A Test in Discernment of Scotch Whiskey and Cognac Brandy Among Untrained Tasters

Adrian Röllin and Martine Zilversmit

Abstract

Scotch whiskey and Cognac brandy are among the most highly prized distilled spirits in Europe and the New World. Centuries of amateur and professional food science have been devoted to developing the ingredients and aging methods that give these beverages their highly prized flavours. Although each of these beverages inspires a devoted and often exclusive following, it has been recently proposed that an untrained individual would not be able to reliably discern these two drinks based only on taste and aroma. Here we examine this hypothesis with a highly sophisticated experimental design using statistical models and experimental procedures.

1 Introduction

The discovery of "Stone Age" beer jugs places the advent of alcoholic beverages at the Neolithic period, circa 10,000 BC. The technology of distillation can be dated to second century BC Babylonia, with the modern form of the process arising in the Middle East in the eighth century [5]. Still, it was not until the mid-12 century that this technology was used to make potable beverages, known as "elixirs" to experimental alchemists of the time [5]. Over the subsequent centuries, alcoholic spirits have evolved into an array of highly divergent and complex beverages, many with specific regional varieties. Two of the most popular forms of European spirits are whiskey (made from fermented grain mash [7]) and brandy (made from fermented fruit, often grapes [3]) with Scotch whiskey and Cognac brandy regarded as some of the finest forms of these liquors.

Although the essential botanical components of Scotch whiskey and Cognac brandy do not overlap, the process of finishing both these potables requires a long aging process in receptacles constructed of the aromatic wood oak [7, 4]. At this stage, the distillates are additionally flavored by the quercus lactone, 3-methyl-4-octanolide, also known as the whiskey lactone (Figure 1), contained in the wood [7]. As lactones are cyclic esters (ester molecules that contain a ring in the structure), and these organic compounds are known for their pungent aromas and flavouring abilities [6], it is not surprising that the quercus lactone contributes so significantly to the flavour of whiskeys and Cognac. It is perhaps this common element to the production of Scotch whiskey and Cognac brandy that lead to the recently purported supposition that an untrained non-specialist, even one with some recreational experience with these spirits, could not tell the difference purely based on taste and aroma [1].

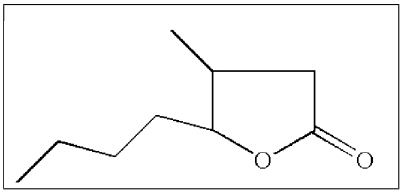

Figure 1: Molecular structure of the organic compound quercus lactone 3-methyl-4-octanolide

2 Material & Methods

We used a single-blind study to test the hypothesis that individuals are not able to reliably discern Scotch Whiskey and Cognac from each other. To avoid a direct comparison of the spirits (a setup where the classification is undoubtedly thought to be much easier) each subject obtained one of the 12 samples on 12 consecutive days. The sample set consisted of 4 different spirits[1] of which each was presented exactly 3 times, but in randomized order. As Cognac tends to have a darker and redder colour than whiskey, all samples where dispensed from identical opaque, bottle-shaped flasks to avoid providing any visual clues. Each sample consisted of approximately 0.75cl of liquid presented in a standard transparent tumbler of about 2dl volumetric capacity. Differences in color between any of the samples were not apparant at the low-volume used for sample size.

All the subjects where informed in advance that each type of spirit would be presented 6 times in total. This leads to difficulties in the statistical inference. Exact p-values can be calculated, or easily simulated, to test whether a specific individual is able to distinguish between the two classes of spirits (a value of p=0.04 for 10 correct answers out of 12 tries under the null hypothesis that an individual uses a randomized sequence of 6 times Cognac and 6 times Whiskey). However, we instead chose a log-linear binomial model as a simple approximation to estimate the effects, where the subjects are treated as random effects. The model fitting was performed using the `lmer`-procedure from the `lme4`-package which is part of the freely available statistical software `R` [8].

Due to the fact that each spirit was presented several times some taste and olfactory learning was expected, even for untrained individuals. Thus, we corrected and tested for this effect. Note that the days were consequently enumerated for each subject individually and missing data gaps were removed.

Other than sex, no other attributes were included in the model, as all of the individuals were drawn from a comparable socio-economic and educational cohort with an age between 25-35 years (with only one exception). Although some of the individuals declared to be well experienced with Whiskey [2], we did not correct for this effect, as some of the authors of the study doubted the reliability of the subjects' self-assessement.

Thus, the full model included the success/failure in each single test, the subject ID, the sex, the type of spirit tested and the sample number for each specific individual.

[1] Whiskies: The Dimple (blended, 15 yrs), Glenfiddich (single malt, 12 yrs); Cognacs: Remy Martin VSOP, Courvoisier VS

	TRUE SEQUENCE AND ANSWERS												CORRECT	RELATIVE
	C3	W2	W1	W1	C4	W2	C3	W2	C4	C4	C3	W1		
1	W	W	W	C	C	W	C	W	C	C	C	W	10	0.83
2	W	W	W	W	C	W	C	C	C	C	C	W	10	0.83
3	W	W	W	W	C	W	W	W	C	C	C	C	9	0.75
4	W		W	C	C	W	C	W	C	W	W		6	0.60
5		W				W	C	C		C	C	W	6	0.86
6		W	W	C	C	W							4	0.80
7		W	C	W									2	0.67
8					W	W	W	W				C	2	0.40
9		W	C										1	0.50
10		W											1	1.00
11									W	C			1	0.50
12											C		1	1.00
13	W			C		C	W			W			0	0.00

Table 1: Raw data matrix, ordered by total amount of correct answers where C=Cognac and W=Whiskey. Also given is the relative amount of correct answers. The numbers behind the true sequence are the actual spirits: 1=The Dimple, 2=Glenfiddich, 3=Remy Martin, 4=Courvoisier.

Model without learning effect	Estimate	SE	p
Overall success log-odds	0.778	0.322	0.016
Population SD of log-odds	0.633	–	–

Model with learning effect	Estimate	SE	p
Success log-odds at day 0	0.295	0.393	0.453
Learning effect per 1 day	0.143	0.087	0.103
Population SD of log-odds	0.416	–	–

Table 2: Estimates of the remaining main effects in the reduced log-linear models with and without correction for the learning effect.

3 Results

Neither sex nor the type of spirit showed a significant effect on the success probability ($p \gg 0.1$). Whereas in the full model the learning effect exhibits a clear tendency of an increase of the success rate ($0.05 < p < 0.1$), in the reduced model without sex and spirit type the effect is less apparent ($p = 0.103$). We thus considered both models, with and without the learning effect.

The estimates of the log-odd ratios are presented in Table 2. The model not allowing for a learning effect shows a significant overall ability of classifying the spirits with a success probability of 68%. Including the population variance, the individual abilities are to be expected to range from 35% up to 89% (not taking into account the additional estimation error).

The results in the model allowing now for a linear learning effect (on the log-odds scale) are clearly different. The overall success probability on the first day is only 57%, not significantly differing from the 50:50 random guessing, with the individual success probabilities ranging from 37% to 75%. However, a positive learning effect can be observed (see Figure 2), so that after 12 days the overall success rate is estimated at 86%, with an individual range from 74% to 94%. Thus most of the individuals seem to be able to distinguish between the two spirits after a learning phase of 6-12 days.

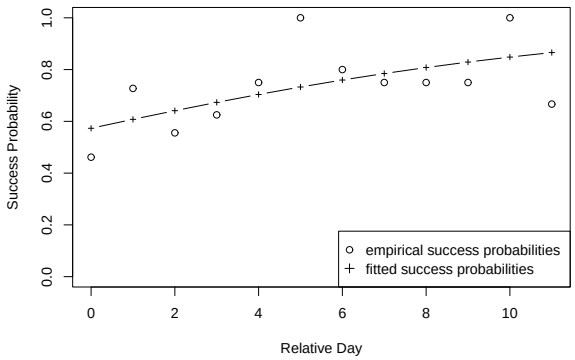

Figure 2: Empirical per-day success probabilities and fitted values from the model with learning effect.

4 Discussion

With our study we could not find conclusive evidence that untrained individuals are able to correctly classify Scotch Whiskey from Cognac. Although we find an overall success probability that is better than 50:50 if the model is not corrected for possible learning, our analysis shows that the short-term learning effect on the olfactory system can have an influence on the performance and should therefore not be neglected. If corrected for learning, the evidence of a better success rate than 50:50 vanishes in the first few days. Further research with a more refined experimental setup is therefore needed to address and circumvent the confounding influence of a possible learning effect.

Acknowledgments

We thank all the persons who participated in the experiment. We thank also Darwin's Cafeteria from the University of Oxford for providing us with the tumblers.

References

[1] A. Röllin (2007). Personal communication.

[2] M. Zilversmit (2007). Personal communication.

[3] Various contributors (2007). Brandy. *Wikipedia: The Free Encyplopedia*, 11 December 2007.

[4] Various contributors (2007). Cognac. *Wikipedia: The Free Encyplopedia*, 11 December 2007.

[5] Various contributors (2007). Distilled Beverages. *Wikipedia: The Free Encyplopedia*, 11 December 2007.

[6] Various contributors (2007). Esters. *Wikipedia: The Free Encyplopedia*, 11 December 2007.

[7] Various contributors (2007). Whisky. *Wikipedia: The Free Encyplopedia*, 11 December 2007.

[8] R Development Core Team (2007). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.