate	Std. Error	z value	$\Pr(>\|z\|)$
(Intercept)	-0.1622	1.2528	-0.129	0.8970
temp	2.2485	1.1112	2.024	0.0430 *
water	1.4000	1.0284	1.361	0.1734
strain	-0.5173	0.9900	-0.523	0.6013
vinegar	-0.5173	0.9900	-0.523	0.6013
vortex	-2.2485	1.1112	-2.024	0.0430 *
block	2.2074	1.0890	2.027	0.0427 *

Table 1: $\text{logit}(\Pr(\text{rating} \geq 3)) = \text{intercept} + \beta_1\text{temp} + \beta_2\text{water} + \beta_3\text{strain} + \beta_4\text{vinegar} + \beta_5\text{vortex} + \beta_6\text{block}$.

	Estimate	Std. Error	z value	$\Pr(>\|z\|)$
(Intercept)	-0.8540	1.0908	-0.783	0.434
temp	-0.3456	0.8377	-0.413	0.680
water	0.3456	0.8377	0.413	0.680
strain	1.0258	0.8550	1.200	0.230
vinegar	-0.3456	0.8377	-0.413	0.680
vortex	-1.6479	0.8795	-1.874	0.061
block	1.6500	0.8808	1.873	0.061

Table 2: $\text{logit}(\Pr(\text{rating} \geq 4)) = \text{intercept} + \beta_1\text{temp} + \beta_2\text{water} + \beta_3\text{strain} + \beta_4\text{vinegar} + \beta_5\text{vortex} + \beta_6\text{block}$.

has separability issues, I could not compare these for the different cut points. However, I ran an ordinal logistic regression on the main effects and all interactions and examined the binary score residuals graphically as suggested by [4]. This visual check suggested that proportional odds is not met for this full model either.

4.2 Inferences

There are three clear takeaways. Vortex is bad, vinegar is not good, and there is a positive block effect. Across all four models, the high level of vortex has a negative coefficient and this is significant at $\alpha = 0.05$ in two of the models and nearly significant for the other two. Vortex also has a negative interaction with water that is nearly significant in model 4. The high level of vinegar has coefficients in all models close to zero (in fact, slightly negative in each one), and in model 3 has a negative interaction with high water that is significant. The high level of block (the second block) has a positive coefficient across all models, and this is significant in two of the models and verging on significant in the other two. I interpret the block effect as evidence that I got better at poaching eggs.

I designed the experiment so that little, outside of the factors I was explicitly varying, could affect the quality of the poached eggs. But I realize after running the experiment that there was a subtle variable I had not carefully controlled, which is the technique I used to add the eggs to the water. There is some skill involved here and different possible approaches. Without carefully thinking about what I was doing, I started to take a more gentle approach over the course of the experiment. At first, I dumped the eggs in rapidly from a few inches over the water. But by the second block, I was gently reversing the ramekin into the water with the bottom of the ramekin touching the water. This is not optimal experimental execution, but I think the inclusion of the blocking variable may mostly prevent this from being a major confound. Ultimately, this hiccup may have been for the best, since it suggests that the method used to add the eggs to the water, which was not highlighted as worth considering by any source I consulted on poaching eggs, could be one of the most significant factors in the final appearance of the eggs.

What about temperature, water, and straining? Here things are murky. There is an indication that the high level of temperature may offer a benefit for achieving decent or better poached eggs (rating ≥ 3). However, for achieving nice or really nice eggs (rating ≥ 4), the high level of temperature seems to offer no benefit, and in interaction with high water could be bad. The effect of water is unclear. For (rating ≥ 3) there is a suggestion that high water could be good, but this diminishes for rating ≥ 4. Moreover, for rating ≥ 4, we see negative interactions with vinegar, temperature, and vortex that are nearly significant or significant. The effect of straining is unclear,

	Estimate	Std. Error	z value	$\Pr(>\|z\|)$
(Intercept)	0.9516	1.6584	0.5738	0.5661
temp	-0.252	1.0748	-0.2344	0.8146
water	0.3613	1.0505	0.3439	0.7309
strain	1.3748	1.1125	1.2358	0.2165
vinegar	-0.252	1.0748	-0.2344	0.8146
vortex	-2.3422	1.1754	-1.9927	0.0463 *
block	2.4903	1.2384	2.0109	0.0443 *
water*vinegar	-2.3785	1.1724	-2.0288	0.0425 *
temp*water	-2.3785	1.1724	-2.0288	0.0425 *

Table 3: $\text{logit}(\Pr(\text{rating} \geq 4)) = \text{intercept} + \beta_1\text{temp} + \beta_2\text{water} + \beta_3\text{strain} + \beta_4\text{vinegar} + \beta_5\text{vortex} + \beta_6\text{block} + \beta_7\text{water*vinegar} + \beta_8\text{temp*water}$.

	Estimate	Std. Error	z value	$\Pr(>\|z\|)$
(Intercept)	0.0412	1.2509	0.0329	0.9737
temp	-0.368	0.9143	-0.4025	0.6873
water	0.1632	0.9398	0.1736	0.8622
strain	1.1946	0.9407	1.27	0.2041
vinegar	-0.368	0.9143	-0.4025	0.6873
vortex	-1.9181	0.9931	-1.9314	0.0534
block	1.9505	0.9971	1.9562	0.0504
water*vortex	-1.9088	0.9908	-1.9265	0.054

Table 4: $\text{logit}(\Pr(\text{rating} \geq 4)) = \text{intercept} + \beta_1\text{temp} + \beta_2\text{water} + \beta_3\text{strain} + \beta_4\text{vinegar} + \beta_5\text{vortex} + \beta_6\text{block} + \beta_7\text{water*vortex}$.

with rating ≥ 3 and rating ≥ 4 providing weak mixed evidence on its effect.

5 Limitations

Other than the small sample size, and the hiccup in experimental protocol, there is an issue of external validity here. None of the 32 eggs I poached resembled even the uglier of the two beauties in Figure 1, which were the product of Thomas Keller's efforts. The explanation I prefer for this is that the eggs he was using were likely extremely fresh, laid just days before he poached them. The eggs I used were as fresh as I could find at my local Whole Foods, but they were laid probably around 3 weeks before I ran my experiment. Every source I have seen claims that egg freshness is the single greatest factor in the quality of poached eggs. I suspect that my results could fail to generalize to very fresh eggs. This has two consequences. First, since very fresh eggs are very hard to come by, my results should generalize better to the setting of most people trying to make poached eggs at home. Of course, this generalization is contingent on my visual tastes in poached eggs matching those of the general public. For the rating ≥ 3 models, this seems like a safe assumption. Second, the title of this project is a jest. In fact, I have a reservation at The French Laundry this Sunday, and my results have done nothing to temper my excitement.

6 References

[1] *A Consumer's Guide to Food Labels and Animal Welfare*. en. URL: https://awionline.org/content/consumers-guide-food-labels-and-animal-welfare (visited on 12/09/2021).

[2] J. M. Coetzee. *Elizabeth Costello*. New York: Penguin Books, 2004. ISBN: 9780142004814 9780670031306.

[3] *Debunking Burger Myths with J. Kenji López-Alt*. en. URL: https://www.youtube.com/watch?v=weFT03McaH0 (visited on 12/09/2021).

[4] Frank E. Harrell. *Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis*. Springer series in statistics. New York: Springer, 2001. ISBN: 9780387952321.

[5] *J. Kenji López-Alt*. en. URL: https://www.seriouseats.com/j-kenji-lopez-alt-5118720 (visited on 12/09/2021).

[6] Thomas Keller. *Thomas Keller Teaches Cooking Techniques I: Vegetables, Pasta, and Eggs*.

[7] Thomas Keller et al. *The French Laundry cookbook*. 2nd ed. New York: Artisan, 1999. ISBN: 9781579651268.

[8] *List of Michelin starred restaurants in New York City*. en. Page Version ID: 1048535265. Oct. 2021. URL: https://en.wikipedia.org/w/index.php?title=List_of_Michelin_starred_restaurants_in_New_York_City&oldid=1048535265 (visited on 12/09/2021).

[9] *List of Michelin starred restaurants in San Francisco Bay Area*. en. Page Version ID: 1047740364. Oct. 2021. URL: https://en.wikipedia.org/w/index.php?title=List_of_Michelin_starred_restaurants_in_San_Francisco_Bay_Area&oldid=1047740364 (visited on 12/09/2021).

[10] J. Kenji López-Alt. *Easy Poached Eggs Recipe*. en. URL: https://www.seriouseats.com/foolproof-poached-eggs-food-lab-recipe (visited on 12/09/2021).

[11] Keith McNulty. *Handbook of regression modeling in people analytics: with examples in R and Python*. First edition. Boca Raton: CRC Press, 2021. ISBN: 9781032041742 9781032046631.

[12] Michael Pollan. "An Animal's Place". en-US. In: *The New York Times* (Nov. 2002). ISSN: 0362-4331. URL: https://www.nytimes.com/2002/11/10/magazine/an-animal-s-place.html (visited on 12/09/2021).

[13] Peter Singer. *Animal liberation: the definitive classic of the animal movement*. Updated ed., 1st Ecco pbk. ed., 1st Harper Perennial ed. New York: Ecco Book/Harper Perennial, 2009. ISBN: 9780061711305.

[14] *Thomas Keller - Biography*. URL: https://www.thomaskeller.com/biography (visited on 12/09/2021).

7 Appendix

7.1 Details of experimental procedure

Following is the experimental procedure I followed for each trial.

1. Write the trial number on the plate.
2. Add correct amount of water to the pot.
3. Add correct amount of vinegar to the pot (stir if adding vinegar).
4. Remove egg carton from the fridge.
5. Crack egg into the strainer or ramekin and put the egg carton back into fridge.
6. If straining, swirl 10 times, let sit for 2 minutes in the strainer (using timer), swirl 10 more times, then transfer to the ramekin.
7. Achieve the water temperature.
8. Add vortex or not.
9. Add egg to the water.
10. Start the timer (3.5 or 4.5 minutes).
11. When the time is up, remove the egg and place it onto the plate.
12. Trim the egg and wipe excess water.
13. Photograph from above including trial information.
14. Remove the egg from the plate and wipe the plate.
15. Dump the water in the pot and rinse with cold water, scrubbing if necessary.
16. Rinse the strainer with cold water if it was used during the trial.
17. Rinse the ramekin with cold water.

The primary bottleneck in this procedure was getting the water to 185 or 200 degrees. I had two electric kettles (1.7 L capacity each) that I used to get the water close to the correct temperature. I pipelined this to the extent possible so that I would be running the kettles during one trial with the water for the next trial. The full experiment took about 10.5 hours, with the first block taking considerably longer than the second. (By the second block, I had gotten better at heating the water ahead of time.)

7.2 Data table

temp	water	strain	vinegar	vortex	block	y
1	1	1	0	0	0	4
0	0	0	0	1	0	1
0	1	0	1	1	0	2
1	0	1	1	0	0	1
1	1	0	0	1	0	3
0	1	1	0	1	0	1
0	0	1	0	0	0	1
0	0	0	1	0	0	2
1	0	0	1	1	0	3
1	1	0	1	0	0	3
1	0	0	0	0	0	3
0	1	1	1	0	0	5
0	0	1	1	1	0	4
1	0	1	0	1	0	2
0	1	0	0	0	0	4
1	1	1	1	1	0	2
0	1	0	0	0	1	5
0	0	0	1	0	1	4
0	1	1	0	1	1	4
1	1	1	0	0	1	4
0	0	0	0	1	1	2
0	0	1	0	0	1	3
1	0	0	0	0	1	4
1	0	1	1	0	1	4
0	1	1	1	0	1	4
1	1	0	1	0	1	3
0	0	1	1	1	1	2
1	0	0	1	1	1	5
1	1	0	0	1	1	3
0	1	0	1	1	1	2
1	1	1	1	1	1	3
1	0	1	0	1	1	5

7.3 My ideal poached egg

Figure 1: My conception of a poached egg that lies in a small neighborhood of perfection (specifically the one on the right) [6].