

Nutella: US vs Europe

Jacopo Mauro, Johanna Stumpf

Department of Informatics, University of Oslo, Norway.

Abstract. Since the dawn of time (or at least since the invention of Nutella) the human kind has strived for the hamletic question “Can human distinguish different version of Nutella sold in different countries.” In this paper we present a simple empirical experiment to shed some light on this important and vital question. Empirical evidence seems to indicate that the Nutella sold in the US can be distinguish from the one sold in Europe. Further investigations are however needed to confirm the statistical significance of the claim.

1 Introduction

Nutella is probably one of the most well-known brand names of a sweetened hazelnut cocoa spread. Since its introduction to the market in the second half of the century it has become the source of happiness (and heart attacks) for countless people. Nutella nowadays is appreciated all over the world and it can be found in many disparate countries.

Nutella has a huge follow up and plenty of fans all over the world. There is even an official Nutella day (5th of February), i.e., a celebration day when many tasting events, thematic parties and challenges are held in honor of Nutella. The spread of Nutella in the entire world triggers the appearance of high intensity discussion about:

1. Who are the best eaters of Nutella?
2. Which is the best hazelnut cream rival of Nutella?
3. Which jar size is better (400 g, 450, . . . , 1Kg, 5kg)?
4. Are the Nutella produced in different countries different, and if so, can the human distinguish them?

These questions leave margin to a huge debate. For instance, the answer to the first question depend on the meaning attributed to the word “best”. If by best we mean “who is eating more” French should be leading since France is the country that consumes the most jars of Nutella. If instead by best we mean the ones appreciating the quality most, probably the answers will be Italians since an Italian was the creator of Nutella and because in Italy people still eat Nutella even though plenty of well-know alternatives are available.

There is no common consensus on the best rival of Nutella. Certainly different prestigious rivals exists such as “Crema cuor di cacao” by Venchi, “Crema Novi” by Novi, or “Crema di Nocciole” by Guido Castagna. Unfortunately previous studies such as [4] were inconclusive in determining the best one. No common consensus exist also for the size of the Nutella jar since clearly a bigger jar usually has a better quantity- price ratio but, thanks to the fact that Nutella can be addictive, having a big jar could also mean eating a lot of Nutella thus increasing the stress for the upcoming swimsuit season.

As far as the last question is concerned, rumors speculate that the Nutella produced in different countries differs [2,5]. In this work we are focusing on this last question trying to discover if those differences in production material allow humans to distinguish them.

We report the result of an empirical testing experiment where test subjects were asked to eat two samples of Nutella, one bought and produced in Germany, and one bought in the US and produced in Mexico. Empirical evidence seems to indicate that the test subjects were able to perceive the differences between the different version of Nutella. Further investigations are however needed to confirm the statistical significance of the experiment.

2 Background

Following World War II, in 1946 the shortage of supplies in the chocolate and cocoa industry had a big impact on European bakeries. An Italian baker named Pietro Ferrero found a solution to the shortage of cocoa and demand for chocolate. Inspired by the Gianduaia chocolate spread invented in Turin during Napoleon's regency, Pietro created a spread mixing sugar, hazelnuts and just the slightest bit of cocoa powder. He called it *Giandujot*. Giandujot became a wild success and the 14th of May 1949 Ferrero became an official trademarked company. Around 1951, the spread was remastered to be easily spread onto bread and called *SuperCrema*.

In 1964, following the death of his father, the son of Pietro Ferrero, Michele, created a new recipe for the SuperCrema: the first jar of Nutella was born. Nutella spread from only Italy to across Europe. Germany fell in love with the spread and lead to its cultural diffusion across the eastern hemisphere. In 1965 Nutella reached France, expanding further into Europe. This was the start of a continuous growth that lead Nutella to be produced in various facilities worldwide. For the European and South Africa markets Ferrero has 2 plants in Italy, one in France, one in Germany, one in Poland, one in Turkey. For the North American market, Nutella is produced in Canada and Mexico. For Australia and New Zealand, Nutella has a plant in Australia. Finally, for the Brazilian market Ferrero has a plant in Brazil.

Nutella is made from sugar, palm oil, hazelnuts, cocoa, skimmed milk powder, whey powder, lecithin, and vanillin. The receipt is kept as a secret by the Ferrero company which also uses special designed machines to guarantee the quality of the production of every Nutella jar.

3 Methodology

The experiment performed was designed to answer to the question if humans are able to distinguish a Nutella spread sold in the US and a Nutella spread sold in Europe. For this reason we bought a random sample of the following jars.

- *US sample*. Made in Mexico. Jar size: 371 g. Ingredients: sugar, palm oil, hazelnuts, cocoa, skim milk, whey (milk), lecithin as emulsifier (soy), vanillin. Best before April 6 2017. Suggested "one serving" size: 37 g.

Considering a portion of 37 g it contains 12 g of fat, 15 mg of sodium, 23 g of carbohydrates, 2 g of proteins.

- EU sample. Made in Germany. Jar size 750 g. Ingredients: sugar, palm oil, hazelnuts (13 %), low fat cocoa, skim milk powder (7.5 %), lecithin as emulsifier (soy), vanillin. Best before May 4 2017. Suggested “one serving” size: 15 g. Considering a portion of 100 g it contains 31.8 g of fat, 102 mg of sodium, 56.9 g of carbohydrates, 6.6 g of proteins.

By the nutrition information given on the jars we can see that the declared ingredients are more or less the same but the suggested “one serving” portion of Nutella in Europe equals 15g, while in America it equals 37g. It is left as an exercise to the reader to draw conclusions from this fact.

To create fair and equal test conditions, we kept the two jars for months in a dark location at room temperature. We then gathered volunteers to be used as test subjects. No compensation was promised to them.

The experiment was constituted by 2 sessions of Nutella testing for every test subject. In every session the test subject was exposed to two tea spoons of Nutella simultaneously. The probability to be exposed to different spreads was 1/2. When the spread selected was the same we select the US spread with probability 1/2 and the EU spread with probability 1/2. To assign the sample to every test subject in a randomic way we used the functionality provided by www.random.org that allows to generate random bytes from atmospheric noise, which for many purposes is better than the pseudo-random number algorithms typically used in computer science. The information about the probabilities was communicated to the test subjects before the experiments.

The tests subjects were first grouped in a room where a questionnaire was presented to them. The questionnaire was given to evaluate their generalities and their attitudes toward chocolate. In particular, the following question were presented to the test subjects:

- Age (expected answer an integer value between 0 and 150)
- Instruction level (expected answer in { none, bachelor degree, master degree, PhD, PhD in a hard science })
- Sex (expected answer a value in { male, female, yet unknown })
- What is your attitude towards chocolate (expected answer in { chocolate what?, I hate it, I am neutral, I like it, I love it, I can not live without it })
- How many times do you eat chocolate in average in a month (expected answer in { never, between 1 and 5, between 6-50, too much })

The test subject having an age below 2 were discarded to avoid possible communication bias. We discarded also the chocolate haters to avoid possible outburst of violence between them and the chocolate lovers. Moreover, to minimize the impact of people not understanding probabilities, we discarded all the subjects not having at least a bachelor degree. The test population thus obtained was constituted by 17 people.

For every one of the two experiments rounds the test subjects were allowed to inspect the two eating samples of Nutella using all their senses with a time limit of 64 seconds. After this period, or when the subject was ready, they were asked the following questions:

Id	First session			Second session		
	Spread	Answer	Preference	Spread	Answer	Preference
1	=	†		≠	✓	EU
2	=	✓		≠	✓	EU
3	≠	✓	EU	=	✓	
4	≠	✓	EU	≠	†	
5	=	†		=	†	
6	=	†		=	†	
7	=	†		≠	✓	
8	=	✓		≠	✓	EU
9	≠	✓	EU	≠	✓	EU
10	≠	✓	US	=	✓	
11	≠	✓	US	=	†	
12	≠	✓	EU	≠	†	
13	≠	✓	EU	=	✓	
14	≠	✓	EU	=	✓	
15	≠	✓	EU	=	✓	
16	=	†		=	✓	
17	-	-		≠	✓	EU

Table 1. Experiment Results.

1. Do you believe that the spreads of Nutella are different (expected answer a boolean value)¹
2. Which sample do you like the most ? (expected answer in { first, second, undefined })

The two testing session were separated by at least 64 seconds. Test subjects were allowed to go to the bathroom and consume liquids, or food between the sessions.

The experiments were conducted at the University of Oslo, Friday the 3rd, 2017 i.e., the last working day before the World Nutella day.

4 Results

The results of the two session are presented in Table 1. For every session we reported in the first column if the samples given to the user were equal (=) or different (≠), while in the second column we report the answer provided (✓ if was correct, † if it was wrong, – if the answer was not provided). The third column shows instead the preference given by the test subject provided that the samples were different and that the test subject identified this correctly.

During the experiments the test subjects could commit two kinds of errors: saying that the samples were different while instead they were equal (false positive), and saying that the samples were equal while they were different (false negative). In the first session there have been 5 false positive and 0 false negatives out of 16 trials, while in the second session there have been 3 false positive and 2 false negative out of 17 trials.

¹ Considering the effect of using three values or fuzzy logics is left as a future work.

Assuming that the null hypothesis is that humans would not be able to distinguish the two chocolate spreads, the answer to the trial questions should be a sequence of Bernoulli trials with probability 0.5. Given n trials, assuming that user can not distinguish the samples, we would therefore expect $n/2$ wrong answers. Luckily this did not happen and in both sessions the number of wrong answers is below half of the trials, thus indicating that the test subjects have a form of discriminating power. Unfortunately, the number of the trials were not big enough to confirm the statistical significance of this claim. To test the significance the binomial test was used [1]. For the first session, the probability for having 5 or less wrong answers out of 16 trials is $\sim 10.51\%$ while for the second session the probability for having 5 or less wrong answers is $\sim 7.17\%$. These values, based on the common consensus that establish that a p-value should be less than 5% to reject the null hypothesis [3], are unfortunately not small enough. More investigations, with more test subjects and a large amount of samples is therefore needed and left as a future work.

Since the test subjects were allowed to mistake two samples of the same spread as different, another interesting question is to find out what happens when the spreads given to the test subject were indeed different. In the first session there have been 9 cases of subjects that received different sample and provided an answer. In all of these cases they correctly identified the difference between the spreads. In the second session instead there were two errors out of 8 trials. The p-value in this case for the first session is $\sim 0.20\%$ while for the second case is $\sim 14.45\%$. We therefore can sustain with statistical significance that in the first round of test, if the spreads are different, humans are able to detect that. We conjecture that the statistical significance was not obtained in the second session due to the diminishing discrimination power of the test subjects. Indeed, based on conversation sustained with the test subject after the experiments, the test subjects reported that they believed that it was harder to distinguishing the taste after eating the first two samples of Nutella. Exemplary is a comment left on the questionnaire in the second round stating “Sensitivity already weaker”.

As far as the preferences between the different sample is concerned we can see a tendency of preferring the EU sample. Unfortunately also in this case the results are not statistically significant and further studies are required. We would like to note also that the preference toward the EU sample may also be due to the familiarity of the test subject to this kind of sample since the majority of the test subjects were living in Europe, having limited access and familiarity with respect to the US spread.

To conclude this section we would like to comment on the answers to the questionnaire. We notice that 2 out of 17 test subjects were “trying to be funny” by providing non typical answers.² In particular these two jokers³ did not tick the option to state that they were not mad or terrorist and they identified their sex as “unspecified”. Interesting

² Being the test subjects coming from a computer science environment— an environment that usually is identified with the nerd culture —we believe that the presence of only 2 jokers is quite an accomplishment and a testimony to the good selection procedure of the test subjects.

³ Warning: the experiments were clearly anonymous but due the size of the test subject population, with the help of the authors, it may be possible to find out with high probability the identity of the jokers. As designers of the experiment we deeply discourage test subject to mess up with Nutella experiments again!

enough, also the jokers did dare to joke on the question related to their attitude towards chocolate. This question was indeed answered by all the test subjects, with no one reporting to hate chocolate and only 3 reporting a neutrality toward it. A deeper study on how people are prone to make fun on chocolate related questions is outside the scopes of this paper and left as a future work.

5 Conclusions

Preliminary investigations show that the Nutella sold in the US and the one sold in Europe could be distinguished. Further investigation and at least 5 kg of Nutella are needed to confirm the statistical significance of the findings.

Test subjects reported that one of the main ways to distinguish the spread was due to their color. The US sample was reported to be lighter in color and more liquid, while the EU sample was darker and more solid. Further investigations are needed to ascertain if the two spread can be distinguish by tasting only.

We can see a clear preference for the EU Nutella in the test subjects. Based on personal conversations conducted after the experiments or comments left on the questionnaire the EU Nutella was described as “richer in taste” while the US version was defined as “too sticky”.

To conclude we would like to warn our fellow scientist that doing experiments with Nutella may be dangerous and have an impact on the immediate life span of the test subjects. For instance, a test subject reported that after the experiments s/he were so full of energy and wanted to stay at work longer. Moreover, as reported by another test subject, Nutella is so addictive that it may trigger the desire to have waffles with it the day after the experiment.

Ethical concerns

No animals were killed for these experiments. To be ecofriendly no Nutella was wasted since the designers of the experiments offered voluntarily to finish what was remaining in the Nutella jars. The designers were not paid to conduct the experiments and sadly are not tied to Ferrero nor any chocolate spread company.

References

1. David C Howell. *Statistical methods for psychology*. Cengage Learning, 2012.
2. Jim Webster. Nutella, imported vs. domestic: Is there a difference? https://www.washingtonpost.com/lifestyle/food/nutella-imported-vs-domestic-is-there-a-difference/2014/05/30/3fe79e68-e5bb-11e3-8f90-73e071f3d637_story.html?utm_term=.2e2fafd8c360.
3. Regina Nuzzo. Statistical errors. *Nature*, 506(7487):150, 2014.
4. Alessandro Rossini. The Chocolate-Hazelnut Spread Experiment. <https://alessandrorossini.org/chocolate-hazelnut-spread-experiment>.
5. Wolf Paulus. The Nutella Experiment. <https://wolfpaulus.com/life/nutella/>.