Designing your sensory test for simple and effective statistical analysis

Chantal C. Gilbert

ESN Seminar "Sensory and consumer science: the interface between research and marketing"

Pretoria, South Africa, 15-16 April 2008

Campden & Chorleywood Food Research Association Group

Experimental Design

- Experimental design is a fundamental aspect of any data collection exercise

- The design of any scientific experiment is intricately linked with the analysis of the data that are collected from the experiment

Campden & Chorleywood Food
Research Association Group

Experimental Design

- Experimental design is more than just sample presentation order, and includes:
  - Specifying of experimental and statistical objectives, including sample size
  - Selection of sensory test method and choice of scales
  - Identification of factors of interest (for example, storage conditions, packaging types, ingredients types and levels)
  - Choice of sample presentation design (such as completely randomised, randomised complete block, or balanced incomplete block designs)
  - etc.

Planning for efficient data analysis

- How you design your experiment will ultimately impact how you analyse your data
- You can help make the analysis easier (yes, really!) by taking the time, before you begin collecting data, to fully plan your experiment
- Of course, a well-planned experiment does not mean things won't go wrong (missing data, etc.), which can cause problems at the analysis stage

Today's focus…

- Understand the issue of 'sample size'
  - Will explore this in the context of consumer (affective) tests

- Demonstrate the link between elements of experimental design and statistical analysis

Sample sizes

C.C. Gilbert
Pretoria, 15-16 April 2008

Campden & Chorleywood Food
Research Association Group

Sample sizes in sensory and consumer research

- Sample size refers to the number of observations required for your study
- Ideally, sample size should be determined before the start of the test
- Sample size for sensory tests - dependent on the method. Examples of recommended sample sizes:
  - Descriptive panel (profiling): usually 8-12 trained assessors
  - Triangle testing for difference – minimum of 18 assessors
  - Triangle testing for similarity – minimum of 30 assessors

  } BS ISO 4120:2004

- Consumer samples sizes:
  - Number of consumers is also dependent on the method
  - Representative sample
  - Random sampling from target group of consumers

Sample sizes con't

- How to calculate sample size is not an easy question!
  - Statistical theory is well established, but can be difficult to understand in practical terms
- Many factors to consider:
  - Objectives of the research
  - Statistical factors
  - Practical considerations

- Sample sizes for hedonic tests will be discussed… but the idea is the same for any type of sensory test

Hypothesis testing

- When performing a statistical test, always be aware that there is an underlying hypothesis test
- Two hypotheses:
  - The null hypothesis (H_0) – usually states 'no difference'
  - The alternative hypothesis (H_A) – usually states 'there is a difference'
- The types of statistical risk:

  α = probability of rejecting H_0 when H_0 true (saying there's a difference when there is none)

  β = probability of failing to reject H_0 when H_0 is false (failing to find an existing difference)

α, β, and power

- Generally, as α decreases, β increases
- Minimising both requires more investment (larger sample size)
- Confidence = 1 - α (usually expressed as a %)
- Power is directly related to β
  - Power = 1 - β
  - Power - probability of correctly rejecting a false null hypothesis (probability that a true difference would be detected)

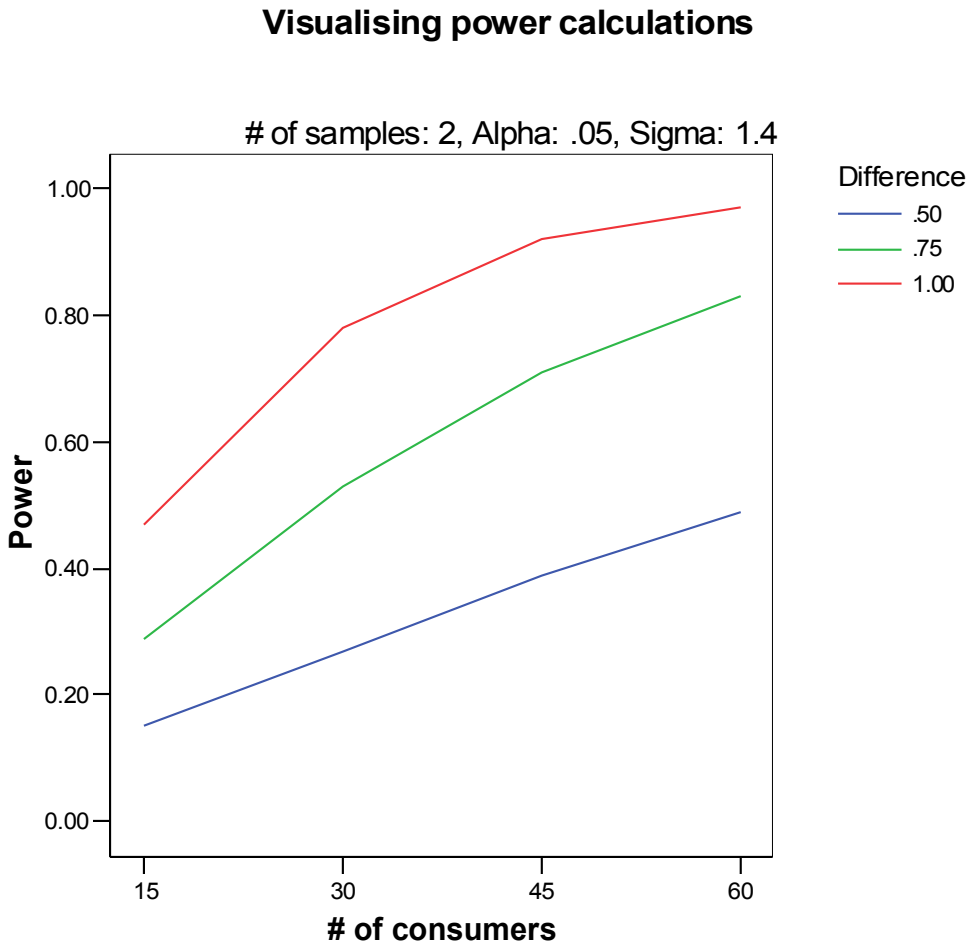
Test sensitivity

- In fact, statistically there are four interacting parameters:
  - Alpha (α) or Confidence level ($1-\alpha$)
  - Beta (β) or Power ($1-\beta$)
  - Effect size (standardised mean difference):

$$\frac{\text{Meaningful difference between samples } (\mu_1-\mu_2)}{\text{Standard Deviation } (\sigma)}$$

  - Sample size (n)
- E.g. more power is required to find smaller differences with the same level of confidence
- Find the sample size that will provide sufficient power (for given α, σ, and size of difference)

Example:
Power vs. Size of difference



C.C. Gilbert
Pretoria, 15-16 April 2008

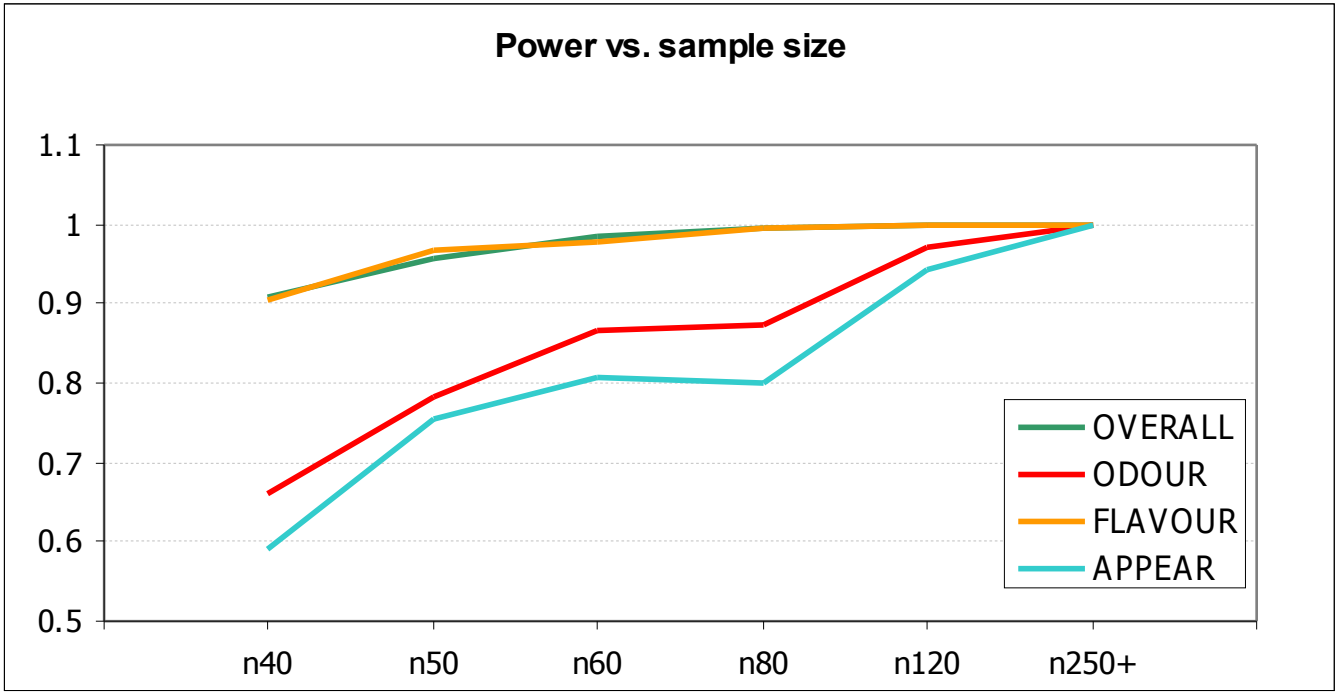
Campden & Chorleywood Food
Research Association Group

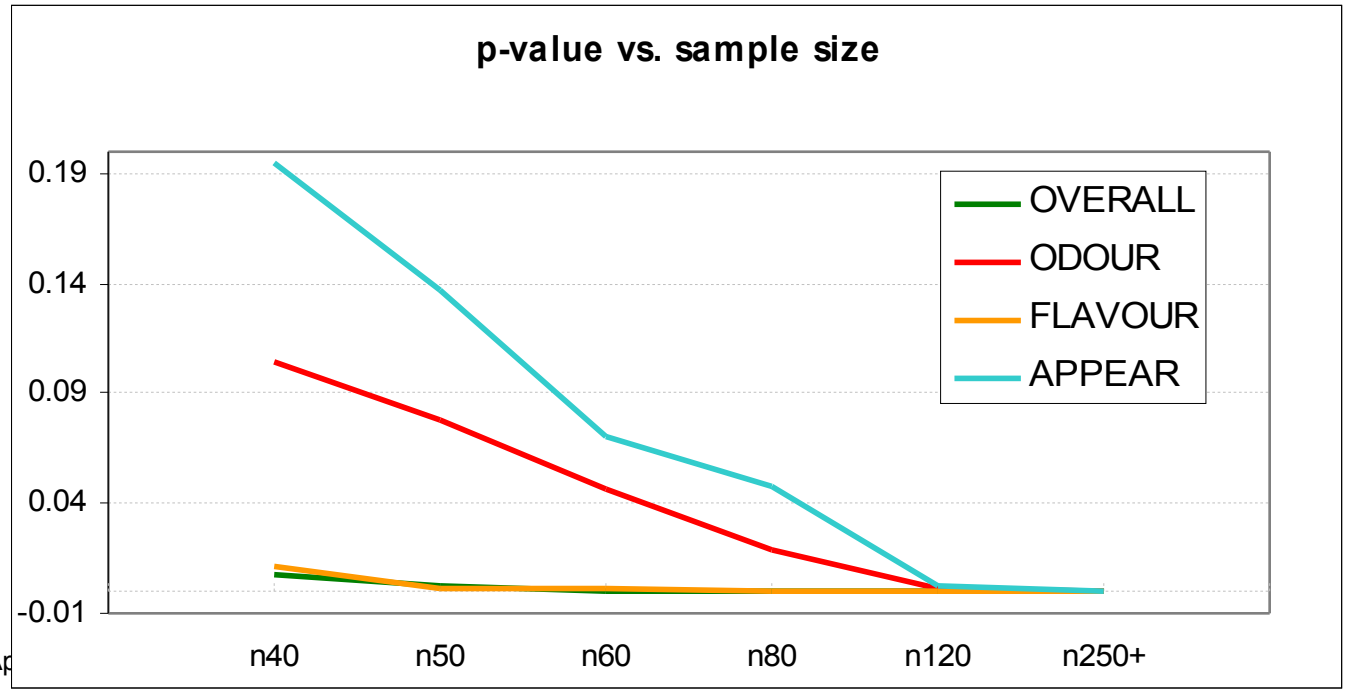
Reasons to increase sample size

- The more confident (1-α) you want to be, the more consumers you need
- The more you want to protect from 'failure' (1-β), the more consumers you need
- More difficult to conclude that small differences are significantly different, so require a larger sample size

Impact of sample size: Example using 'real' data...

- As an example, we've taken a real set of consumer data, where the total sample size was 'over 250'
- 6 samples of stirfry sauce
- 4 attributes (liking): overall, odour, flavour, appearance
- Analysis (ANOVA) was conducted using the total sample size (250+); statistical parameters including power and p-value were recorded
- Data set was then progressively reduced to the following samples sizes: 120, 80, 60, 50, 40
- 10 different random samples were taken at each sample size; analysis and recording of statistical parameters was conducted for each random sample, then summarised


Power vs. sample size
1.1
1
0.9
0.8
0.7
0.6
0.5
n40
n50
n60
n80
n120
n250+
OVERALL
ODOUR
FLAVOUR
APPEAR


p-value vs. sample size
0.19
0.14
0.09
0.04
-0.01
n40
n50
n60
n80
n120
n250+
OVERALL
ODOUR
FLAVOUR
APPEAR

Other factors that influence sample size

- Number of products
- Statistical hypothesis (e.g. 1 or 2 sided test)
- Sample presentation design and planned analysis (e.g. one-way or two-way ANOVA)
- In addition, a variety of practical issues can impact sample size:
  - Budget
  - Time
  - Type of product, amount of product available
  - etc.

Link between design elements and data analysis

C.C. Gilbert
Pretoria, 15-16 April 2008

Campden & Chorleywood Food
Research Association Group

Link between sensory method, scales and analysis

- Different sensory scales yield data with different properties - this will impact which statistical tests that can be used for analysis.

Examples:

- Classification/sorting → 'nominal' data → frequency counts & %
- Ranking → 'ordinal' data → rank sums; non-parametric statistics e.g. Friedman analysis of rank
- Line scales (scoring) → 'interval' data → means & variance; parametric methods (e.g. analysis of variance), depending on the 'shape' of the data
- Category scales (rating) → 'ordinal' data → medians & Inter-quartile ranges; non-parametric methods of analysis (although can often use ANOVA to analyse)

Link between sample presentation design and analysis

Independent Samples Design

	Product 1	Product 2
Assessor 1	X	
Assessor 2	X	
Assessor 3	X	
Assessor 4	X	
Assessor 5		X
Assessor 6		X
Assessor 7		X
Assessor 8		X

E.g. Independent samples T-test, or one-way ANOVA for ≥ 3 samples

Related Samples Design

	Product 1	Product 2
Assessor 1	X	X
Assessor 2	X	X
Assessor 3	X	X
Assessor 4	X	X
Assessor 5	X	X
Assessor 6	X	X
Assessor 7	X	X
Assessor 8	X	X

E.g. Paired T-test, or two-way ANOVA for ≥ 3 samples

Example:

Motivation for using a repeated measures design & two-way ANOVA

C.C. Gilbert
Pretoria, 15-16 April 2008

Campden & Chorleywood Food
Research Association Group

One-way ANOVA

- Treating data as 'Independent samples'

Tests of Between-Subjects Effects

Dependent Variable: GRAINY/T

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1823.440[a]	4	455.860	.796	.531
Intercept	85322.410	1	85322.410	148.932	.000
SAMPLE	1823.440	4	455.860	.796	.531
Error	54425.150	95	572.896		
Total	141571.000	100			
Corrected Total	56248.590	99			

a. R Squared = .032 (Adjusted R Squared = -.008)

Two-way ANOVA

- Same data - treating data as 'Related samples'

Tests of Between-Subjects Effects

Dependent Variable: GRAINY/T

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	45360.130[a]	13	3489.241	27.559	.000
Intercept	85322.410	1	85322.410	673.899	.000
SAMPLE	1823.440	4	455.860	3.601	.009
JUDGE	43536.690	9	4837.410	38.207	.000
Error	10888.460	86	126.610		
Total	141571.000	100			
Corrected Total	56248.590	99			

a. R Squared = .806 (Adjusted R Squared = .777)

Link between 'product design' and data analysis

Example:

Effect of replacing gelatin in mousses on sensory and rheological properties

CCFRA Research Project

C.C. Gilbert
Pretoria, 15-16 April 2008

Campden & Chorleywood Food Research Association Group

Aims

- In the past few years numerous ingredients have been used in desserts as replacements for gelatin
- The aim of this cross-departmental research project was to study how changing the type and levels of various stabiliser/emulsifier systems, along with changes to the sugar level, would affect the rheology and sensory properties of dairy mousses

Product design for development of mousse samples

A full factorial experimental design was used to combine type of stabiliser, concentration, and sugar, creating 18 samples of mousses

		Sugar	
Stabiliser	**Concentration**	Low	High
Gelatin	*2.4 (high)*	*Gel 2.4 L*	*Gel 2.4 H*
	2 (medium)	*Gel 2.0 L*	*Gel 2.0 H*
	1.5 (low)	*Gel 1.5 L*	*Gel 1.5 H*
Replacer A (alginate/ carrageenan/ pectin)	*2.7 (high)*	*A 2.7 L*	*A 2.7 H*
	2.4 (medium)	*A 2.4 L*	*A 2.4 H*
	2 (low)	*A 2.0 L*	*A 2.0 H*
Replacer B (fat and pectin based)	*2.7 (high)*	*B 2.7 L*	*B 2.7 H*
	2.4 (medium)	*B 2.4 L*	*B 2.4 H*
	2 (low)	*B 2.0 L*	*B 2.0 H*

Data collection & analysis

Data Collection

- Descriptive profiling of the 18 mousse samples was performed by a trained sensory panel
- In addition, the rheology of the mousses was measured with a rheometer

Analysis (sensory)

- Understanding differences between sample vs understanding the effect and interaction between ingredients
- Examples from two different approaches…

Analysis of Variance
Data treated as “18 samples”

Tests of Between-Subjects Effects

Dependent Variable: Rubbery_Tx

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	155134.289[a]	215	721.555	6.648	.000
Intercept	43872.793	1	43872.793	404.241	.000
Sample	86688.506	17	5099.324	46.985	.000
Judge	17014.854	11	1546.805	14.252	.000
Sample * Judge	39493.960	187	211.198	1.946	.000
Error	15628.500	144	108.531		
Total	221790.000	360			
Corrected Total	170762.789	359			

a. R Squared = .908 (Adjusted R Squared = .772)

Sample	Mean	Groupings
Gel_H_H	52.5	A
Gel_M_H	44.8	AB
Gel_H_L	37.6	B
Gel_M_L	20.4	C
Gel_L_H	15.7	CD
Gel_L_L	14.9	CDE
ReplA_H_H	7.6	DE
ReplA_H_L	6.2	DE
ReplA_M_H	4.9	DE
ReplA_L_L	2.1	E
ReplA_M_L	1.3	E
ReplA_L_H	0.4	E
ReplB_H_H	0.3	E
ReplB_M_L	0.2	E
ReplB_H_L	0.2	E
ReplB_M_H	0.1	E
ReplB_L_L	0.1	E
ReplB_L_H	0.1	E

C.C. Gilbert
Pretoria, 15-16 April 2008

Campden & Chorleywood Food
Research Association Group

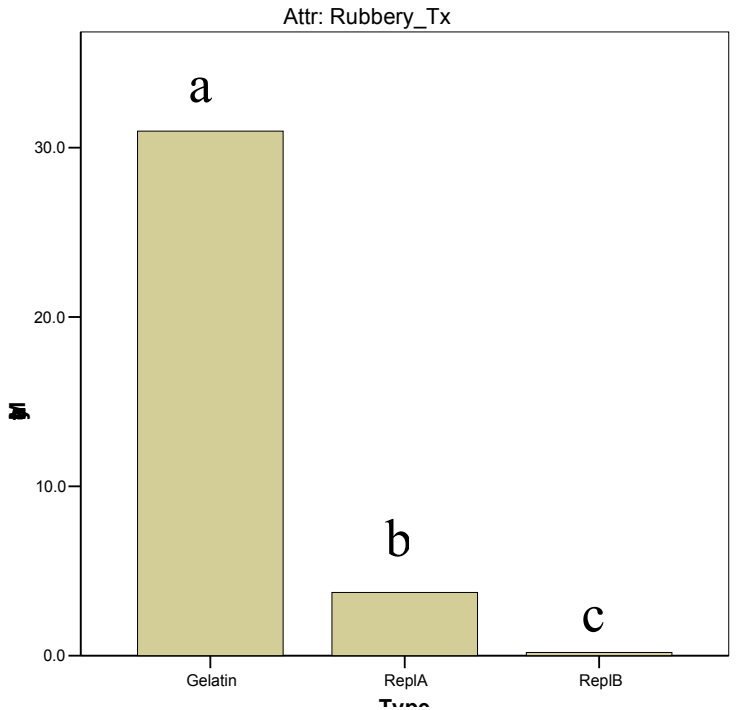
Analysis of Variance - Using the design of products

Tests of Between-Subjects Effects

Dependent Variable: Rubbery_Tx

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	81467.233[a]	79	1031.231	17.609	.000
Intercept	29248.529	1	29248.529	499.442	.000
Type	40861.787	2	20430.894	348.874	.000
Conc	5052.669	2	2526.334	43.139	.000
Sugar	1254.260	1	1254.260	21.417	.000
Judge	9835.374	11	894.125	15.268	.000
Type * Conc	6057.713	4	1514.428	25.860	.000
Type * Sugar	1975.194	2	987.597	16.864	.000
Conc * Sugar	840.465	2	420.233	7.176	.001
Type * Judge	13167.852	22	598.539	10.221	.000
Conc * Judge	1989.720	22	90.442	1.544	.069
Sugar * Judge	432.198	11	39.291	.671	.764
Error	7964.488	136	58.562		
Total	118680.250	216			
Corrected Total	89431.721	215			

a. R Squared = .911 (Adjusted R Squared = .859)

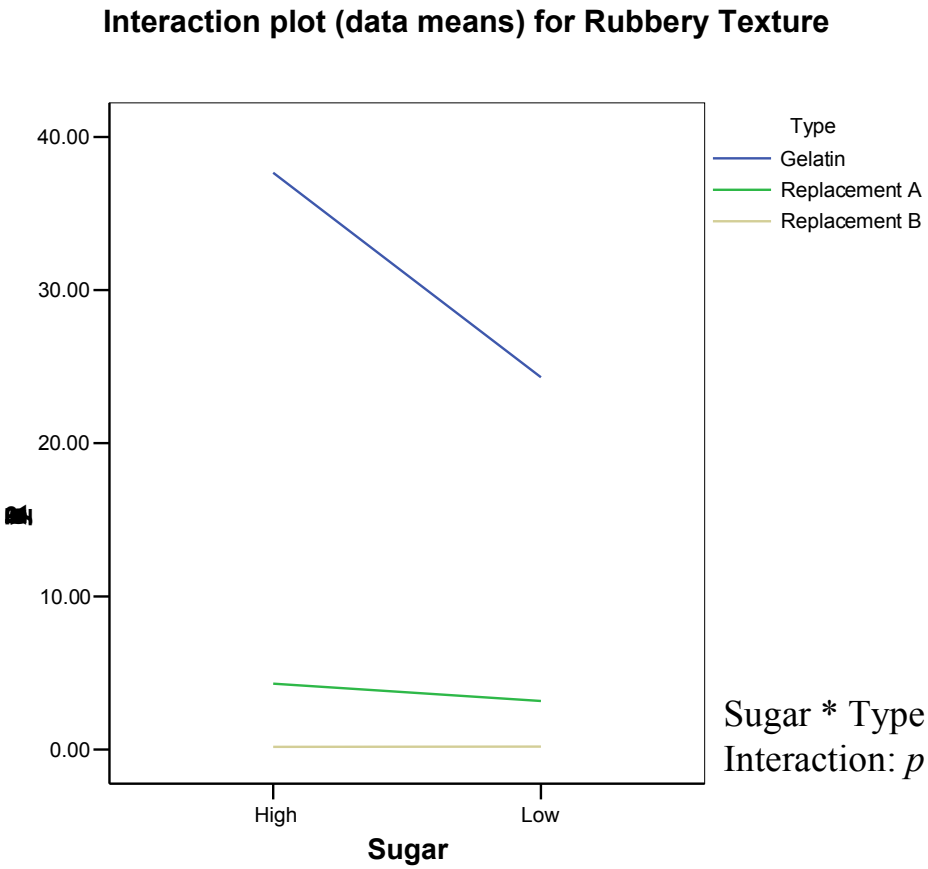


C.C. Gilbert
Pretoria, 15-16 April 2008

Campden & Chorleywood Food Research Association Group

Ingredient interactions

- Understanding of ingredient interactions and their impact on sensory perception…
- Example:



Sugar * Type
Interaction: $p < 0.001$

Conclusions

- Experimental design is an essential part of sensory and consumer research
- Sample size considerations are an important part of this planning phase
- A carefully planned and executed sensory experiment will save you time at the analysis stage
- It is important (as well as beneficial and efficient) to directly link the statistical analysis to the experimental design

Thank you for your attention!

Questions?

Chantal Gilbert

Email: c.gilbert@campden.co.uk

Tel: +44 (0)1386 842256

Campden & Chorleywood Food Research Association Group

Chipping Campden
Gloucestershire
GL55 6LD
England
www.campden.co.uk

Campden & Chorleywood Food Research Association Group