"Colour training and colour differences thresholds in orange juice".

[1]Rocío Fernández-Vázquez, [1]Carla M. Stinco, [2]Dolores Hernanz, [1]Francisco J. Heredia, [1]Isabel M. Vicario

[1]Food Colour & Quality Laboratory, Department of Nutrition & Food Science. Universidad de Sevilla. Facultad de Farmacia, 41012 Sevilla, Spain. www.color.us.es

[2]Department Analytical Chemistry. Universidad de Sevilla. Facultad de Farmacia, 41012 Sevilla, Spain.

* Corresponding author:

Isabel M. Vicario

Food Colour & Quality Lab., Dept. Nutrition & Food Science. Facultad de Farmacia.
Universidad de Sevilla. 41012-Sevilla, Spain

Tel.: +34 954556339 Fax: +34 954557017

Email: vicario@us.es

Abstract

This study was aimed at training a panel of assessors to evaluate specifically orange juice colour, and to establish the colour difference threshold in orange juice for a trained and untrained panel. Panellists were first preselected using Farnsworth-Munsell 100-Hue Test and then trained with a specific method for orange juice colour. This training allowed assessors to evaluate visually orange juice samples in hue and intensity. The final selection of assessors was a panel of 8 trained observers with reproducibility and repeatability, and a significant discrimination among the samples *($p<0.05$).*

On the other hand, commercial orange juices were evaluated both instrumentally by image analysis and visually by the trained panel, and the untrained panel. Instrumental colour measurements and visual evaluation were correlated. Values around 1.5 and 2.8 CIELAB units could be consider the threshold for colour differences between two orange juices for the trained and untrained panel, respectively.

Keywords: Colour, orange juice, sensory training, colour differences

1. **Introduction**

Colour is one of the most important visual attributes in food and usually is the first one evaluated by consumers and is associated to the concept of quality (Huggart, Petrus, & Buzz Lig, 1977; Tepper, 1993; Pangborn, 1960). In orange juices, the natural bright colour is considered one of their main advantages over other juices (Barron, Maraulja, & Huggart, 1967) and has attached great importance since some studies have probed that it may influence flavour perception and other quality attributes (Tepper, 1993; Fernández-Vázquez et al., 2012).

Colour can be evaluated by instrumental or visual analysis. Humans and instruments measure colour in different ways. Human perception of colour is based on responses of photoreceptors in the retina of the eye and the way they are interpreted within the brain. These perceived colours are often characterised by physical scientists using three dimensions: lightness, hue and chroma. Instruments, on the other hand, are capable of seeing pure values of the colorimetric coordinates CIELAB L*, a*, and b*. Nowadays, there are new advances in image acquisition technology that offer the possibility of using technically sophisticated apparatus available at relatively low cost to evaluate colour in terms of millions of pixels. In comparison with the traditional light sensors, the main advantage is that they allow making a detailed evaluation of a wider area of any food product, with inhomogeneous colour possible. Every different colour in the image of the analyzed food matrix can be accounted for by one or more pixels (Antonelli et al., 2004). Furthermore, it is based upon digital cameras, which can quickly capture images in digital format (DigiEye®) (Luo, Cui, & Li, 2001) and offers a more reliable measurement of the food colour, which can be correlated with sensory analysis and other colour measurements (Fernández-Vázquez, Stinco, Melendez-Martinez, Heredia, & Vicario I.M, 2011) .

Anyway, colour measurement usually requires instruments that are not always available in small and medium size companies and visual evaluation could be an alternative. Human colour vision is a quite complex process and colour is undoubtedly a perception, a virtual property of the material. In order to use the visual analysis as an objective quality control, it is necessary to standardize the measurement conditions to be able to compare with the instrumental measurement. Previous studies have shown that a good correlation can be achieved when the instrumental and sensory measurements are done considering different aspect such as background, surround or illumination (Fernández-Vázquez, Stinco, Melendez-Martínez, Heredia, & Vicario I.M, 2011; Meléndez-Martínez, Vicario, & Heredia, 2005).

Although colour evaluation is included in many sensory studies (Poelman & Delahunty, 2011; Calvo, Salvador, & Fiszman, 2001; Frata, Valim, & Monteiro, 2006), there are very few studies which specially train the panellists to do the visual evaluation of food with more details. An example of an specific training in visual evaluation was done by Gambaro *et al*. (2001) for strawberry yogurt. Based on this experience, we have particularly trained a panel to evaluate orange juice colour in a reproducible and repeatable way.

On the other hand, the evaluation of colour differences has had a high interest for long time. Specifically, 'just noticeable differences' have been very important in the development of the Colorimetry. There are equations to find out the colour difference between two stimuli in the CIELAB space, which are mathematical expressions which allows us to obtain the number ΔE^*_{ab}. This is the positive number which stays invariable when the products are exchanged (Melgosa, Pérez, Yebra, Huertas, & Hita, 2001). At the present time, calculation of colour differences has many applications in Colorimetry, such as the reproducibility of colour in manufactured products and communication systems, or in the study of the colour fading in food, works of art, etc. or more recently to determine colour tolerance in orange juice (Wei, Ou, Lou & Hutchings, 2012).

One of the key problems in the visual evaluation is establishing the threshold for colour differences. Previous studies have explored colour threshold using colour standards (Berns, Alman, Reniff, Snyder, & Balonon-Rosen, 1991). An attempt to stablish the thresholds for visual discrimation between wines was also published by Martínez *et al.* (2001). Recently Wei *et al*. (2012) established the colour of an ideal orange juice and the colour tolerance, using a digital display. However, so far, literature on orange juice colour does not provide data on the colour differences that can be visually detected between two orange juices (based on real samples) by consumers, although this type of information could be very useful for the orange juice industry.

The objectives of this study were: (1) to train a panel of assessors to evaluate specifically the orange juice colour, and (2) to study the visually perceived colour difference by the observers' panel (trained and untrained) in a complete range of orange juices of different colours to establish the colour difference threshold in this popular beverage.

2. Material and methods

2.1. Instrumental colour measurement

The DigiEye imaging system was used to capture the digital images (Luo, Cui, & Li, 2001). The latter system includes a calibrated digital camera 10.2-megapixel Nikon D80 (Nikon Corporation, Tokyo, Japan) and an objective Nikkor 35-mm f/2D (Nikon Corporation), a colour sensor for display calibration, and an illumination box designed by VeriVide Ltd. (Leicester, UK) (Figure 1). For objective colour specifications, the samples were placed in 75 mL capacity transparent plastic bottles (Figure 2) and measured against a grey surround ($L^* = 50$) and white background. Digital images were made in order to obtain the total appearance of juice at depths observed by consumers. In these measurements, the samples were illuminated by a diffuse D65 simulator. For obtaining CIELAB coordinates, we used the DigiFood software (Heredia, González-Miret, Álvarez, & Ramírez, 2006), which allows the transformation of RGB values into the CIELAB colour parameters, based on computational solutions (León, Mery, Pedreschi & León, 2006).

From the CIELAB uniform colour space, the psychophysical parameters chroma (C^*_{ab}) and hue (h_{ab}) are defined as:

$$C^*_{ab} = \sqrt{(a^*)^2 + (b^*)^2}, \quad h_{ab} = \arctan(b^*/a^*)$$

Chroma (C^*_{ab}) is used to determine the degree of difference of a hue in comparison with a grey colour with the same lightness, and is considered the quantitative attribute of colourfulness. Hue (h_{ab}) is the attribute according to which colours are usually defined as reddish, greenish, etc. and is used to define the difference of a colour with reference to a grey colour with the same lightness. This attribute is related to the differences in reflectance at different wavelengths and is considered the qualitative attribute of colour.

Colour differences (ΔE^*_{ab}) were calculated as the Euclidean distance between two points in the 3-D space defined by L^*, a^* and b^*:

$$\Delta E^*_{ab} = \sqrt{(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2}$$

2.2. Colour training

Briefly, a protocol was designed for the selection and training of assessors based on the methodology proposed by Gambaro *et al*. (2001). First, a preliminary panellist selection was made using the Farnsworth-Munsell 100-Hue Test. Then, those panellists who did not present a good skill to discriminate light differences in tone and intensity using blended of colouring dilutions were rejected. Finally, the selected panel was training using two commercial samples and a serie of dilutions of one of them.

2.2.1. Panel Selection

A panel of 12 assessors were recruited from students and staff at the University of Seville and then preselected according to their normal colour vision following ISO 11037 (1999). The Farnsworth-Munsell 100-Hue Test for the examination of Colour Discrimination (Farnsworth 1957) was used to verify the normal vision. It allows to separate persons with normal colour vision into classes of superior, average and low colour discrimination, and to measure the zones of colour confusion in colour defective persons.

To determine the ability to discriminate among slight tone and intensity differences in orange colour, aqueous orange-coloured solutions were prepared using two food dyes (red and yellow food dyes from McCormick, Spain S.A) blended in different proportions: 500 and 12000 μL/L of red and yellow food dye respectively, for yellowish dilutions; 1250 and 12000 μL/L of red and yellow food dye for orangish dilutions; and 3000 and 12000 μL/L of red and yellow food dye for reddish dilutions. Ten aqueous solutions of each blend of colourings were prepared (100, 78, 47, 36, 29, 22, 17, 14, 10, and 5 %). These solutions (75 mL) were placed in bottles of transparent plastic and coded with 2 digit random numbers.

The test was carried out using a VeriVide CAC Portable cabinet (dimension of viewing area: 635 mm width, 280 mm height, and 280 mm depth) to control illumination and observation conditions. D65 was used as source of illumination (the same used in the instrumental measurements) (CIE, 2007), a white background and a grey surround were selected to simulate the objective measurements made by image analyses (Stinco et al., 2012).

The evaluation sessions were organized as follows: in a first stage (hue classification) assessors dealt with the samples (n=3) corresponding to the each dilution level separately and were asked to sort them in yellowish, orangish, and reddish hues (10 evaluations). In a second stage (intensity ranking), the assessors were given the whole dilution series (yellowish, orangish and reddish) separately, and they were asked to rank them according to the increasing intensity.

To select the assessors, the criteria used were: (1) reject those who were unable to accomplish sorting the tubes into the three tone groups and (2) reject those whose Spearman's ranked correlation coefficients *(p)* of sensory ranking versus colour concentrations were not significant ($p > 0.05$) (O'Mahony, 1986).

2.2.2. Panel training

Two nonstructured 10 cm long scales, anchored at the end, were used to train the assessors in colour evaluation. The colour attributes trained were hue and intensity. In this study, chroma and lightness were not considered separately as individual attributes because in previous studies it was observed that panellists had difficulties to understand and evaluate chroma (Fernández-Vázquez, Stinco, Melendez-Martínez, Heredia, & Vicario I.M, 2011). For this reason, intensity was assayed as the best way to evaluate both parameters visually. Hue was evaluated from yellowish to reddish and intensity was evaluated from low to high.

A collection of samples were selected to encompass the full range of colour intensity and hue in commercial orange juice. Two commercials orange juice samples (COJ I and COJ II), and six samples prepared from dilutions of COJ II (6, 10, 30, 50, 60, and 80 %) were used. These samples were evaluated in 15 min sessions, and at the end of each session a meeting of 30 minutes was done by the leader of the panel and all the panellists to unify the criteria of evaluation. In each session, assessors evaluated duplicates of the commercial samples and a couple of diluted samples.

The samples evaluated in each session were the same for all assessors, being the order randomized across assessors. This design allowed training of the assessors in evaluating orange colour intensity and hue as well as determining their reproducibility and performance consistency. A three factor ANOVA (assessor, sample, and repetition) for samples COJ I and COJ II, and a two factor ANOVA (assessor and sample) for these two samples and all the dilutions were performed on the data obtained (O'Mahony 1986). According to the results obtained, those assessors with the highest variance and greater judgement dispersion were withdrawn from the panel.

2.3. Colour differences thresholds

2.3.1. Orange juices samples

16 commercial orange juices (5 from concentrated, 6 from squeezed oranges and stored at 4ºC, and 5 from squeezed oranges and stored at room temperature) were purchased from different supermarkets in Spain. These samples were chosen in order to collect the variety of the orange juices colour available in the supermarket. Each sample was placed in 75 mL capacity transparent plastic bottles to measure its colour by image analyses and then to evaluate the colour differences.

2.3.2. Sensory evaluation

The samples were compared by pairs (120 comparisons) by the trained panel of 8 observers with normal colour vision and previously trained in colour discrimination experiments.

Afterwards, eight panellists recruited also from students and staff at the University of Seville, with normal vision (according to the Farnsworth-Munsell 100-Hue Test) but no previous knowledge in colour science, repeated the experiment with the aim of establishing the colour difference threshold for untrained observers.

120 pairs of samples were displayed on the centre of the VeriVide CAC 120 cabinet. Observers were situated 50 cm in front of the samples, with white background and grey surround.

The task of each observer was to judge whether they could notice the colour difference between the two orange juice samples (1) or if they could not (0). They did the test twice per pair of samples in two different sessions: one with the couple of samples side by side (experiment a); and another one with the samples separated by 15 cm (experiment b). In this way, each observer in the trained panel (8 panellists) evaluated the 120 pairs of samples once (960 judgments in total), similarly did the group of 8 untrained observers (1200 judgments in total). The final results were expressed as "visual colour difference" (ΔV) which was calculated as the percentage of positive panellists' responses (1).

The visual judgments were made immediately after the instrumental colour measurements in order to avoid the colour variation of the samples.

The correlation between the visually perceived and instrumentally measured colour differences were explored following procedures previously described elsewhere (Davidson & Friede, 1953; Kuehni, 1976; Strocka, Brockes, & Paffhausen, 1983; Martínez et al., 2001). The percentage of the positive colour differences perceived by both panels (ΔV) were plotted against the CIELAB colour differences ((ΔE^*_{ab}) instrumentally measured. Then, an S-shaped curve (y = A/[1 + exp(B +Cx)]) was fitted using an iterative algorithm of successive approximations to the function and its derivatives, until maximizing the value of r^2 (Martínez et al., 2001). The software MATLAB R2011b (The MathWorks Inc., Natik, Massachusetts) was used for this purpose. For the final difference threshold, the 50% of positive responses by the observers was consider as the typical measurement of tolerance or acceptability of colour differences perceived (Alman, Berns, Snyder, & Larsen, 1989; Berns et al., 1991; Qiao, Berns, Reniff, & Montag, 1998).

2.4. Data analysis

All statistical analyses were performed using the the program Statistica 8 for Windows (StatSoft, 2007)

3. Results and discussion

3.1. Colour training

3.1.1. Panel Selection

The Farnsworth-Munsell 100-Hue Test was applied to the assessors and results showed that all passed the test with punctuation lower than 48. This mean that some of them was in the group of superior discrimination (scores lower than 16) and others were in the group of average discrimination (scores lower than 100). An example of the results is shown in Figure 3. The severity of the defect can be gauged by the extent of the 'bulge', a severe degree of defect would show clear bipolarity with high error scores; moderate cases would show small 'bulges' and lower total error scores; mild cases with good colour discrimination would show no 'bulge' and cannot be identified by this test. The food dye solutions used for tone separation and intensity ranking were analysed by image analysis. The yellowish serie had a lightness ranging from 56.02 to 76.17; the orangish from 48.14 to 74.01 and finally the reddish serie from 43.51 to 70.11. Chroma ranges were 44.45-65.23, 46.57-62.29, and 44.12-60.48 for the yellowish, the orangish and the reddish series, respectively. Finally, hue angle data ranged from 58.1º to 95.9º in yellowish serie, from 37.0º to 91.3º in the orangish serie and from 30.8º to 83.3º in the reddish serie (Figure 4).

After the panellists sorted the samples out and ranked them according to increasing intensity, just one panellist was removed from the panel following the criteria used for the selection (Spearman's ranked correlation coefficient was not significant).

3.1.2. Panel training

The objective colour of the training solutions, both commercial orange juices and their dilutions, were analysed by image analysis. The values of the coordinate L* ranged from 64.70 to 72.03; hue ranged from 78.5 (the most reddish OJ) to 92.8 (the most yellow) and the chroma ranged from 42.65 to 60.79. Figure 5 shows the samples in the CIELAB space and the coordinate L*.

Panellists were trained in different sessions until a consistent panel of assessors was obtained. Figure 6 shows the evolution in the scores for the sample COJ I along the sessions. Standard deviations decreased from the first to the last session (from 1.66 to 0.55 and from 1.75 to 0.58 in hue and intensity respectively). The error between the last and the penultimate session was 6.6 % in hue and 4.5 % in intensity (both cases less than 10 %) in agreement with an increasing uniformity of the panel. Finally, eight panellists were selected to be part of the panel.

A three factor ANOVA (assessor, sample, and repetition) for samples COJ I and COJ II, and a two factor ANOVA (assessor and sample) for all the samples were performed. Non-significant effects ($p>0.05$) were obtained for assessors and repetition, however, sample effects were obtained in both tests, as it was the objective (Table 1).

In order to ascertain if the panellists were correctly trained in colour evaluation, the sensory score were compared to the instrumental values. Table 2 shows the final scores given by the panel to the samples evaluated and their standard deviations. Surprisingly, from samples COJ II to COJ II 50% panellists did not find significant differences in hue and intensity. This could be explained by the fact that they evaluated the samples separately, not comparing between them, which supposed an extra difficulty in the evaluator's task. Furthermore, differences among these samples in hue, lightness and chroma, were not too pronounced (for example among COJ II and COJ II 50%, difference in lightness was 1.8 CIELAB units). However, among samples whose differences in colorimetric parameter measured instrumentally were higher (such us COJ II and COJ II 30%) panellists did find significant differences in hue and intensity in their visual evaluations.

In order to probe the reliability of the panel, it is important to highlight its uniformity and also to evaluate the visual and instrumental correlations. The correlation between hue and the h_{ab} parameter was explored, resulting r^* statistically significant ($r^* = -0.97$). As it was explained previously, in this study colour intensity is proposed as an attribute related to both lightness and chroma. In accordance, correlation coefficients between intensity and both lightness and chroma were explored, resulting significant in both cases ($r^* = -0.89$ and $r^* = 0.74$ respectively). Equations including C^*_{ab} and L^*, like $C^*_{ab}(100\text{-}L^*)$ and C^*_{ab}/L^* were explored in a try to relate both parameters in a unique correlation with the intensity. Statistically significant correlation coefficient were obtained for both expressions ($r^* = 0.87$ and $r^* = 0.83$, respectively). In visual analysis, the observers frequently interpret chroma and lightness as the quantitative expressions of colour. In this sense, it is noteworthy to mention that the correlations found between perceived intensity and the proposed equations (which relate quantity of colour to the inverse of lightness (C^*_{ab} /L*) or in a similar way to 'darkness' (C^*_{ab}·(100-L*)), attach great relevance to establish correlations between the instrumental and sensory evaluation of colour.

Finally, a significant (*r* = 0.93*) correlation between intensity measured by the panel and the orange juice dilutions was found.

3.2. Colour differences thresholds

In Figure 7, the CIELAB colour space (*a*b** plane) *(a)* and L* *(b)* illustrates the colour of the samples included in this study and measured by digital image analyses. It can be observed that they encompassed a wide range of colour in commercial orange juices. The values of the coordinate L* ranged from 61.32 (the darkest) to 66.87 (the lightest). Hue ranged from 77.6º (the most reddish OJ) to 83.7º (the most yellowish). Range for the coordinate C^*_{ab} was 54.23-59.10. CIELAB colour differences ΔE^*_{ab} were calculated in each pair of orange juice and ranged from 0.47 to 8.53.

The results of the correlation between the instrumental measurements and the visual evaluations done by the trained and untrained panels are shown in Figures 8 and 9, respectively. The CIELAB colour-differences instrumentally measured for each pair of samples were plotted against their visually perceived colour differences (ΔV). Both experiments (*a)* side by side and *(b)* separated samples are presented as separated graphs. As mentioned above, the equation of the fitted curve for all the situations was $\Delta V = A/[1 + \exp(B + C\Delta E^*_{ab})]$. The coefficients of the fitted curves corresponding to each experiment are shown in Table 3.

For the final threshold calculations, the 50% probability was considered as a typical measurement of tolerance or acceptability of colour differences perceived by the observers (Alman, Berns, Snyder, & Larsen, 1989; Berns et al., 1991; Qiao, Berns, Reniff, & Montag, 1998). The values of ΔE^*_{ab} corresponding to 50% of colour differences perceived (ΔV) by the trained panel were 1.63 CIELAB units for the side by side experiment and 1.44 CIELAB units in the separated observations experiment (Figure 8).

Both values are very similar (only show a difference of 0.19 CIELAB units). It might be because the distance between the samples (15 cm) was not enough to change the panellists' perception comparing separated and side by side evaluation. Moreover, training of the panellists did that these small differences were not enough to change their responses. According to these results a suprathreshold of 1.5 CIELAB units could be proposed for a trained panel.

However, in the untrained panellists the values of ΔE^*_{ab} corresponding to 50% of colour differences perceived were 2.78 CIELAB units for the side by side and 2.27 CIELAB units for the separated observations (Figure 9). These higher values could be explained by the lack of training and knowledge about colour theory. Since this would be the case of the main potential consumer this is an interesting result for the citrus industry.

The higher threshold in the case of side by side samples could be related to an increase in the sensibility of the panel due to the closeness of the samples.

Up to now, no experiments on colour-differences perceived by observers in real samples of orange juice have been reported. Previous studies based on standard propose to consider a range of 0.38–0.73 CIELAB units, and over 1.75 CIELAB units as the threshold and suprathreshold colour difference, respectively, while over 5 CIELAB units could be refer as 'big colour differences' (Macadam, 1942; Brown & MacAdam, 1949; Brown & MacAdam, 1949; Brown, 1957; Wyszecki & Fielder, 1971; Witt, 1990; Luo & Rigg, 1986; Cheung & Rigg, 1986; Berns et al., 1991; Melgosa, Hita, Poza, Alman, & Berns, 1997). Considering the results previously discussed we propose a value of around 2.8 CIELAB units as a preliminary estimate of colour-difference threshold in orange juice's consumers. These results are in accordance with previously published studies in red wine which reported a value around 3.0 CIELAB units (Martínez et al., 2001). However, a lower colour difference of 1.5 CIELAB units is proposed for a trained panel.

4. Conclusions

To sum up, in this study a specific training method for visual evaluation adapted to orange juice was set up, demonstrating its utility and efficiency. Intensity is proposed as a new attribute to jointly evaluate chroma and lightness, with good correlations with the instrumental colour parameters. Furthermore, for the first time a colour-difference threshold of 1.5 CIELAB units for a trained panel and 2.8 CIELAB units for untrained panellists are proposed.

Acknowledgments

This work was supported by funding from the Consejería de Innovación Ciencia y Empresa, Junta de Andalucía by the project P08- AGR-03784. RFV holds a grant from the Consejería de Innovación Ciencia y Empresa, Junta de Andalucía. Special thanks to Francisco José Rodríguez Pulido and Jose Miguel Hernández Hierro for their collaboration on the use of software MATLAB R2011b.

References

Alman, D. H., Berns, R. S., Snyder, G. D., & Larsen, W. A. (1989). Performance testing of color-difference metrics using a color tolerance dataset. *Color Research & Application, 14,* 139-151.

Barron, R. W., Maraulja, M. D., & Huggart, R. L. (1967). Instrumental and visual methods for measuring orange juice color. Florida State Horticultural Society. In (pp. 308-311).

Berns, R. S., Alman, D. H., Reniff, L., Snyder, G. D., & Balonon-Rosen, M. R. (1991). Visual determination of suprathreshold color-difference tolerances using probit analysis. *Color Research & Application, 16,* 297-316.

Brown, W. R. J. (1957). Color Discrimination of Twelve Observers. *Journal of the Optical Society of America, 47,* 137-143.

Brown, W. R. J. & MacAdam, D. L. (1949). Visual Sensitivities to Combined Chromaticity and Luminance Differences. *Journal of the Optical Society of America, 39,* 808-823.

Calvo, C., Salvador, A., & Fiszman, S. (2001). Influence of colour intensity on the perception of colour and sweetness in various fruit-flavoured yoghurts. *European Food Research and Technology, 213,* 99-103.

Cheung, M. & Rigg, B. (1986). Colour-difference ellipsoids for five CIE colour centres. *Color Research & Application*, 11, 185-195.

CIE (2007). Commission internationale de l'Eclairage. Standard Illuminants for Colorimetry. ISO, 11664-2:2007.

Davidson, H. R. & Friede, E. L. A. I. (1953). The Size of Acceptable Color Differences. *Journal of the Optical Society of America, 43,* 581-589.

Fernández-Vázquez, R., Stinco, C., Melendez-Martínez, A. J., Heredia, F. J., & Vicario I.M (2011). Visual and instrumental evaluation of orange juice color: a consumers' preference study. *Journal of Sensory Studies, 26,* 436-444.

Fernández-Vázquez, R., Fisk, I., Hort, J., Hewson, L., Hernanz Vila, D., Heredia, F. J. et al. (2012). Colour influence on the perception of flavour, sweetness, sourness and overall liking in orange juice. *Proceedings of 5th European Conference on Sensory and Consumer Research, Bern, Switzerland*.

Frata, M., Valim, M. F., & Monteiro, M. (2006). Preference mapping of commercial orange juices. *Fruit processing, 2,* 116-121.

Gambaro, A. D. R. I., Giménez, A., & Burgueño, J. (2001). Sensory and instrumental evaluation of strawberry yogurt color. *Journal of Sensory Studies, 16*, 11-22.

Heredia, F. J., González-Miret, M. L., Álvarez, C., & Ramírez, A. (2006). DigiFood® (Análisis de imagen). [No. SE-01298.].

Ref Type: Patent

Huggart, R. L., Petrus, D. R., & Buzz Lig, B. S. (1977). Color aspects of Florida commercial grapefruit juices. *Proceedings of the Florida State Horticultural Society, 90,* 173-175.

Kuehni, R. G. (1976). Color-tolerance data and the tentative CIE 1976 L*a*b* formula. *Journal of the Optical Society of America, 66,* 497-500.

León, K., Domingo M., Franco P., & León J. (2006). Color measurement in L*a*b* units from RGB digital images. *Food Research International, 39,* 1084-1091.

Luo, M. R., Cui, C. G., & Li, C. (4-10-2001). British Patent Entitled apparatus and method for measuring colour (DigiEye System). 0124683.4.

Luo, M. R. & Rigg, B. (1986). Chromaticity-discrimination ellipses for surface colours. *Color Research & Application, 11,* 25-42.

Macadam, D. L. (1942). Visual Sensitivities to Color Differences in Daylight. *Journal of the Optical Society of America*, *32*, 247-273.

Martínez, J. A., Melgosa, M., Pérez, M. M., Hita, E., & Negueruela, A. I. (2001). Note. Visual and Instrumental Color Evaluation in Red Wines. *Food Science and Technology International*, *7,* 439-444.

Meléndez-Martínez, A. J., Vicario, I. M., & Heredia, F. J. (2005). Correlation between visual and instrumental colour measurements of orange juice dilutions: effect of the background. *Food Quality and Preference, 16*, 471-478.

Melgosa, M., Hita, E., Poza, A. J., Alman, D. H., & Berns, R. S. (1997). Suprathreshold color-difference ellipsoids for surface colors. *Color Research & Application, 22,* 148-155.

Melgosa, M., Pérez, M. M., Yebra, A., Huertas, R., & Hita, E. (2001). Some reflections and recent international recommendations on color-difference evaluation. *Óptica Pura y Aplicada, 34,* 1-10.

O'Mahony, M. (1986). *Sensory Evaluation of food: Statistical methods and procedures.* (Marcel Dekker ed.) New York.

Pangborn, R. M. (1960). Influence of Color on the Discrimination of Sweetness. *The American Journal of Psychology, 73,* 229-238.

Poelman, A. A. M. & Delahunty, C. M. (2011). The effect of preparation method and typicality of colour on children's acceptance for vegetables. *Food Quality and Preference, 22,* 355-364.

Qiao, Y., Berns, R. S., Reniff, L., & Montag, E. (1998). Visual determination of hue suprathreshold color-difference tolerances. *Color Research & Application, 23*, 302-313.

StatSoft, I. (2007). STATISTICA, version 8.0 [Computer software].

Stinco, C., Fernández-Vázquez, R., Meléndez-Martínez, A. J., Heredia, F. J., Bejines-Mejías, E., & Vicario, I. M. (2012). Influence of Different Backgrounds on the Instrumental Color Specification of Orange Juices. In J.Caivano & M. Buera (Eds.), *Color in Food. Technologigal and Psychophysical Aspects* (pp. 143-150). New York: CRC Press. Taylor & Francis Group.

Strocka, D., Brockes, A., & Paffhausen, W. (1983). Influence of experimental parameters on the evaluation of color-difference ellipsoids. *Color Research & Application, 8,* 169-175.

Tepper, B. J. (1993). Effects of a slight color variation on consumer acceptance of orange juice. *Journal of Sensory Studies, 8,* 145-154.

Wei, S.T., Ou, L.C. Luo, M.R. & Hutchings, J.B. (2012). Optimisation of food expectations using product colour and appearance. *Food Quality and Preference 23*, 49–62.

Witt, K. (1990). Parametric effects on surface color-difference evaluation at threshold. *Color Research & Application, 15*, 189-199.

Wyszecki, G. N. T. & Fielder, G. H. (1971). New Color-Matching Ellipses. *Journal of the Optical Society of America*, *61*, 1135-1152.

Figure Captions

Fig 1. Scheme of the DigiEye System.

Fig 2. Characteristics of the bottles used for containing the samples in the instrumental colour measurements and the visual evaluation.

Fig 3. Examples of the results for the Farnsworth-Munsell 100-Hue Test conducted in panellists with (a) average discrimination and (b) superior discrimination.

Fig 4. Hue corresponding to the three series of colorant dilutions (yellowish, orangish and reddish) used in the selection of the panellists to take part in the training sessions.

Fig 5. Representation of the colour coordinates of the commercial orange juice samples used in the training sessions of the panel (a) in the a*b* plane and (b) lightness values.

Fig 6. Evolution of the evaluations done by the panel for hue and intensity (mean and standard deviation) along the training sessions.

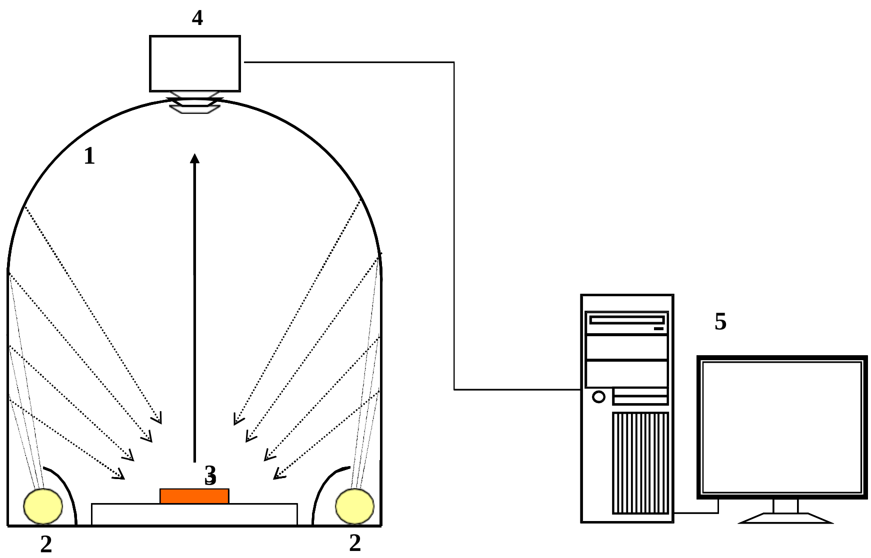
Fig 7. Representation of the colour coordinates of the orange juice samples used for the colour threshold study (a) in the a*b* plane and (b) lightness values.

Fig 8. Trained panellists ability to perceive colour differences when presented with pairs of orange juices. Sample pairs were presented to panellist in a side-by-side (a) or separated (b) condition and each point represents the number of panellists (as a percentage) who perceived a difference (ΔV) between two juices that differ in objective colour measurement as defined by ΔE^*_{ab}.

Fig 9. Untrained panellists ability to perceive colour differences when presented with pairs of orange juices. Sample pairs were presented to panellist in a side-by-side (a) or separated (b) condition and each point represents the number of panellists (as a percentage) who perceived a difference (ΔV) between two juices that differ in objective colour measurement as defined by ΔE^*_{ab}.

FIGURES

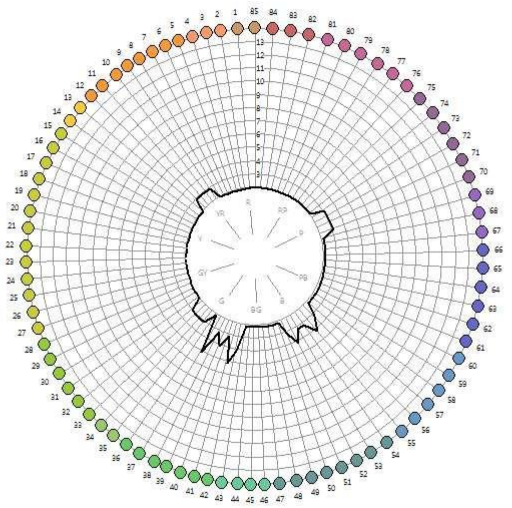
Figure 1.



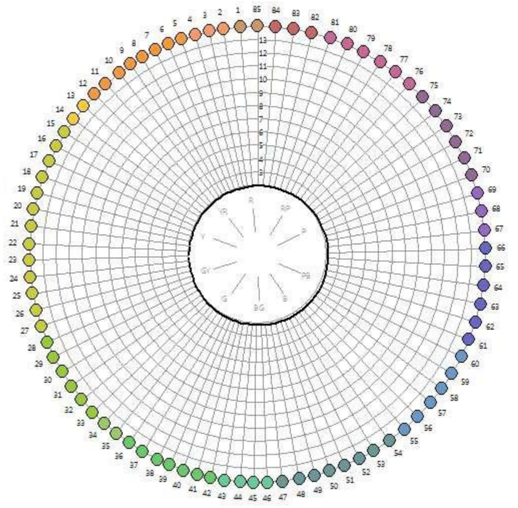
1: Domed cabinet
2: Fluorescent tubes. D65 simulator
3: Sample
4: Digital camera
5: PC Digifood® Software

Figure 2.

Figure 3.



(a)



(b)

Figure 4.

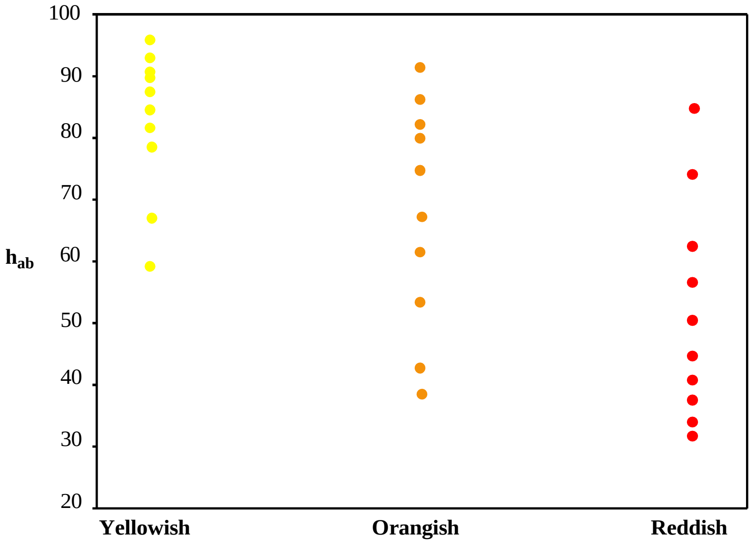


Figure 5.

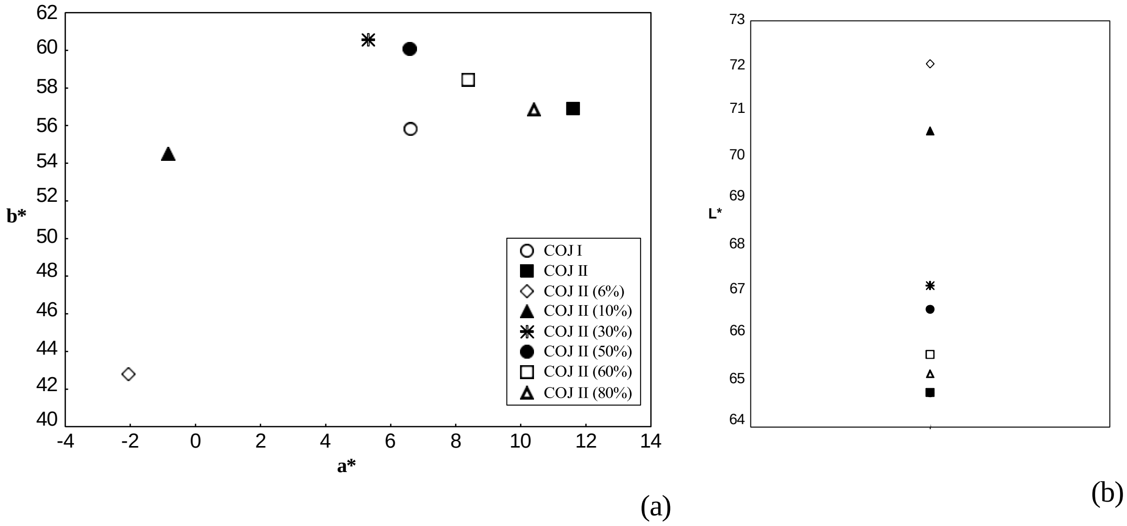

b*
a*
L*
COJ I
COJ II
COJ II (6%)
COJ II (10%)
COJ II (30%)
COJ II (50%)
COJ II (60%)
COJ II (80%)
(a)
(b)

Figure 6.

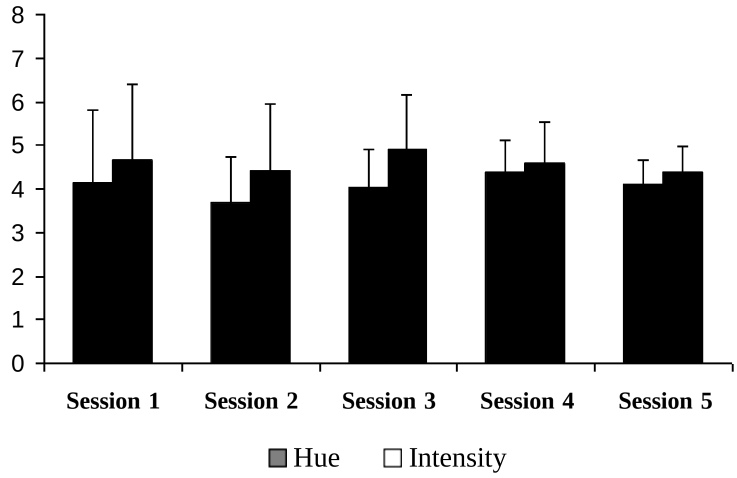
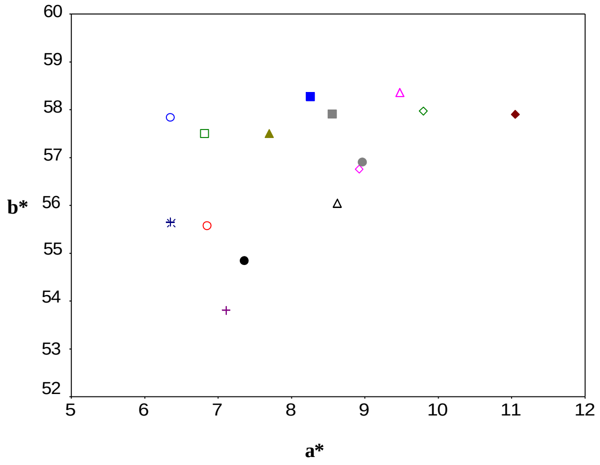
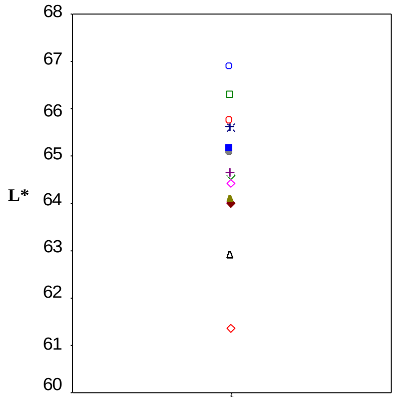

8
7
6
5
4
3
2
1
0
Session 1
Session 2
Session 3
Session 4
Session 5
Hue
Intensity

Figure 7.



60
59
58
57
56
55
54
53
52
b*
5
6
7
8
9
10
11
12
a*

(a)



68
67
66
65
64
63
62
61
60
L*

(b)

Figure 8.

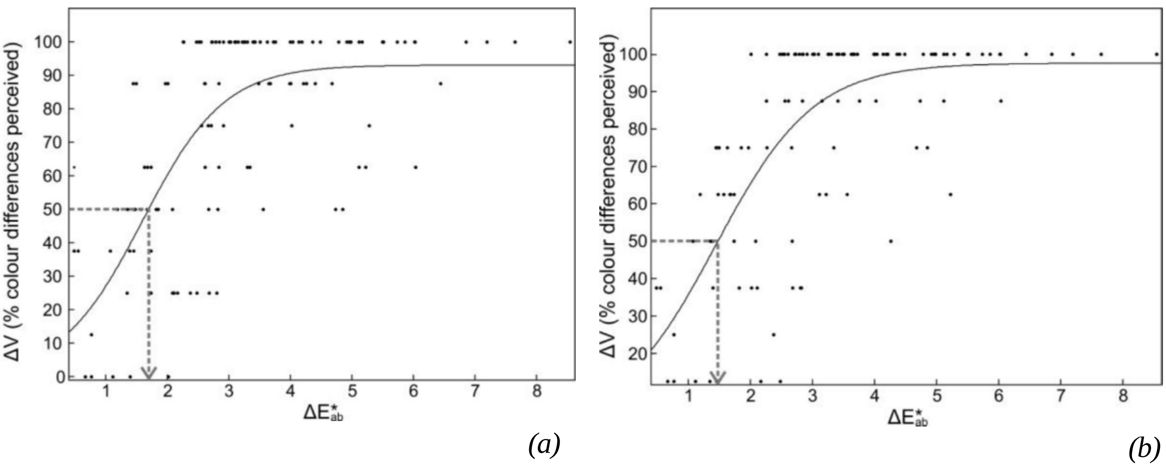
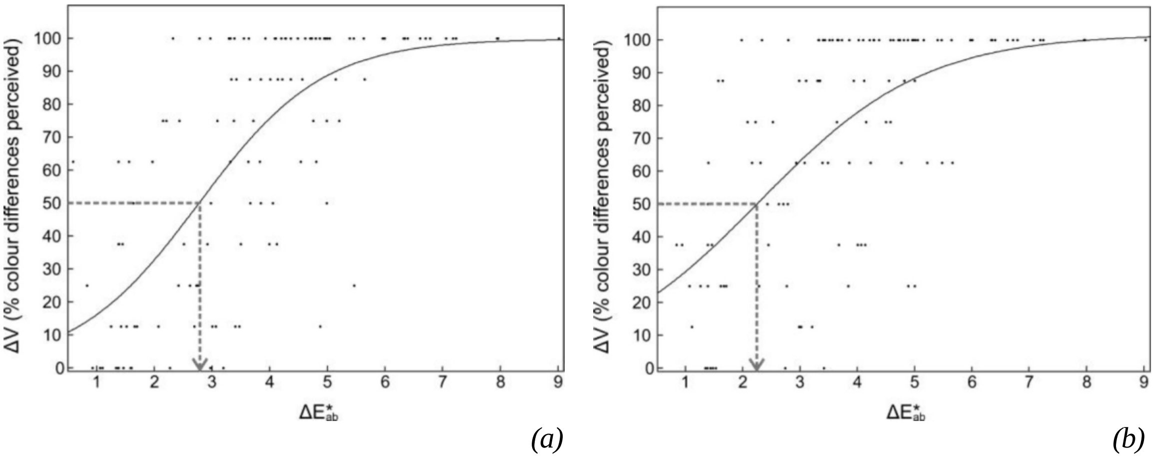

ΔV (% colour differences perceived)
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
5
6
7
8
ΔE^*_{ab}
(a)
ΔV (% colour differences perceived)
100
90
80
70
60
50
40
30
20
1
2
3
4
5
6
7
8
ΔE^*_{ab}
(b)

Figure 9.



ΔV (% colour differences perceived)
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
5
6
7
8
9
ΔE*ab
(a)
(b)

TABLES

Table 1. Results of the three factor ANOVA analysis for the colour attributes hue and intensity evaluated by the trained panellist.

Effect	Level of Significance	
	Hue	Intensity
Assessor	0.159	0.281
Sample	0.002	0.009
Repetition	0.540	0.724
Assessor - Sample	0.073	0.479
Repetition – Assessor	0.245	0.308
Repetition - Sample	0.252	0.242

Table 2. Final scores (mean and standard deviation) for hue and intensity given by the panel for the different commercial orange juices (COJI and COJII) and the corresponding dilutions, used in the training sessions.

Samples	Hue	Intensity
COJ I	4.10 ± 0.55^{a}	4.39 ± 0.58^{a}
COJ II	5.90 ± 0.60^{c}	6.88 ± 0.45^{c}
COJ II (80%)	5.66 ± 1.03^{c}	6.91 ± 0.84^{c}
COJ II (60%)	5.65 ± 1.07^{c}	6.19 ± 1.02^{c}
COJ II (50%)	5.26 ± 0.93^{c}	6.19 ± 1.01^{c}
COJ II (30%)	4.00 ± 0.34^{a}	4.43 ± 0.63^{a}
COJ II (10%)	0.84 ± 0.55^{b}	2.02 ± 0.53^{b}
COJ II (6%)	0.75 ± 0.62^{b}	1.53 ± 0.88^{b}

$^{a-c}$ Different superscripts within columns indicate statistically significant differences ($p<0.05$)

Table 3. Coefficients for the fitted equations ($\Delta V = A/[1 + \exp(B + C\Delta E^*_{ab})]$) resulting from each of the thresholds experiments (side by side and separated samples) in both panels (trained and untrained).

	Trained panel				Untrained panel			
	A	B	C	r^2	A	B	C	r^2
(a) Side by side	93.0	2.38	-1.5	0.49	99.77	2.59	-0.93	0.50
(b) Separated	97.6	1.80	-1.26	0.51	101.7	1.60	-0.70	0.40