# Is triangle test more powerful than tetrad test in case of high alcoholic beverages?

- Tran Thi Hong Cam.\*,
- Nguyen Quang Phong,
- Pham Thanh Quang
- Nguyen Hoang Dzung

Ho Chi Minh city University of Technology, VNU-HCM, Vietnam

\*Email: trancambt@gmail.com, Tel: +84 909572714

(Manuscript Received on September 20<sup>th</sup>, 2014; Manuscript Revised December 5<sup>th</sup>, 2014)

### ABSTRACT

Tetrad test is theoretically more powerful than triangle test. However, in the presence of sensory fatigue-caused foods, the advantage of the tetrad test is lost. In case of alcohol beverage containing 50%abv, triangle test is recommended as a standard protocol because the sensory fatigue caused by ethanol is great enough that leads tetrad test be less powerful than triangle test.

In this study, the association between alcohol concentration and power of tetrad as well as triangle was established in cases of alcoholic beverages lower than 50%abv. Based on this result, the alcohol concentration at which the tetrad test is more powerful than triangle test was determined.

The power comparisons between triangle test and tetrad test were extended to some alcohol beverages containing 40 and 30%abv. 240 panelists were divided into 8 panels. In each session, both triangle test and tetrad test were performed by a same panel (N = 30) on the samples made from the same alcohol beverage basic with and without citric acid. There were

product four sessions for each that corresponded with four increasingly levels of added citric acid (2.192, 3.1, 4.384, 6.2g/L). By applying logistic regression, the power comparison between two tests was based on the correct response proportions (Pc) which were predicted from two variables: acid concentration and protocol.

Multivariate logistic regression analysis indicated that there was 84% chance and 38% chance for the tetrad's Pc to be higher than triangle's Pc in case of beverage containing 30 and 40%abv, respectively. The results shown that increase alcohol concentration will reduce the power of both triangle and tetrad tests. It was noteworthy that tetrad test was affected more strongly by alcohol than triangle test.

These findings suggest that the triangle test can be used as a suitable testing methodology for alcoholic beverages which alcohol concentration is not lower than 40%abv. In case of beverage containing 30%abv, triangle test should be replaced by tetrad test due to the increased reliability of information.

Keywords: power, triangle test, tetrad test, logistic regression, high alcoholic beverage.

#### **1. INTRODUCTION**

Power analysis can evaluate the efficiency of a discrimination test. The more powerful a discrimination test, the less likely it is to lead to a type II error. In foods and consumer products industries, making a type II error could lead to miss a positive significant change in formula modification or in processing, or fail to detect a negative sensation which could lead to consumers rejecting purchase or consuming products. It is therefore critical to choose a discrimination test which minimizes type II error (i.e., that is higher in power). In fact, the power of a discrimination test is considered as the most important issue to be considered when choosing methods [1].

There is a large body of work evaluating discrimination testing methods by comparing their power. For example, Masuoka, et al. [2] compared two pairs of tests on performed with beers and concluded that the 3-AFC test was more powerful than the triangle test but could find differences in power between the specified tetrad or the unspecified tetrad. In a study with basic yogurt, the triangle test was found to be less powerful than the same-different test by Rousseau, et al. [3]. In a more comprehensive study, Bi [4], Bi and Ennis [1] ranked discrimination tests in order of increasing statistical power as follow: duo-trio, triangle, A-not A, 2-AFC, 3-AFC. In general, specified tests are more powerful than unspecified tests [5]. This suggests that specified tests are preferable to unspecified tests when statistical power is concerned but are less applied in food evaluation practice because food is considered a complex and multi-dimensional system in fact [6].

To enhance the power of these unspecified tests without increasing the number of panelists, some modified discrimination tests were suggested: the modified triangle test, the degree of difference test, the double discrimination test. The replicated discrimination test was also concerned as a solution to get sufficiently high testing power when the number of available panelists is not large enough [4].

In a different approach, researchers tried to find sensitive testing methods through comparing the sensitivity of various testing methodologies [7, 8]. Several notable papers have focused on the power comparison between tetrad and other methods. For example, Ennis and Jesionka [9] calculated the power as a function of sample size at different effect size (d') to compare the power of tetrad, triangle, duo-trio and 2-AFC. In the work reported by Garcia, et al. [10] the specified tetrad test revealed a larger difference between the stimuli than the 2-AFC test in case of large sample size. In a study with apple juice, Garcia, et al. [7] and Ennis [8] concluded that the power of the tetrad test was still higher than the power of the triangle test even though the effect size of tetrad was smaller. In another study with apple juice and orange juice product categories, Ishii, et al. [11] reported that the tetrad test could have higher power than the triangle test for small effect sizes and for some resampling conditions. The scientists recommended a switch from triangle test to tetrad test because the tetrad test is therefore superior to the triangle test, with lower sample size requirements, higher power and greater sensitivity [8].

However, the result of Ennis [8] investigating effects of four types of salsa on the power of the tetrad test and the triangle test shown that the tetrad test performed well for the mild and medium salsa, but poorly for the hot and very spicy varieties. This indicates that the relationship between the powers of triangle and tetrad test may not be the same for all kinds of product but depends on the levels of sensory fatigue. The lower level of sensory fatigue is, the higher chance that the tetrad test to be more powerful than the triangle test is.

In a study with 50% abv beverage, the result shown that this chance was low (12%). It is recommend that the triangle test should be chosen in case of beverages containing 50% abv [12]. Our purpose was to gauge whether or not the triangle test is a suitable test for beverages containing lower than 50% abv. In this study, some Vietnamese alcoholic beverages whose alcohol concentrations are lower than 50% abv were concerned. We assumed that the lower alcohol concentration is, the smaller sensory fatigue caused by alcohol is. When the sensory fatigue caused by ethanol is low enough, the tetrad is more powerful than triangle test. The power comparison between tetrad test and triangle test was carried out to point out which protocol, triangle or tetrad, is suitable for each kind of product.

#### 2. MATERIALS AND METHODS

#### 2.1 Stimuli

Two commercial alcoholic beverages were obtained from the local supermarkets in Ho Chi Minh city, Vietnam: the first one is Bau Da 40% abv (Tam Huong company Inc., Binh Dinh province, Vietnam), the second one is Bach Ma 30% abv (Hoa Viet company Inc., Ho Chi Minh city, Vietnam).

Each of product, the stimuli for the experiment consisted of a basic alcoholic beverage and the same alcoholic beverage with added commercial citric acid (Hoa Nam Inc., Ho Chi Minh city, Vietnam) at different concentrations: 2.192, 3.1, 4.384, 6.2 g/L. The first stimulus will be referred to as sample A, and the series with citric acid as samples Bx.

Alcoholic beverage bottles were chilled until their temperature was below 5°C before being mixed into the homogenous sample. This sample was divided into two parts: one for sample A and the other for samples B. Citric acid was added into the Bx samples to get the expected concentrations. Both sample A and samples B were stirred until citric acid in samples B was dissolved completely. The desired samples were poured in plastic bottles and kept in cold water (approx.  $1-5^{\circ}$ C).

Before starting an experimental session, 10ml samples were dispensed in plastic-lidded cups and kept in fridge for at least 5 minutes to ensure that all samples had the same temperature (approx. 8-10°C). When panelists evaluated samples, the temperature of the samples was about 12-15°C.

#### 2.2 Subjects

240 participants from Ho Chi Minh city University of Technology and Industrial University of Ho Chi Minh city participated in this study. They were 18 to 35 years old, willing to consume alcohol, not suffering from periodontal disease and not be allergic to any of the ingredients in the product. This information was collected by questionnaires before conducting the experiment.

#### 2.3 Testing Procedure

Each kind of product, we carried out 4 experimental sessions corresponding to 4 citric acid concentrations added to samples Bx. There were 8 experimental sessions in total for two kinds of alcoholic beverages.

In an experimental session, each panel (N=30) was performed only one time with both tetrad and triangle tests. Half the panel (15 panelists) started with the tetrad test, while the other half started with the triangle test. Each of the six possible presentation orders of both the triangle test and tetrad test was presented 5 times.

The panelists were explained the instructions, rinsed their mouth three times with distilled water, tasted the samples from left to right, swallowed the whole 10 mL of each sample, and finally gave their answer upon tasting the last sample. For the triangle test, the instructions were to select the odd sample which was different from the others [13]. For the tetrad test, panelists were asked to divide the four samples into two groups of two based on similarity [2]. Before continuing the second test, panelists had a 10 minute rest to reduce the effect of sensory-fatigue. The testing time for the each test lasted approximately 10 minutes on average and 30 minutes for the complete experiment. The same procedure was repeated with another panelists for other citric acid concentrations.

#### 2.4 Statistical Analysis

Chi square test was used to compare Pc-values.

Models developed by regression analysis allow us to observe simultaneously the influence of both protocol and acid concentration variables on the Pc[14]. Therefore regression analysis was used in our study to give a more accurate picture of the power relationship between protocol and acid concentration variables.

A multivariate logistic regression analysis was carried out to test the association between independent variables and Pc with and without interaction. The independent variables tested in the model were acid concentration and protocol (triangle test or tetrad test).

The statistical software R (version 3.1.0) was used for all statistical analyses. All reported p-values were two-tailed, and p-values lower than .05 were considered statistically significant.

#### **3. RESULTS**

Results of 8 experimental sessions are showed in Figure 1. In general, the Pc values of two protocols increase when adding more citric acid in the same product while they decrease when increase alcohol at the same added citric concentration. In case of beverage containing 40% abv, there were two sessions that triangle's Pc values were greater than the tetrad's. It was noteworthy that the tetrad's Pc values were not lower than triangle's in case of beverage containing 30% abv. However, in each session, p-values (using Chi square test) greater than .05 indicated that there was no significant difference between triangle's and tetrad's Pc values that were used to comparing power of protocols.



Figure 1. The correct response proportion (Pc) of triangle test and tetrad test in 8 experiment sessions

Using the *Pc*-ratio comparison approach [12], the ratio of tetrad's *Pc* to triangle's *Pc* can be estimated and compared to 1. The confidence levels of the ratios which were greater than 1 were shown in Table 1 and Table 2.

In case of beverage containing 40% abv, we had approximately 50% chances that the triangle test returned greater numbers of correct responses than the tetrad test (the chance values ranged from 30.2% to 63.1%). In addition, Chi square test shown that the chance that *Pc* values of two tests in each session

were not significant different were higher than 60%. A trend was drawn that the power of triangle and tetrad tests were almost equivalent in this kind of product.

In case of beverage containing 30% abv, we have more than 50% chances that the tetrad test returned greater numbers of correct responses than the triangle test (the chance values ranged from 50.0% to 80.6%). A trend was drawn that the power of tetrad test was greater than the one of triangle test in this kind of product

 Table 1. The ratios of the tetrad's Pc to the triangle's and the confidence levels that ratio is greater than 1 for each stimuli's concentration in case of beverage containing 40%abv

Citric acid concentration (g/l)	Protocol	Pc	Var	p-value (Chi square test)	P <sub>c.tetrad</sub> P <sub>c.triangle</sub>	Confidence that Pc ratio is greater than 1 (%)	
2.192	Triangle	0.467	0.091	1	1 000	50.0	
	Tetrad	0.467	0.091	1	1.000		
3.100	Triangle	0.567	0.090	0.605	0.000	30.2	
	Tetrad	0.500	0.091	0.005	0.882		
4.384	Triangle	0.567	0.090	0.705	0.041	39.8	
	Tetrad	0.533	0.091	0.795	0.941		
6.200	Triangle	0.800	0.073	0.720	1.042	63.1	
	Tetrad	0.833	0.068	0.739	1.042		

 Table 2. The ratios of the tetrad's Pc to the triangle's and the confidence levels that ratio is greater than 1 for each stimuli's concentration in case of beverage containing 30%abv

Citric acid concentration (g/l)	Protocol	Рс	Var	p-value (Chi square test)	P <sub>c.tetrad</sub> P <sub>c.triangle</sub>	Confidence that Pc ratio is greater than 1 (%)	
2.192	Triangle	0.500	0.091	1	1.000	50.0	
	Tetrad	0.500	0.091	1	1.000		
3.100	Triangle	0.567	0.090	0.508	1 110	70.1	
	Tetrad	0.633	0.088	0.398	1.110		
4.384	Triangle	0.667	0.086	0.300	1 150	80.6	
	Tetrad	0.767	0.077	0.390	1.150		
6.200	Triangle	0.833	0.068	0.448	1.080	77.7	
	Tetrad	0.900	0.055	0.448	1.000		

P-value\*: is our measure of statistical significance and will tell us whether it is likely that we would have found a relationship of this size in the sample if there was no relationship in the population.

P-value\*\*: used for evaluation the goodness of fit of used model. When P-value is higher than 0.05 which suggests that our model fits the data.

#### **Trang 36**

Logistic regression was used to determine which variables related to *Pc* values and also to estimate the magnitude of the overall effect of the explanatory variables on the outcome of our study. Table 3 and Table 4 shows the  $\beta$  coefficients and involving statistics in the logistic regression models for beverages containing 40% abv and 30% abv, respectively. In model 1, predictive variables are acid concentration, protocol, and interaction between acid concentration and protocol; while acid concentration and protocol are in model 2.

In both of beverage containing 40% abv and 30% abv, the interaction between citric acid concentrations and testing protocols was not found significantly. Thus, model 2 of two products whose predictive variables are acid concentration, protocol was used for result analysis.

In both of beverage containing 40% abv and 30% abv, the  $\beta$  coefficient of acid concentration variable was positive with the 95% confidence interval (Table 5). This positive value indicates that the *Pc* values increased when the added citric acid

concentration increased. In case of beverage containing 40%abv, the mean of  $\beta$  coefficient of protocol variable was negative while positive in beverage containing 30%abv. This positive value indicates that the *Pc* when the switching from the triangle test to the tetrad test or vice versa. Table 5 provides the  $\beta$  coefficients and the 95% confidence intervals of  $\beta$  coefficients.

The 95% confidence interval of  $\beta$  coefficient of protocol ranges from -0.608 to 0.459 in case of beverage containing 40% abv, while it ranges from - 0.274 to 0.858 in case of beverage containing 30% abv. At the 62<sup>nd</sup> percentiles, the  $\beta$  coefficient of protocol was positive in case of beverage containing 40% abv. Because the tetrad was coded as 1, the positive  $\beta$  coefficient of protocol meant that the tetrad test returned greater *Pc* than the triangle test. In other words, there was a 38% chance that the tetrad's power could be greater than the triangle's power in case of beverage containing 40% abv. This chance was 84% in case of beverage containing 30% abv (Figure 3).



Figure 2. The logistic regression model without interaction (model 2) in case of beverage containing 40% abv and 30% abv

	Beverage 40%abv			Beverage 30% abv			
	Mean	Lower	Upper	Mean	Lower	Upper	
Acid concentration	0.368	0.186	0.559	0.46	0.259	0.676	
Protocol	-0.074	-0.608	0.459	0.29	-0.274	0.858	

Table 3. The 95% confidence intervals of the  $\beta$  coefficients in model 2 in case of beverage containing 40% and 30% abv

#### β- coefficients of Acid concentration



Figure 3. β coefficients of acid concentration and protocol variables in case of beverage containing 40% abv and 30% abv (simulated by using R program with 1000 replication)

#### 4. DISCUSSIONS

In this research, the powers of the triangle and tetrad test were compared through Pc values. According to the results shown in Figure 3, the differences of Pc values of two protocols were not significant in all of products. Thus, tetrad's and triangle's power is non-significant difference. However, when calculating the confidence levels of Pc ratios in all of products, there is an increate trend in chance that the triangle's Pc was higher than the tetrad's.

By logistic regressive analysis, the results showed that there was no interaction between citric acid concentrations and testing protocols. The association between triangle's Pc and tetrad's was therefore not influenced by the acid citric concentration.

β- coefficients of Protocol

A result of analyzing  $\beta$  coefficient show that there are 38% and 84% chance that the tetrad's power could be greater than the tetrad's power in case of beverage with 40% alcohol and 30% alcohol, repetitively. Combining this result with result of Tran in a study with beverage containing 50% abv [12] that this chance is 12%, we realized that *the lower alcohol concentration is, the higher this chance is.* Moreover, this chance decrease when increase alcohol concentration at the same citric acid concentration in both tetrad and triangle test. Tetrad's power variability is higher than triangle's.

Alcohol elicits trigeminal sensations whose characteristics are numbing, burning, tingling, and irritation [15]. Alcoholic beverage can easily cause sensory fatigue [8], when alcohol concentration increate lead to increase the variance of perceptual distribution. This in turn would decrease power in each of tests. The effect of alcohol on power of tetrad that there are the 4<sup>th</sup> sample to evaluate, is larger than triangle test. This is reason why chance that the triangle's power could be greater than the tetrad's changes towards alcohol concentration.

This results is supported by the research of Ennis [8] who used salsa product that had the different spiced levels. Both alcohol drinks and salsa are food products causing sensory fatigue. This suggests that sensory properties of food products must be taken into account when comparing triangle and tetrad tests with the intent of reducing type II error. With this in mind, a tetrad procedure is not recommended in case of beverage that have alcohol concentration is not lower than 40%abv. Even if the tetrad's power is equivalent to triangle's, the triangle test is also recommended because of a simple protocol with less sample.

In this study, the alcoholic beverages are almost pure (water and ethanol are the major ingredients). Some other ingredients such as  $CO_2$  (in beer, sparkling wine), tannin (in red wine) also elicit trigeminal sensations [15]. They maybe affect to the power of triangle and tetrad tests. Therefore, further studies are needed to employ other beverages with  $CO_2$  or tannin content to have a guideline for sensory-field practitioners who are considering a switching from the triangle test to the tetrad test.

#### **5. CONCLUSION**

Power of the tetrad and triangle tests were compared at different alcohol concentrations. Specifically, when increasing alcohol concentration, the chance of the tetrad test to be more powerful than the triangle test decrease. It is possible that our results reflect sensory fatigue caused by high alcoholic beverages. The evaluation of the fourth sample in tetrad protocol increased the variance of perceptual distribution that led to decrease its power advantage. This suggests that tetrad test is not suitable for all kinds of alcoholic beverages. The triangle test is recommended as a standard protocol in cases of beverages containing not lower than 40% abv. In case of 30% abv beverages, the tetrad should be chosen to minimize type II error. Because the sensory fatigue caused by alcohol in this case is small enough that the tetrad test is more powerful than the triangle test.

# Phải chăng năng lực của phép thử tam giác luôn cao hơn tetrad trên những đồ uống có độ cồn cao?

- Trần Thị Hồng Cẩm\*,
- Nguyễn Quang Phong,
- Phạm Thanh Quang
- Nguyễn Hoàng Dũng

Trường Đại học Bách khoa, ĐHQG-HCM

## TÓM TẮT

Trên lý thuyết, năng lực của phép thử tetrad được đánh giá cao hơn phép thử tam giác. Tuy 77nhiên, đối với những dạng thực phẩm dễ gây cảm giác mệt mỏi như thức uống có cồn, năng lực của phép thử tetrad có thể bị giảm đi. Những nghiên cứu trước đây đề nghị sử dụng phép thử tam giác thay thế phép thử tetrad đối với những sản phẩm đồ uống có nồng độ cồn 50% v/v.

Trong nghiên cứu này, chúng tôi thiết lập mối quan hệ giữa nồng độ cồn và năng lực của hai phép thử tetrad và tam giác cho những sản phẩm đồ uống có nồng độ cồn thấp hơn 50% v/v. Nghiên cứu kỳ vọng sẽ xác định được nồng độ cồn mà tại đó phép thử tetrad có năng lực cao 83hơn phép thử tam giác.

Việc so sánh năng lực của hai phép thử được được nghiên cứu trên những sản phẩm có nồng độ cồn 40% và 30% v/v. Người thử (n=240) được sắp xếp ngẫu nhiên vào 8 hội đồng, 30 người cho một hội đồng. Ở mỗi đợt thí nghiệm, mỗi hội đồng thực hiện cả hai phép thử trên hai loại mẫu được chuẩn bị từ cùng một loại đồ uống; trong đó, một loại mẫu không bổ sung và loại kia có bổ sung acid citric. Với mỗi loại đồ uống có cồn, 4 đợt thí nghiệm được tiến hành tương ứng với 4 nồng độ acid citric bổ sung tăng dần theo tỉ lệ 2.192, 3.1, 4.384, và 6.2g/L. Việc so sánh năng lực của hai phép thử dựa trên các giá trị tỷ lệ câu trả đúng (Pc) được ước lượng từ mô hình hồi quy logistic thông qua hai biến: nồng độ acid và loại phép thử.

Kết quả chỉ ra rằng: có 84% và 38% cơ hội, tương ứng với sản phẩm có nồng độ cồn 30% và 40% v/v, để giá trị Pc của phép thử tetrad cao hơn Pc của phép thử tam giác. Kết quả cũng cho thấy khi nồng độ cồn tăng, năng lực của cả hai phép thử đều giảm, và năng lực phép thử tetrad giảm mạnh hơn năng lực của phép thử tam giác.

Với những kết quả thu được, chúng tôi đề nghị: để tăng độ tin cậy của thông tin thu được từ phép thử trên đối tượng nghiên cứu là đồ uống có cồn, phép thử tam giác nên được sử dụng với những sản phẩm có nồng độ cồn lớn hơn hoặc bằng 40% v/v, đối với những sản phẩm có nồng độ cồn 30% v/v, phép thử tam giác nên được thay thế bằng phép thử tetrad

#### REFERENCE

- J. Bi and D. M. Ennis, "The Power Of The "A"-"Not A" Method," Journal of sensory studies, vol. 16, pp. 343-359, 2001.
- [2] S. Masuoka, D. Hatjopoulos, and M. O'mahony, "Beer bitterness detection: Testing Thurstonian and sequential sensitivity analysis models for triad and tetrad methods," Journal of Sensory Studies, vol. 10, pp. 295-306, 1995.
- [3] B. Rousseau, A. Meyer, and M. O'mahony, "Power And Sensitivity Of The Same-Different Test: Comparison With Triangle And Duo-Trio Methods," Journal of Sensory Studies, vol. 13, pp. 149-173, 1998.
- [4] J. Bi, Sensory discrimination tests and measurements: Statistical principles, procedures and tables: John Wiley & Sons, 2008.
- [5] D. Van Hout, M. J. Hautus, and H. S. Lee, "Investigation Of Test Performance Over Repeated Sessions Using Signal Detection Theory: Comparison Of Three Nonattribute-Specified Difference Tests 2-AFCR, A-Not A And 2-AFC," Journal of Sensory Studies, vol. 26, pp. 311-321, 2011.
- [6] M. C. Meilgaard, B. T. Carr, and G. V. Civille, Sensory evaluation techniques: CRC press, 2006.
- [7] K. Garcia, J. M. Ennis, and W. Prinyawiwatkul, "A Large-Scale Experimental Comparison Of The Tetrad And Triangle Tests

In Children," Journal of Sensory Studies, vol. 27, pp. 217-222, 2012.

- [8] J. M. Ennis, "Guiding the switch from triangle testing to tetrad testing," Journal of Sensory Studies, vol. 27, pp. 223-231, 2012.
- [9] J. M. Ennis and V. Jesionka, "The power of sensory discrimination methods revisited," Journal of Sensory Studies, vol. 26, pp. 371-382, 2011.
- [10] K. Garcia, J. M. Ennis, and W. Prinyawiwatkul, "Reconsidering the specified Tetrad test," Journal of Sensory Studies, vol. 28, pp. 445-449, 2013.
- [11] R. Ishii, M. O'Mahony, and B. Rousseau, "Triangle and tetrad protocols: Small sensory differences, resampling and consumer relevance," Food Quality and Preference, vol. 31, pp. 49-55, 2014.
- [12] C. H. T. Tran et al, "Power comparison of triangle and tetrad tests applied to high ethanol beverages", SPISE Proceedings 4 (2014), pp. 81-87.
- [13] ASTM, "E1885 04(2011) Standard Test Method for Sensory Analysis - Triangle Test," ed, 2011.
- [14] D. W. Hosmer Jr and S. Lemeshow, Applied logistic regression: John Wiley & Sons, 2004.
- [15] H. Lawless and H. Heymann, "Sensory Evaluation of Food: Practices and Principals," ed: Chapman and Hall: New York, NY, USA, 1988.