AIFST – Sensory analysis and consumer research

Getting value out of your sensory testing – the difference from control test

by

Stephen Nottingham

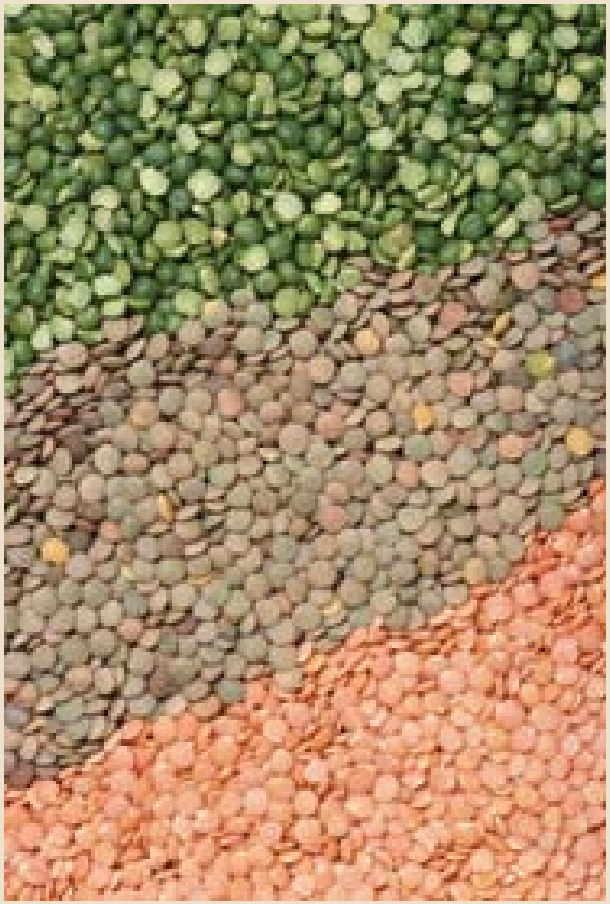
Sensory and Consumer Science

Innovative Food Technology – Emerging Technologies

Outline of presentation

- ✓ the difference from control (dfc) test
- ✓ background
- ✓ how it works
- ✓ advantages
- ✓ application in practice particularly quality control

Where is it currently being used?

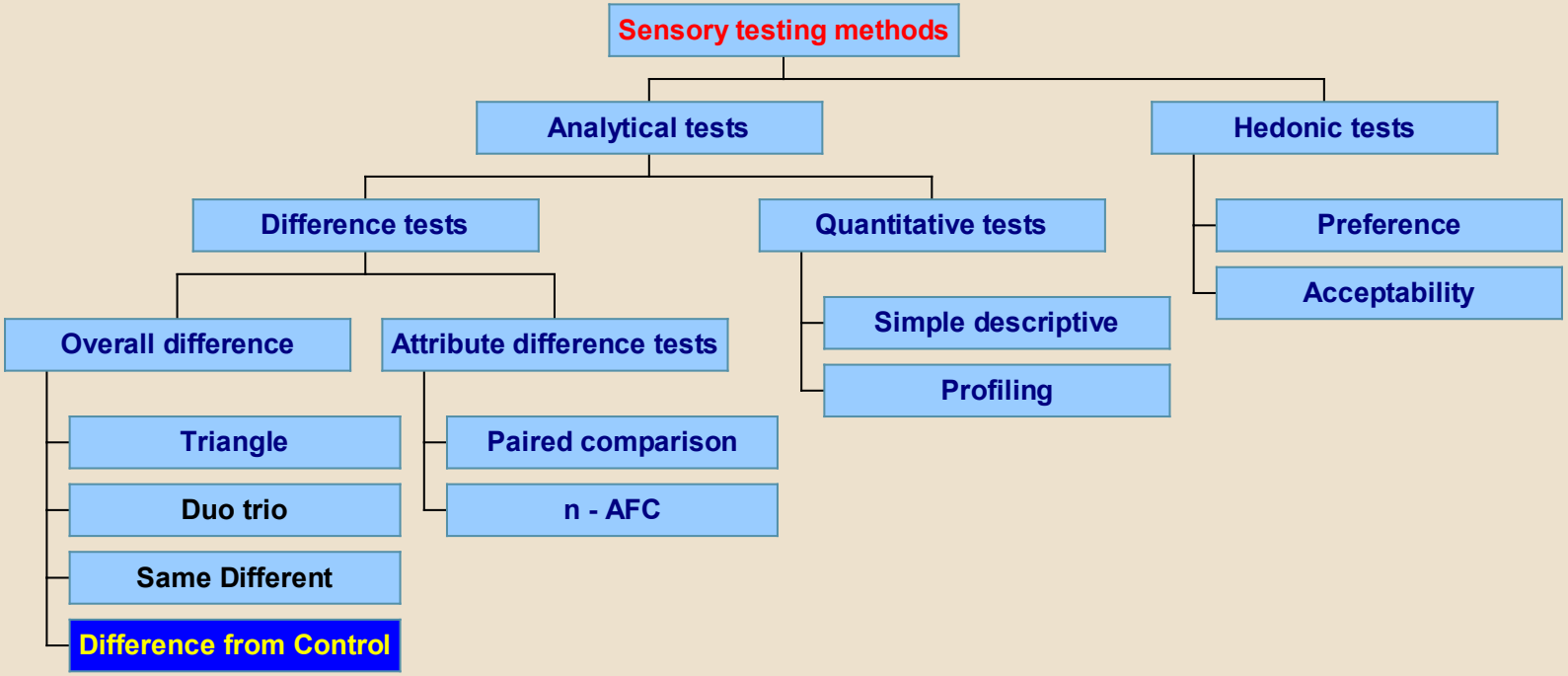


- ✓ testing several lots of an ingredient. 2000
- ✓ effect of portion size and heterogeneity on the sensory analysis of firmness using peas and biscuits as examples. 2000
- ✓ roasted peanut flavour intensity variations among US genotypes. 1995

Where is it currently being used? cont.

- ✓ evaluation of flavour and aroma of genetically modified wine strains using DFC with an experienced panel. 1992
- ✓ shelflife extension of Michigan apples using sucrose polyester. 1992
- ✓ ingredient substitutions: how to know when you have made a “match”. 1992
- ✓ comparison of control vs prototype products applied to the skin. 2000

Where does it fit?



Where does it fit?

- ✓ part of difference tests
- ✓ more specifically part of overall difference tests like triangle and duo-trio tests
- ✓ extension of the same-different or simple difference test

Typical triangle test questionnaire

TRIANGLE TEST

Name…… Date……

Product: Flavoured Milk

Two samples are the same and one is different.

Assess the samples in the order shown from left to right and circle the code of the sample that you think is different.

Please rinse your palate with water between samples.

Sample 243 Sample 344 Sample 265

You must make a choice

Comments……

Interpreting triangle test results

- ✓ analysis is straight-forward
- ✓ alpha risk 0.05 to 0.20
- ✓ similarity testing – usually more interest in this
- ✓ beta risk
- ✓ p_d the proportion of distinguishers

Interpreting triangle test results

- ✓ only answer the question - is there a significant difference between the samples?
- ✓ no idea of the nature of the difference.
- ✓ no idea of the size of the difference
- ✓ no idea of the direction of the difference
- ✓ comments may give a guide

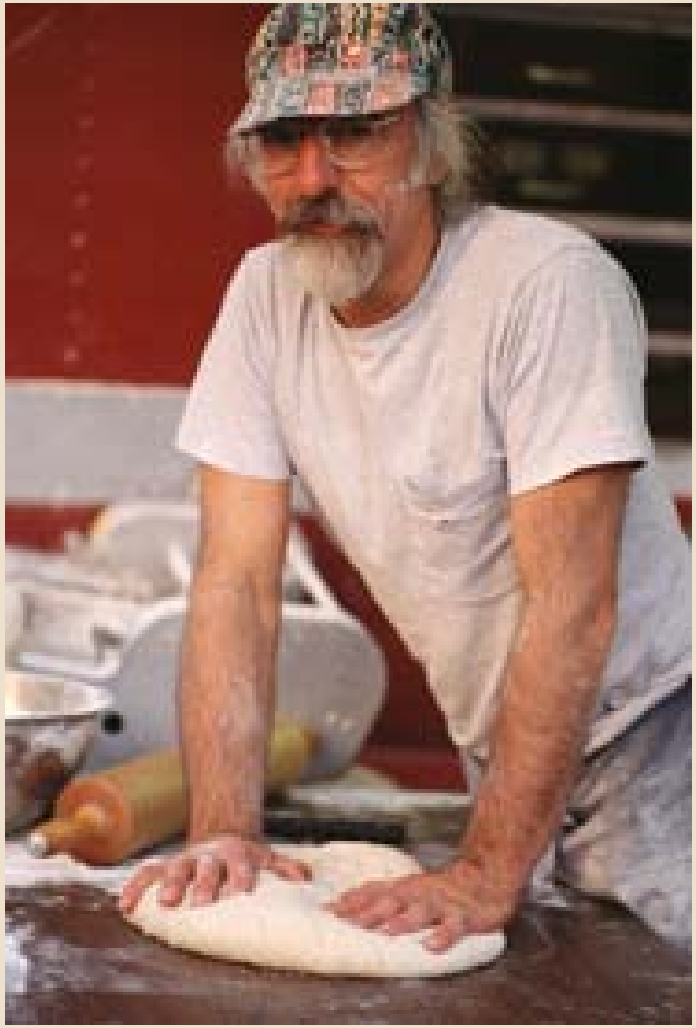
What are our options?

- ✓ descriptive analysis eg profiling, QDA
- ✓ will tell us for example product B is significantly ($p<0.05$) sweeter than product A
- ✓ gives us nature i.e. sweetness
- ✓ gives us size i.e. significance
- ✓ gives us direction i.e. B is sweeter than A
- ✓ comes at a price in terms of time and resources
- ✓ use DFC

The degree of difference

- ✓ first proposed in 1985 as degree of difference (DOD) by Aust et al in Scottsdale Arizona
- ✓ range of non-homogeneous products including devilled ham, chilli sausage, pizza and Vienna sausage
- ✓ within batch and between batch variation and test product
- ✓ concluded that DOD was more appropriate test to use on non-homogeneous product than a triangle test

Why DFC?



- ✓ batch to batch variation between days or within a day due to natural variation
- ✓ a triangle test may produce “falsely” significant differences due simply to this variation
- ✓ this can be misleading in interpretation of the result

When to use the DFC?

- ✓ if the product is non-homogeneous (can also use if homogeneous)
- ✓ we wish to determine whether a difference exists between one or more samples and a control
- ✓ we wish to estimate the size of any such differences
- ✓ in QA/QC the relative size of a difference is important for decision-making

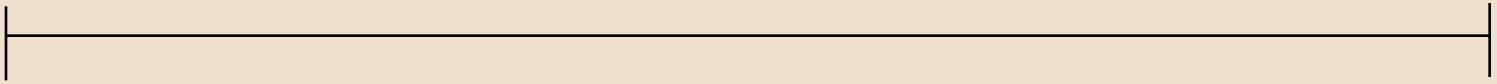
How does the DFC work?

- ✓ identified reference or control sample
- ✓ one or more coded test samples that include the reference or control (blind control)
- ✓ panellists rate the degree of difference from the identified reference on a line or category scale

Examples of scales

Verbal Category Scale	*Numerical Category Scale*
No difference	0 = No difference
Very slight difference	1
Slight/moderate difference	2
Moderate difference	3
Moderate/large difference	4
Large difference	5
Very large difference	6
	7
	8
	9 = Very large difference

Line Scale



no difference very large difference

How does the DFC work?

- ✓ 20 – 50 presentations required
- ✓ can be trained or untrained but not a mixture of the two
- ✓ all panellists should be familiar with the test format
- ✓ panellists should also be aware that some of the samples will be blind controls

How to analyse the results?

- ✓ easiest approach is to apply parametric tests such as Anova or t-tests
- ✓ calculate the mean difference from the identified control for each test sample and the blind control
- ✓ for paired data using categorical scales use non-parametric chi-squared analysis.

DIFFERENCE FROM CONTROL TEST

Name: ***Date:*** ***Time:***

Product: Beer ***Test Sample Code:***

Instructions:

1. *You have received two samples, a control sample labelled C and a test sample labelled with a 3-digit number.*
2. *Evaluate the control sample, cleanse your palate and then the test sample.*
3. *Indicate the size of the difference of the test sample relative to the control on the scale below.*

Assess sample 386 and score the overall sensory difference between the two samples using the scale below.

_____ *0 = No difference*
_____ *1*
_____ *2*
_____ *3*
_____ *4*
_____ *5*
_____ *6*
_____ *7*
_____ *8*
_____ *9 = Very large difference*

Remember that a duplicate control may be the test sample some of the time

*Comments:*_____

Panellist	Blind control	Beer A	Beer B
1	1	2	4
2	3	3	6
3	0	3	3
4	2	2	3
5	1	3	6
6	4	3	6
7	2	4	3
8	1	2	3
9	0	1	5
10	0	3	4
11	1	2	6
12	3	4	5
13	5	4	5
14	0	1	2
15	1	4	4
16	2	3	6
17	1	2	4
18	1	3	8
19	0	4	6
20	3	3	4
21	3	2	5
22	3	4	7
23	2	2	3
24	2	2	5
25	4	4	5
26	0	2	2
27	3	2	3
28	3	4	4
29	2	1	2
30	1	1	2
Mean[a]	**1.90a**	**2.67b**	**4.37c**

Analysis of DFC data

Source	Degrees of freedom	Mean square	F value	P Value
Panellists	29	3.15	2.72	0.0006
Samples	2	47.81	41.36	0.0001
Error	58	1.16		
Total	89			

LSD **= 2.02 √(2)(1.16)/30**
= 0.56

So what does this mean?

Blind Control	Beer A	Beer B
1.90a	2.67b	4.37b

Mean for the blind control sample was 1.90 so this is a measure of the so called “placebo” effect. This is the result when no treatment is applied and the sample is in fact a duplicate control.

Beer A and Beer B are significantly ($P<0.05$) different from the control and the difference is greater for Beer B as 4.37 – 2.67 = 1.70 is larger than LSD value of 0.56

How to make it work?

- ✓ adjusted mean test sample scores denoted by delta Δ expressed as

$$\overline{X}_{test} - \overline{X}_{control}$$

(average test score minus the average control score)

Δ Beer A = 2.67 – 1.90 = 0.77

Δ Beer B = 4.37 – 1.90 = 2.47

- ✓ removes the "placebo" effect and also removes the heterogeneity of the control product

Applications in quality control

- ✓ track production in terms of the amount that any finished product differs from a control or standard
- ✓ accept product that falls below some critical difference value and is closer to the control
- ✓ reject product that falls above the critical difference value

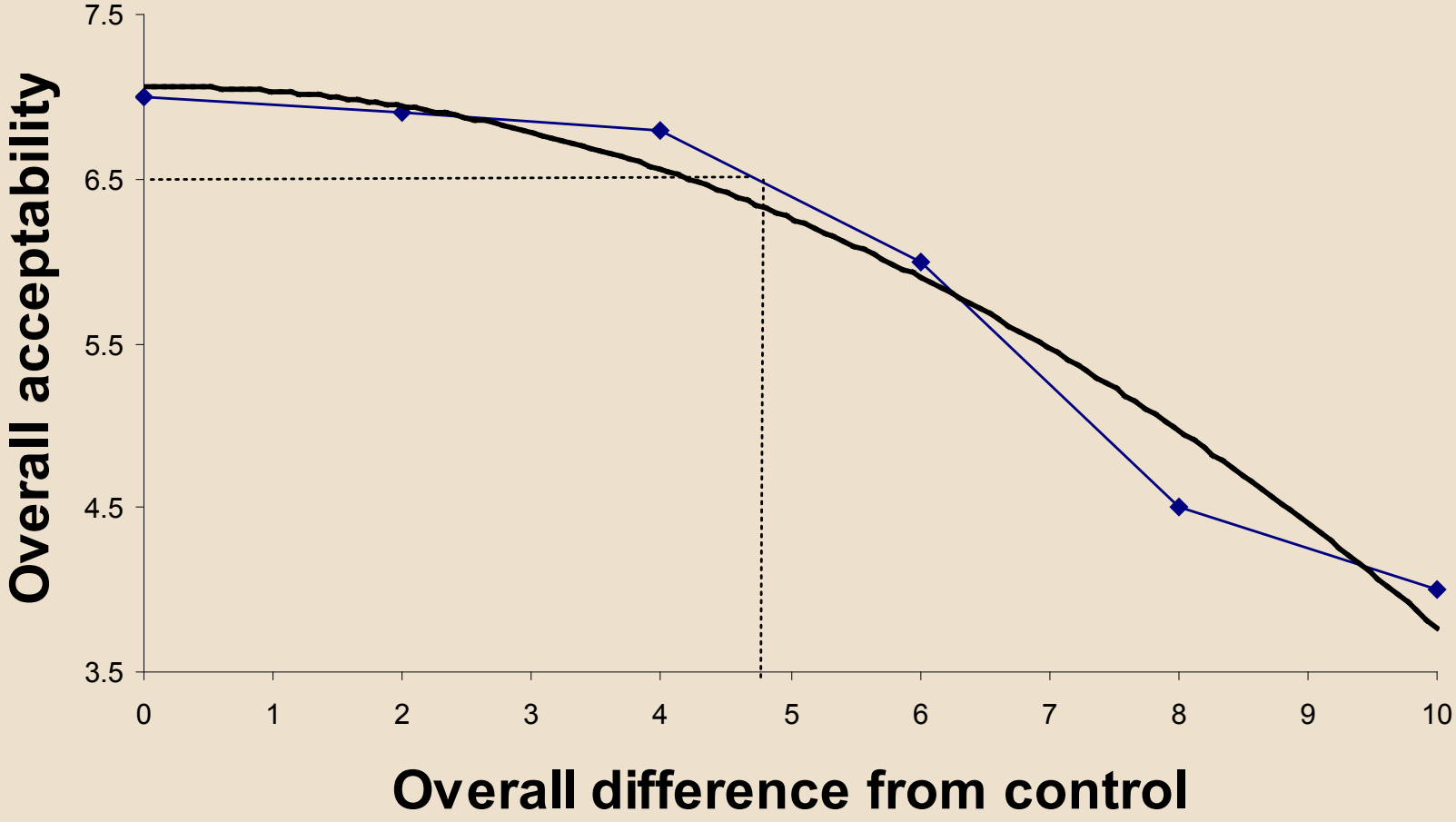
What is this critical difference value?

Suggestion by Munoz, Civille and Carr that it is set by

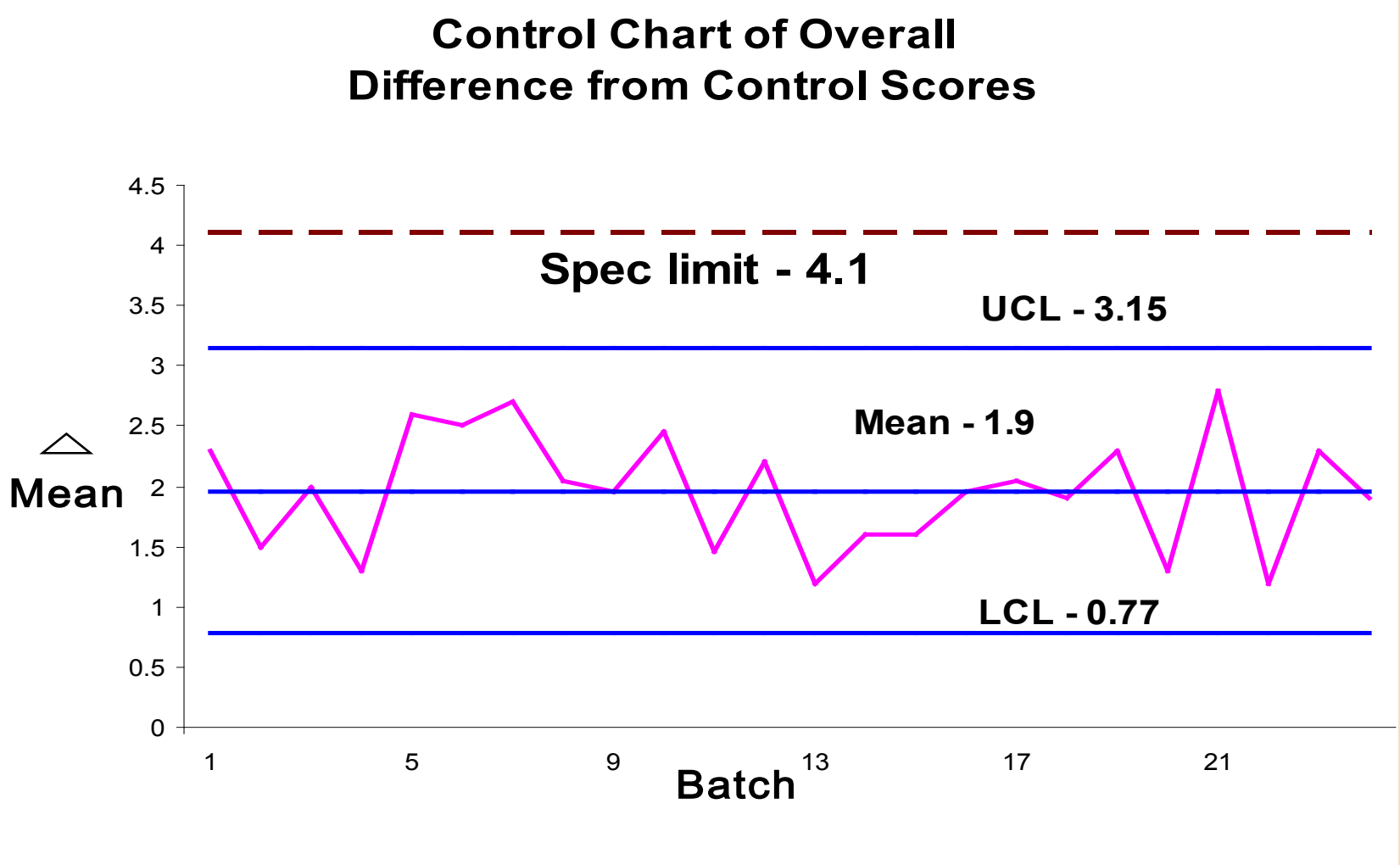
1. using responses from both consumers and management to an array of products that differ from the control or
2. using management responses only

Preference is given to option 1.

Managements primary consideration is the relationship between consumer acceptance and overall difference from control

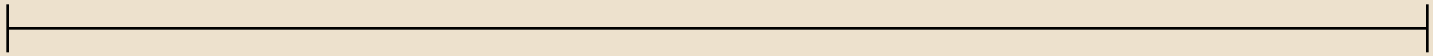


Statistical Process Control
I chart



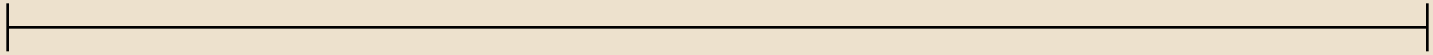
Does it stop here?

Sweetness



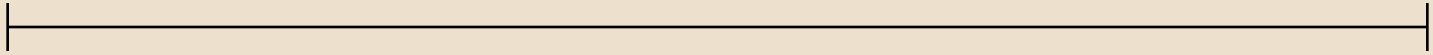
no difference extreme difference

Sourness



no difference extreme difference

Viscosity



no difference extreme difference

Does it stop here?

Sweetness

less intense more intense

Sourness

less intense more intense

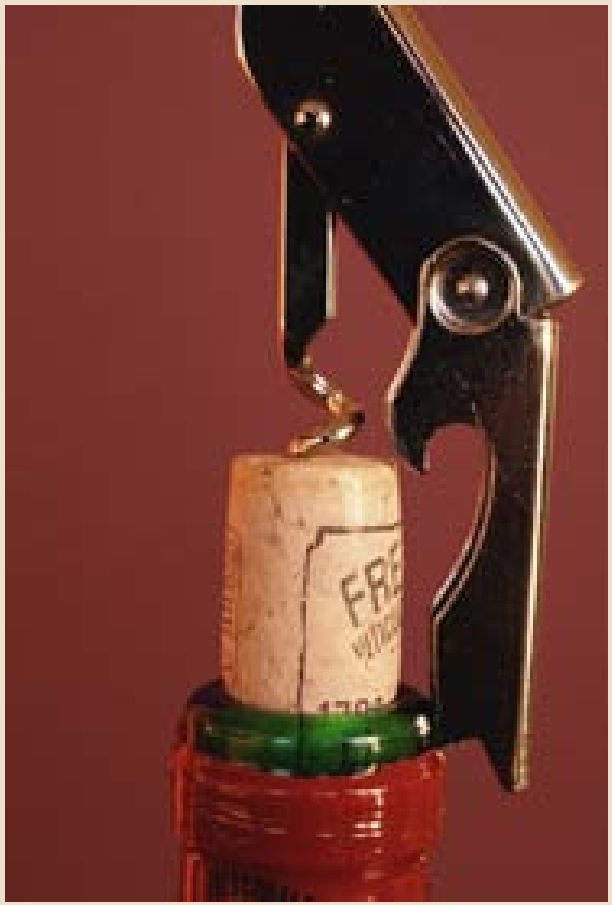
Viscosity

less intense more intense

Defining the standard

- ✓ definition of the quality standard is critical in implementing a quality control program
- ✓ with foods often difficult to obtain a suitable reference
- ✓ use mental standard
- ✓ use written standard

Mental standards



- ✓ usually created by one or more experts
- ✓ more reliable for products such as wine, coffee, olive oil etc
- ✓ perception is that the expert's opinions are not representative of consumer's opinions
- ✓ exceptions may be with raw materials or ingredients that are not directly evaluated by the consumer
- ✓ small differences between high quality levels may be decisive in their market price

Written standards

- ✓ quality grading method when constant reliable standard not available
- ✓ most recognizable is the quality index method (QIM) scheme
- ✓ developed in Europe mainly for evaluation of fish species
- ✓ demerit point system based on freshness as a key quality parameter

Quality Index Method (QIM) Scheme for Farmed Salmon

Quality parameter		Description	Score
Skin	Colour/ appearance	Pearl-shiny all over the skin	0
		The skin is less pearl-shiny	1
		The fish is yellowish, mainly near the abdomen	2
	Mucus	Clear, not clotted	0
		Milky, clotted	1
		Yellow and clotted	2
	Odour	Fresh seaweedy, nutral	0
		Cucumber, metal, hay	1
		Sour, dish cloth	2
		Rotten	3
	Texture	In rigor	0
		Finger mark disappears rapidly	1
		Finger leaves mark over 3 seconds	2
Eyes	Pupils	Clear and black, metal shiny	0
		Dark grey	1
		Matt, grey	2
	Form	Convex	0
		Flat	1
		Sunken	2
Gills	Colour	Red/dark brown	0
		Pale red, pink/light brown	1
		Grey-brown, brown, grey, green	2
	Mucus	Transparent	0
		Milky, clotted	1
		Brown, clotted	2
	Odour	Fresh, seaweed	0
		Metal, cucumber	1
		Sour, mouldy	2
		Rotten	3
Abdomen	Blood in abdomen	Blood red/not present	0
		Blood more brown, yellowish	1
	Odour	Neutral	0
		Cucumber, melon	1
		Sour, fermenting	2
		Rotten/rotten cabbage	3
Quality Index			0-24

Summary

- ✓ the DFC test can be used on both homogeneous and non-homogeneous product, particularly the latter
- ✓ it can provide not only an indication of any difference but also the size of that difference
- ✓ if we expand our scales by specifying sensory attributes and using the midpoint control scale we can get an indication of direction and the nature of the difference
- ✓ we can assess more than one pair of samples at a time, depending on sensory fatigue

Summary cont

- ✓ at the basic level no intensive training is required but as the questionnaire becomes more specific, training is required
- ✓ the test lends itself well to the quality assurance/control area and if a reliable suitable reference or control is not available then we have options using a mental standard or written standard
- ✓ sensory QC using DFC should be incorporated into an overall QC program

Summary cont

- ✓ quote from Gail Civille from Sensory Spectrum Inc who said " I love the DOD/DFC test for QC and shelf-life and in conjunction with descriptive when we want a sense of the distance of products to a "control" and you do need a controlled control."
- ✓ this really sums up the DFC test

Recommendations

- ✓ to prepare a draft Australian standard on DFC to compliment the existing standards in sensory analysis.
- ✓ there is scope for more research into developing or adapting sensory methods that are applicable in the food industry. e.g. appropriateness of small scale consumer panels