

Existing and new approaches for the analysis of CATA data



Michael Meyners
Procter & Gamble



B. Thomas Carr
Carr Consulting



John C. Castura
Compusense Inc.



Gain insights from CATA data

- Are products different?
- How do consumers perceive the products?
- What are the relationships between variables?
- What variables drive hedonic response?
- What variables should be the focus if attempting to improve a product?

Are 2 products different?

		Product B	
		Selected	Not Selected
Product A	Selected	a	b
	Not Selected	c	d

Sign Test (exact)

McNemar's Test (approximation)

Are 3 (or more) products different?

Cochran's Q Test (approximation)

extends McNemar's Test

Testing strategy

Attributes

Products

Test

1

All

Cochran's Q Test

Significant?



1

2

Sign Test

Testing strategy

Attributes

Products

Test

All

All

? Global Test ?

Significant?



1

All

Cochran's Q Test

Significant?



1

2

Sign Test

Global Test 1: Approximation

Sum of Q statistics

$$Q^* = \sum_{a=1}^{n_A} Q$$

Problem: attributes not independent

Global Test: (Quasi-)Exact

Randomization test approach

- Obtain (Q) test statistic for observed data
- Compare with null distribution based on random allocations of products to observed data *within each panelist*

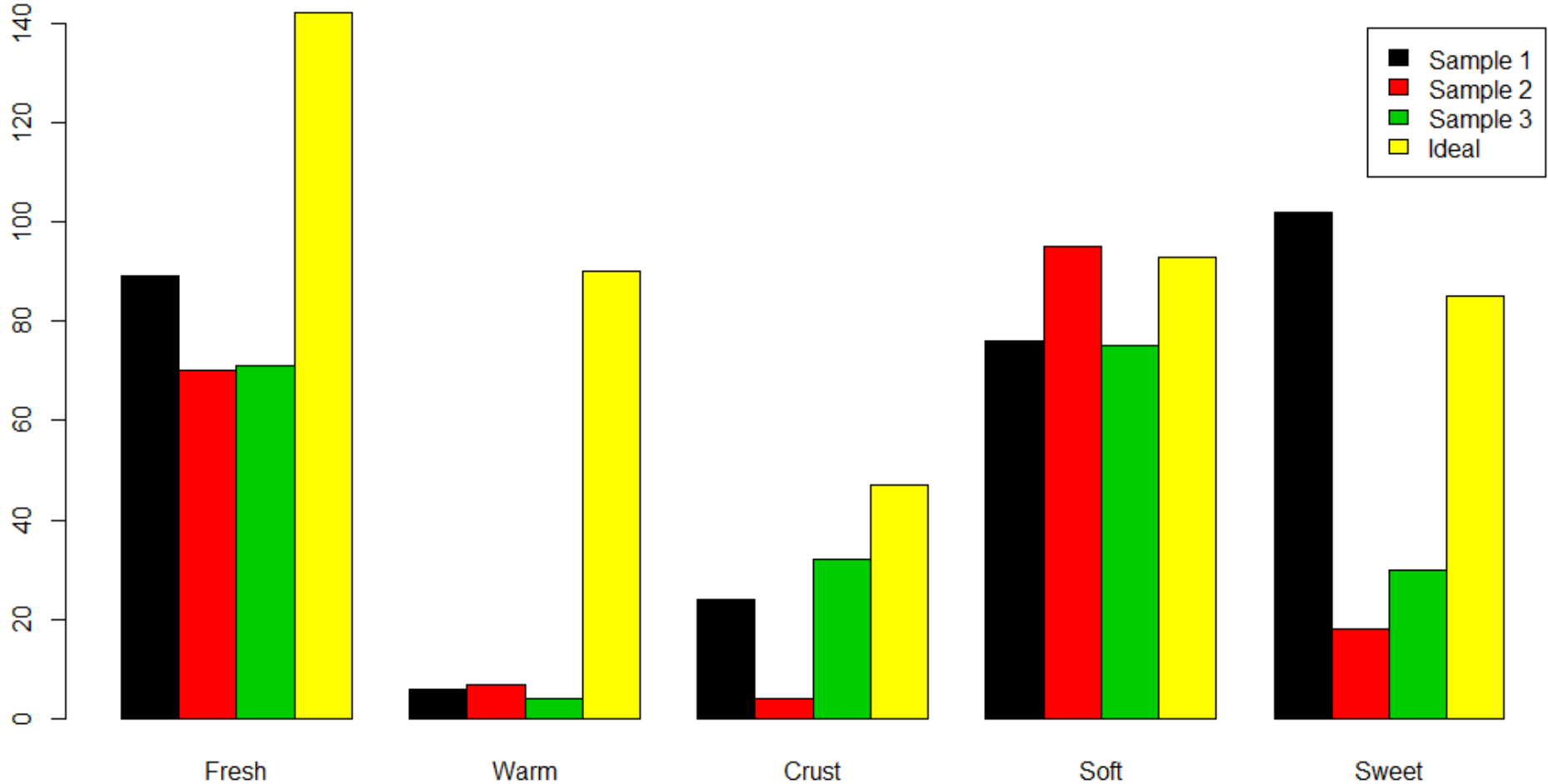
Agreement between randomization and asymptotic global tests excellent in whole grain bread data

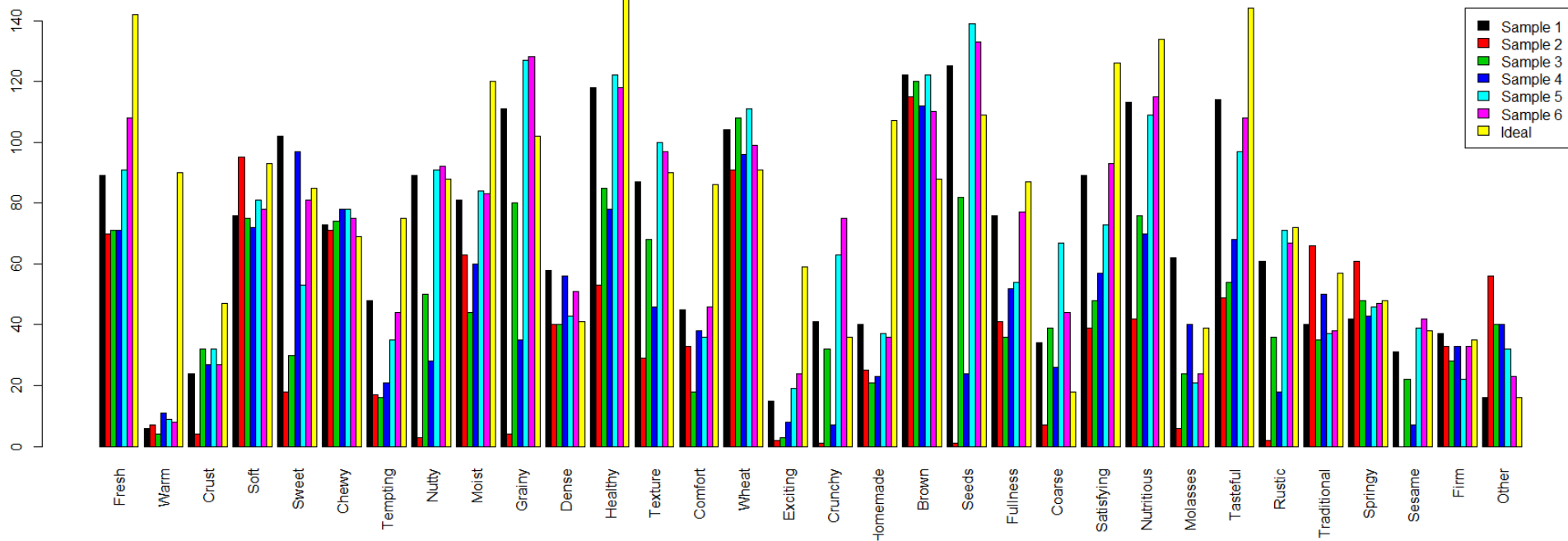
How do consumers perceive the products?

Contingency table

	1	2	3	...	Ideal	Total
Fresh	89	70	71	...	142	642
Warm	6	7	4	...	90	135
Crusty	24	4	32	...	47	193
Soft	76	95	75		93	570
Sweet	102	18	30		85	466
...
<i>Total</i>	<i>2169</i>	<i>1144</i>	<i>1539</i>	<i>...</i>	<i>2579</i>	

Bar charts

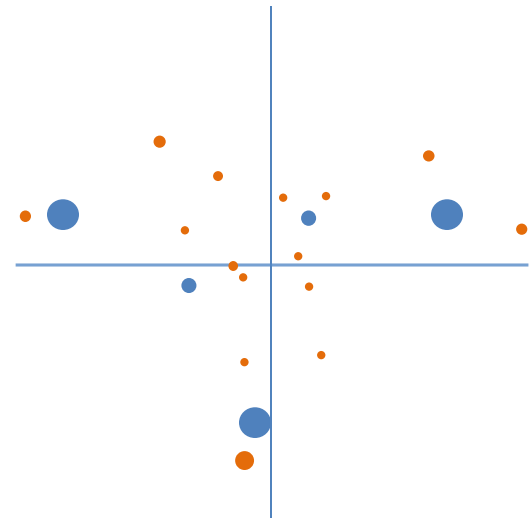
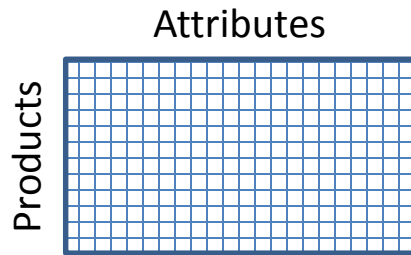




Problem: This is rather difficult to read.

Correspondence Analysis

Reduces dimensionality of contingency tables
(*Think: PCA for ordinary data*)



Correspondence Analysis

Could use...

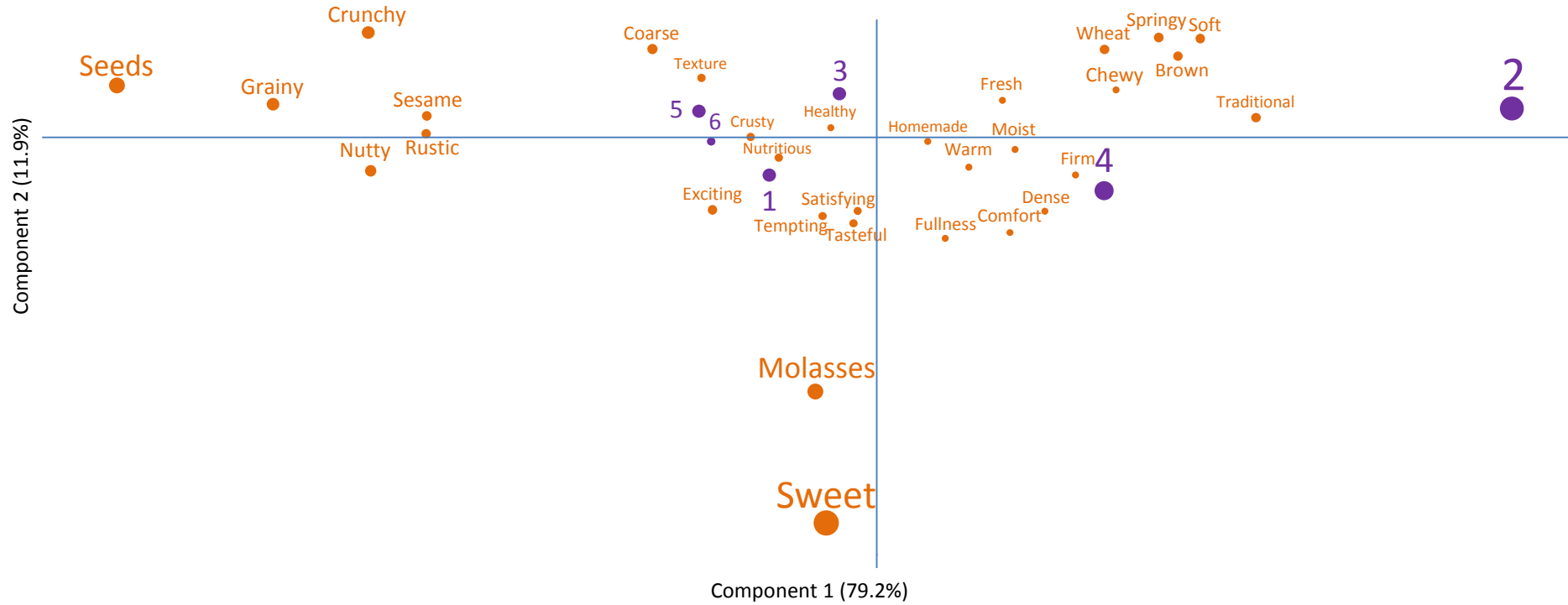
Chi-square distances

Sensitive to attributes with low counts

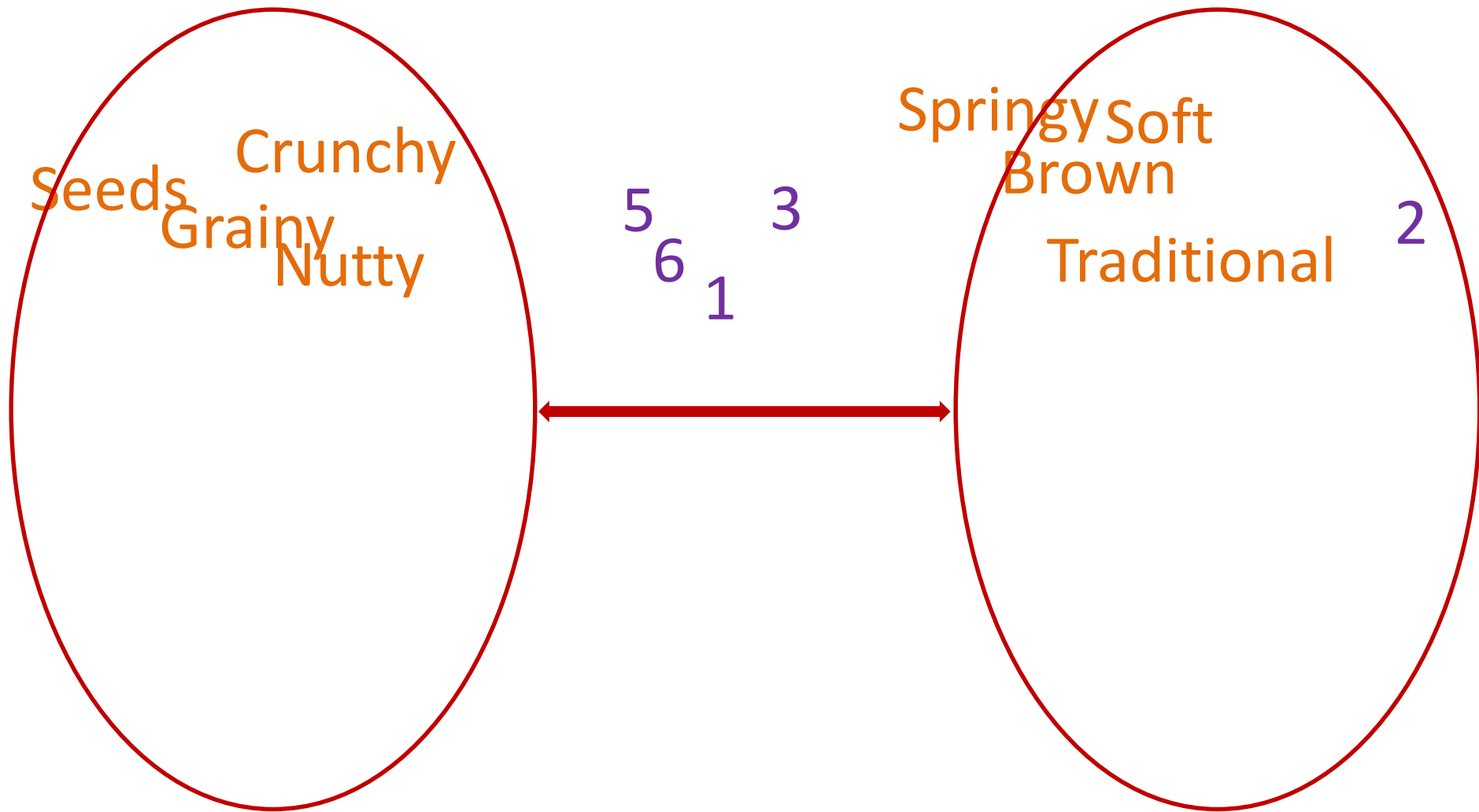
Hellinger distances

Low counts not a problem

Correspondence Analysis

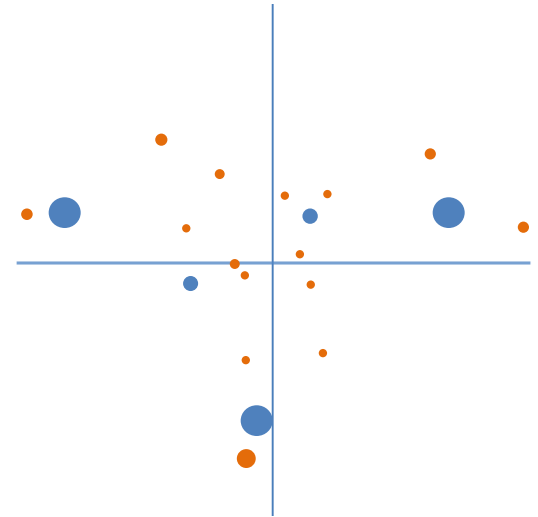


Component 1 (79.2%)



Correspondence Analysis

Can only look at 2 (or 3) dimensions at once



Problem: True relationships between products and attributes might be obscured.

Multidimensional Alignment (MDA)

Solution: Look at each product individually



Crusty

Seeds
Nutty

Crunchy
Grainy
Nutritious

Texture
Sesame
Healthy

Rustic
Course

Exciting

Tempting

Tasteful

Molasses

Homemade
Fresh
Brown

Warm
Springy
Chewy

Dense

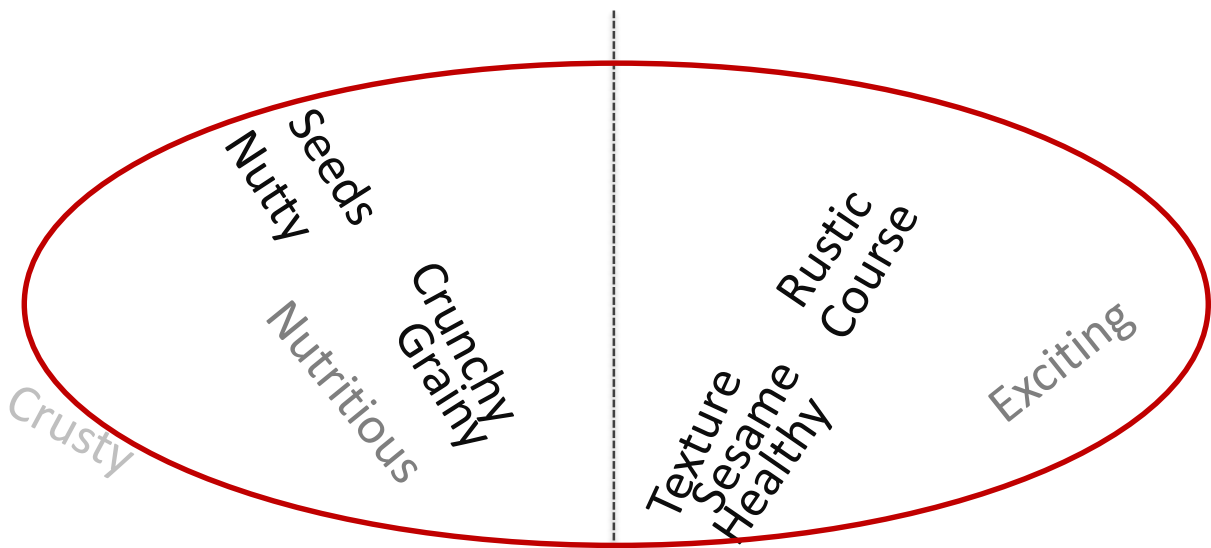
Sweet

Satisfying

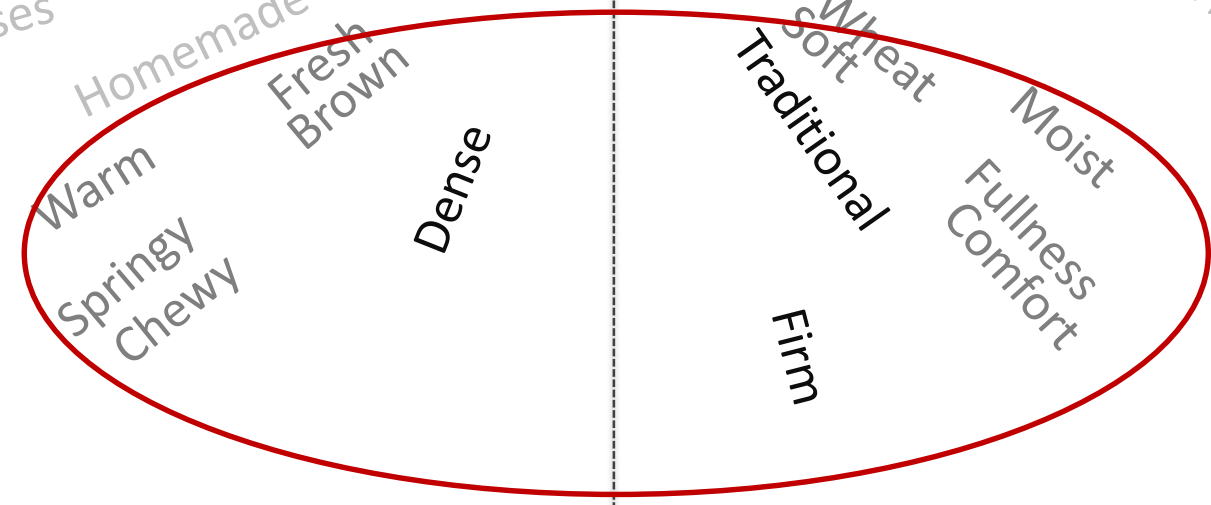
Wheat
Soft
Traditional

Moist
Fullness
Comfort

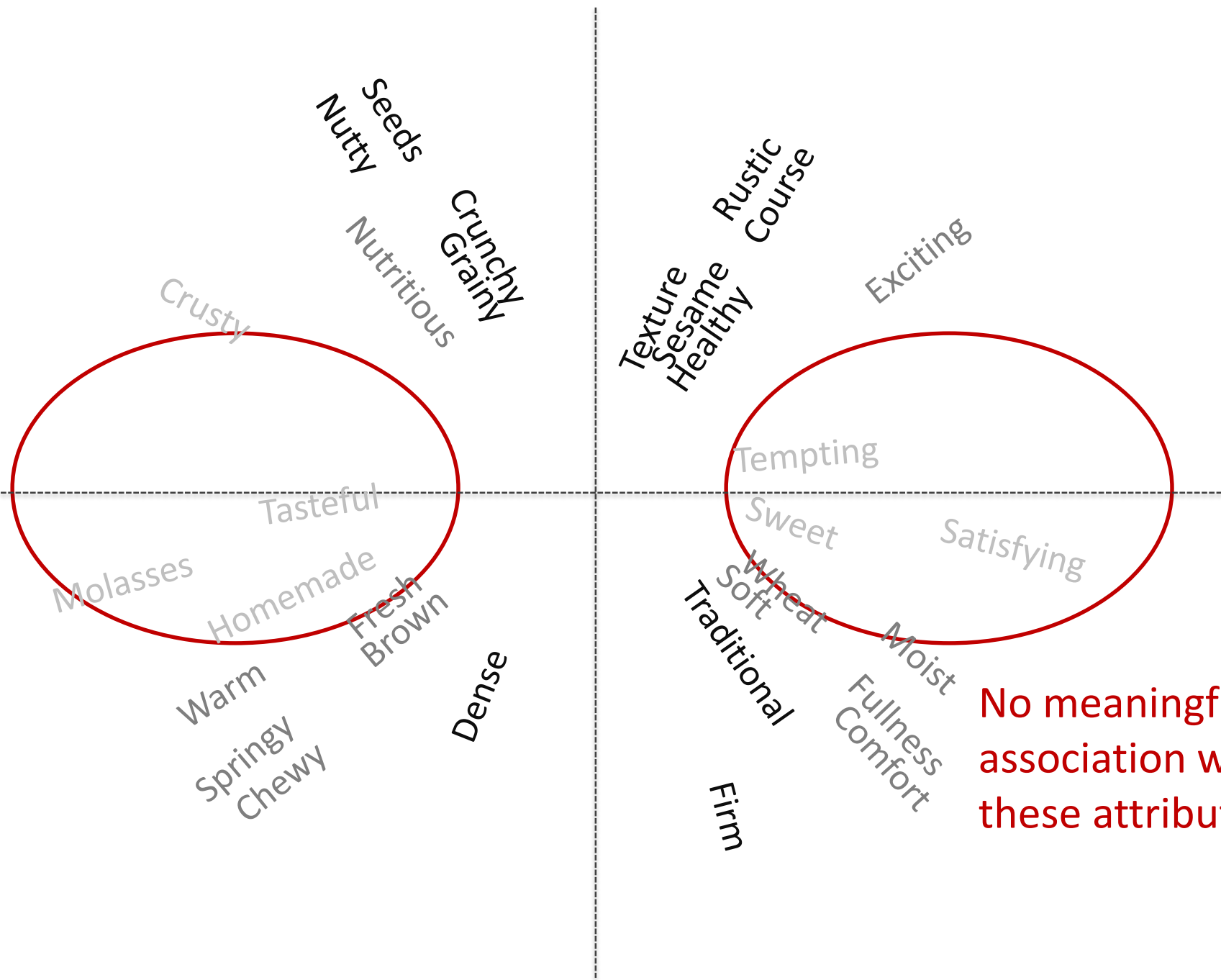
Firm



Positive association with these attributes



Negative association with these attributes



No meaningful association with these attributes

How are CATA attributes related?

Solution:

ϕ -coefficient measures correlation between 2 binary attributes

(related to χ^2 for 2x2 contingency table)

Classical
MDS on
matrix
of $1 - \varphi$
distances



Classical
MDS on
matrix
of $1 - \varphi$
distances



CATA + Liking

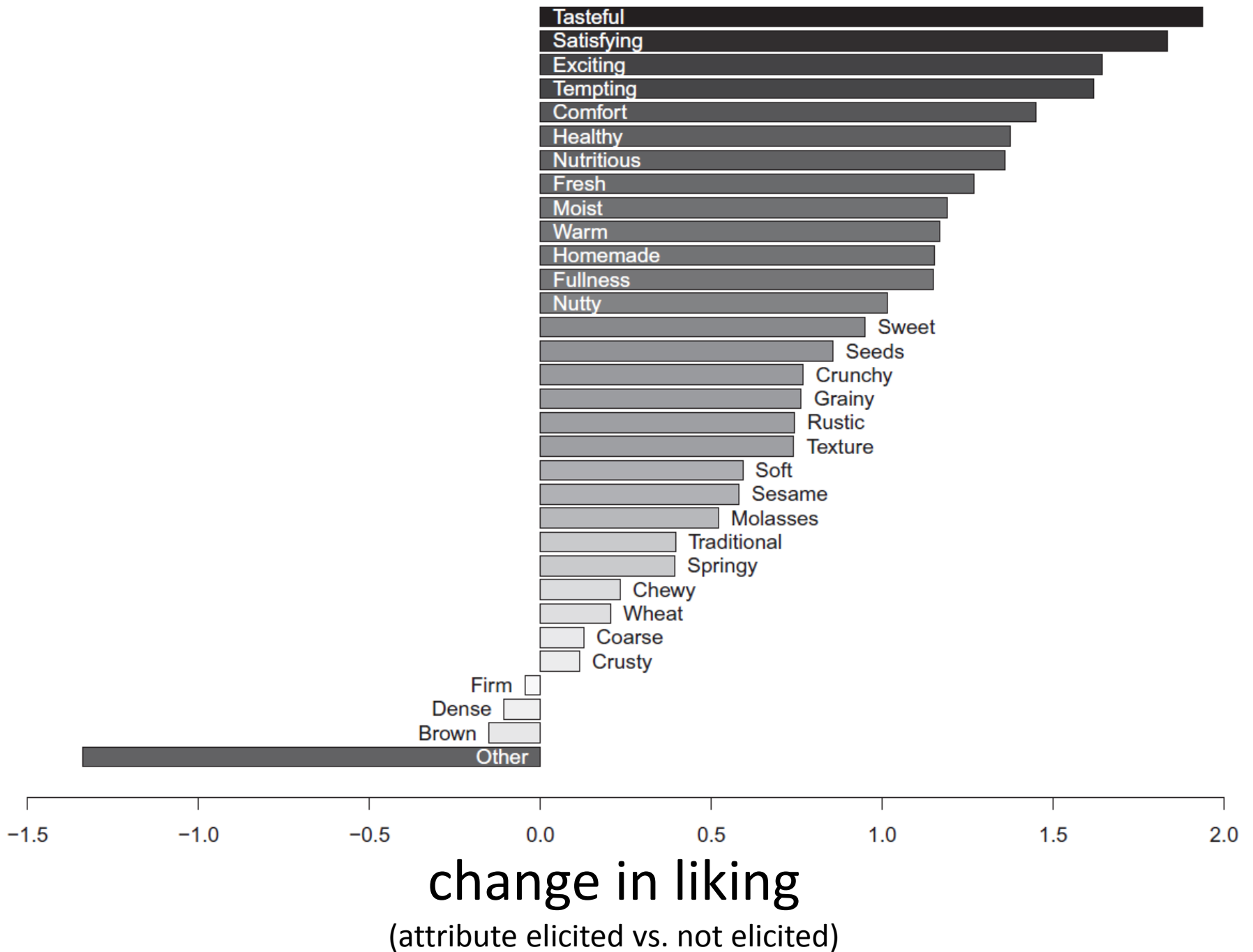
Problem:

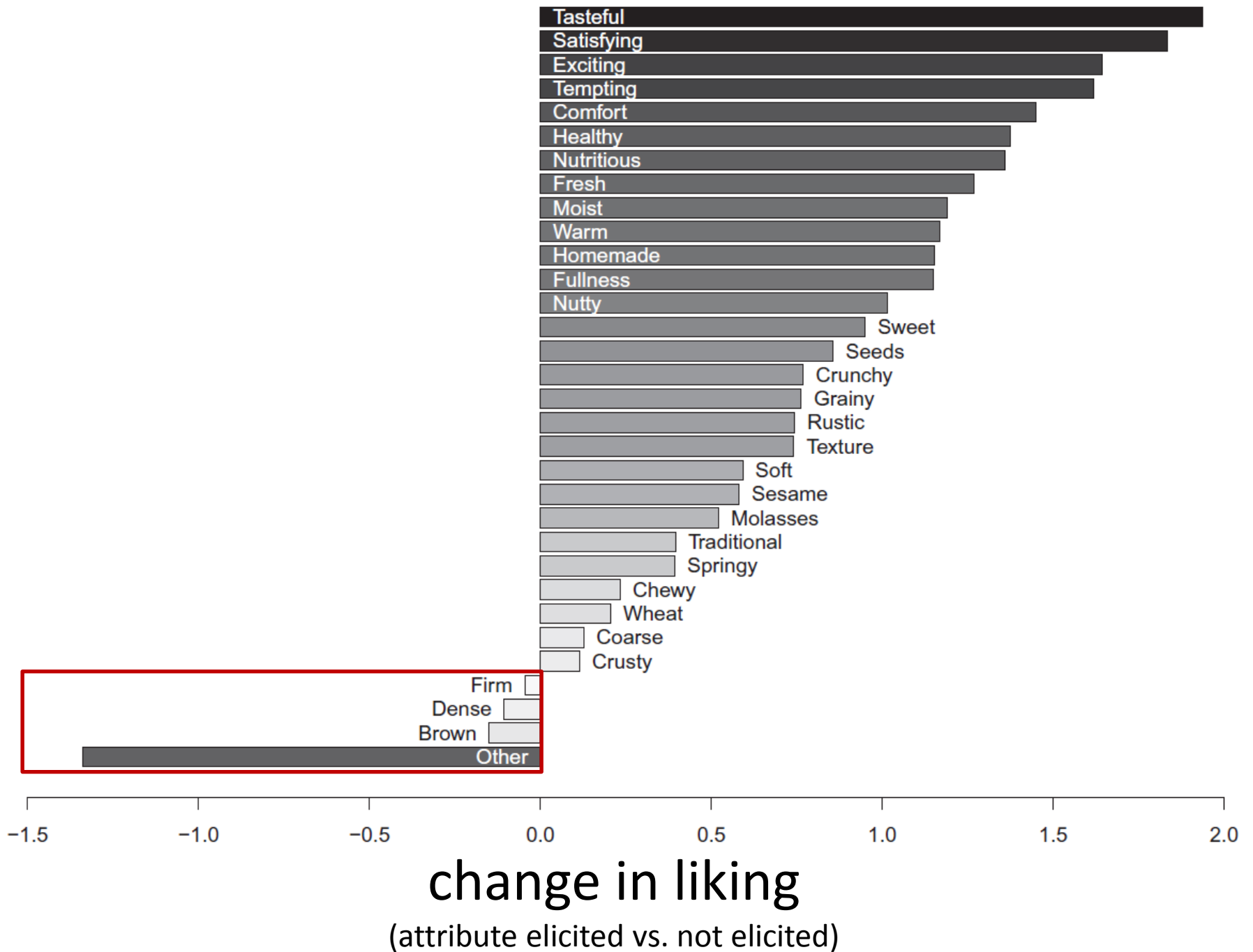
We want to understand the relationship between **what consumers perceive** and their **hedonic response**

Penalty-Lift Analysis

Solution:

Look at elicitations of attributes and association with liking





Comparison with Ideal

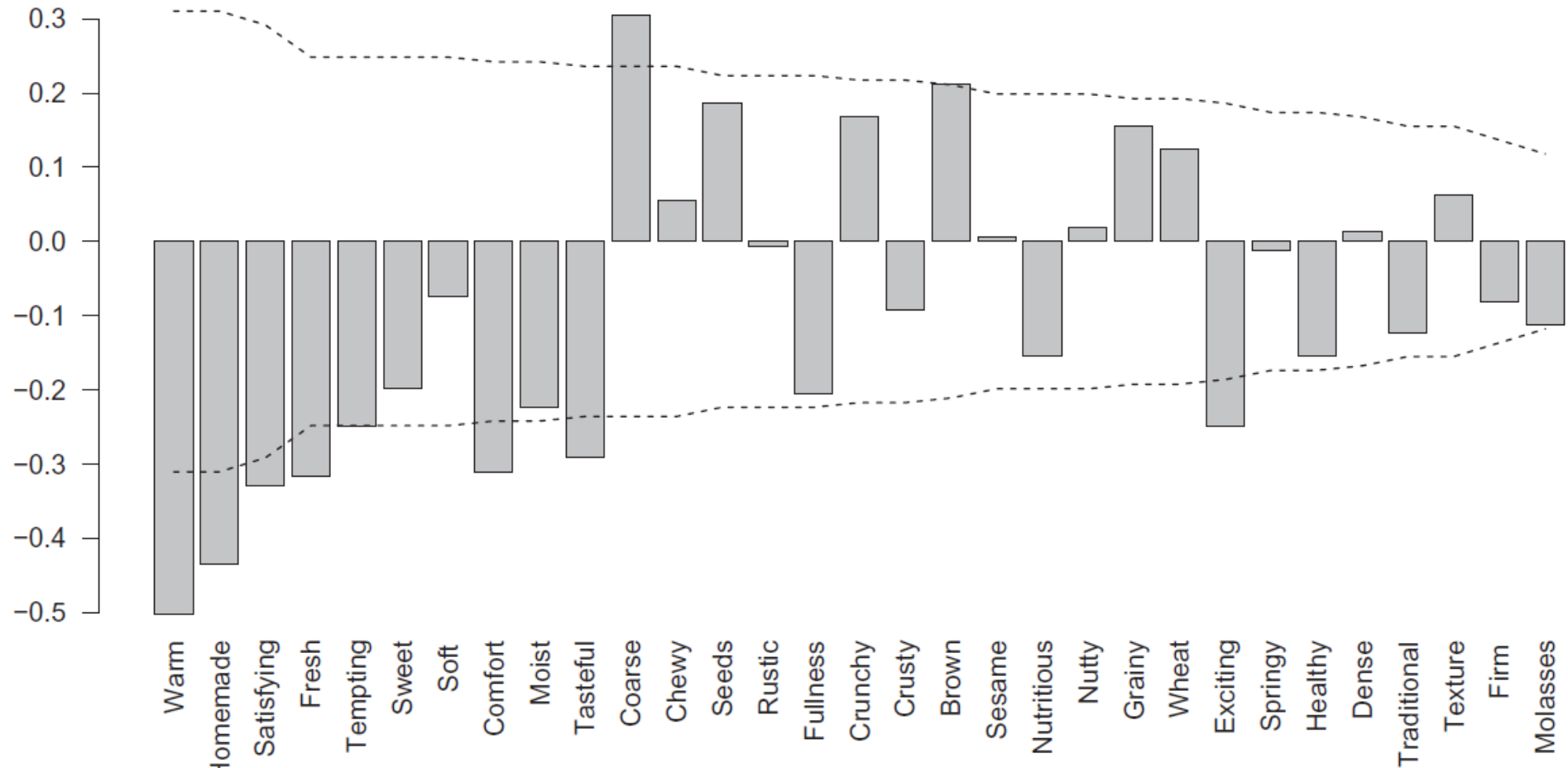
Problem: How to modify a product to delight consumers?

Solution 1: Close gap between product and ideal

(accounting for natural variability!)

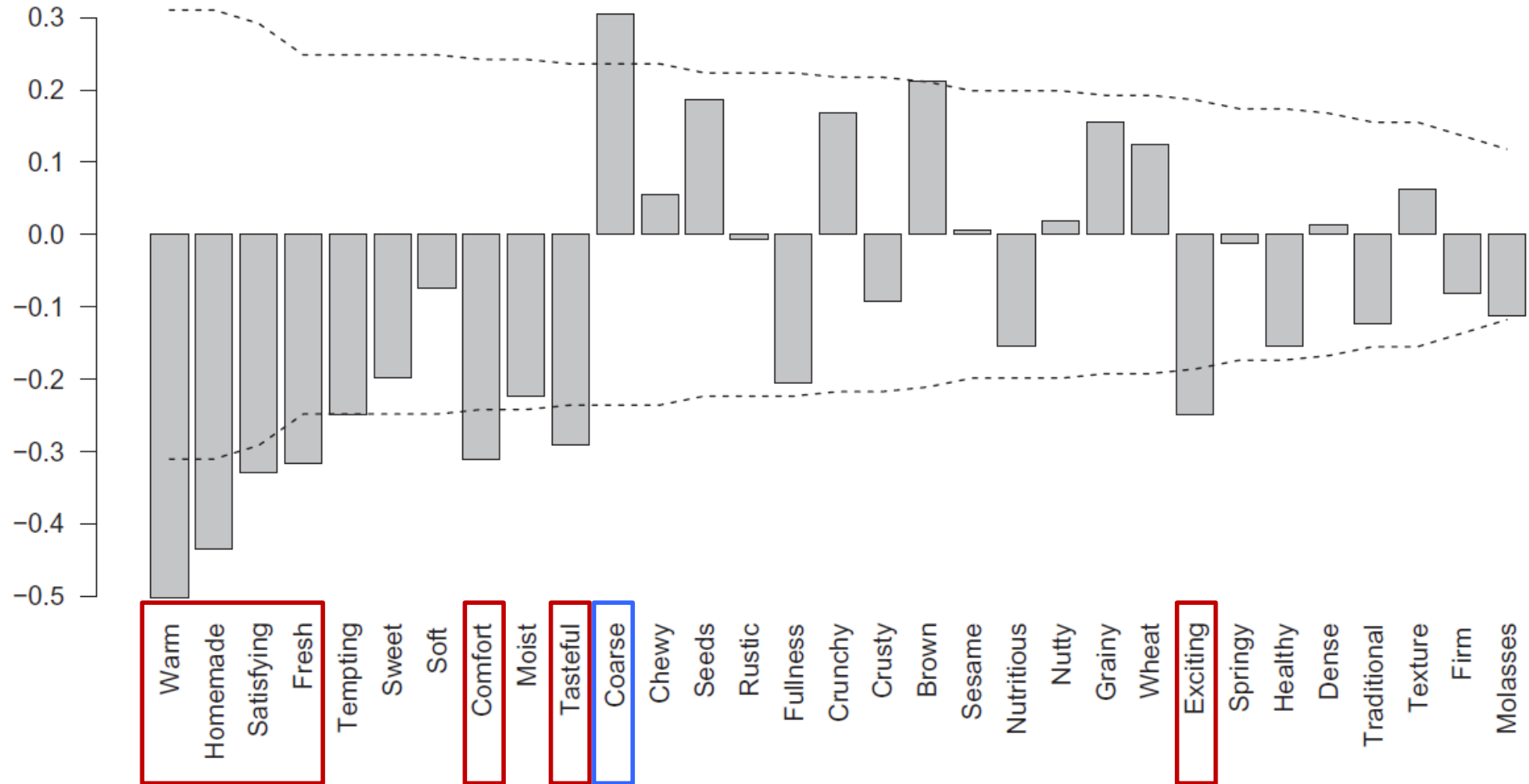
Difference in elicitation rates

Product – Ideal



Difference in elicitation rates

Product – Ideal

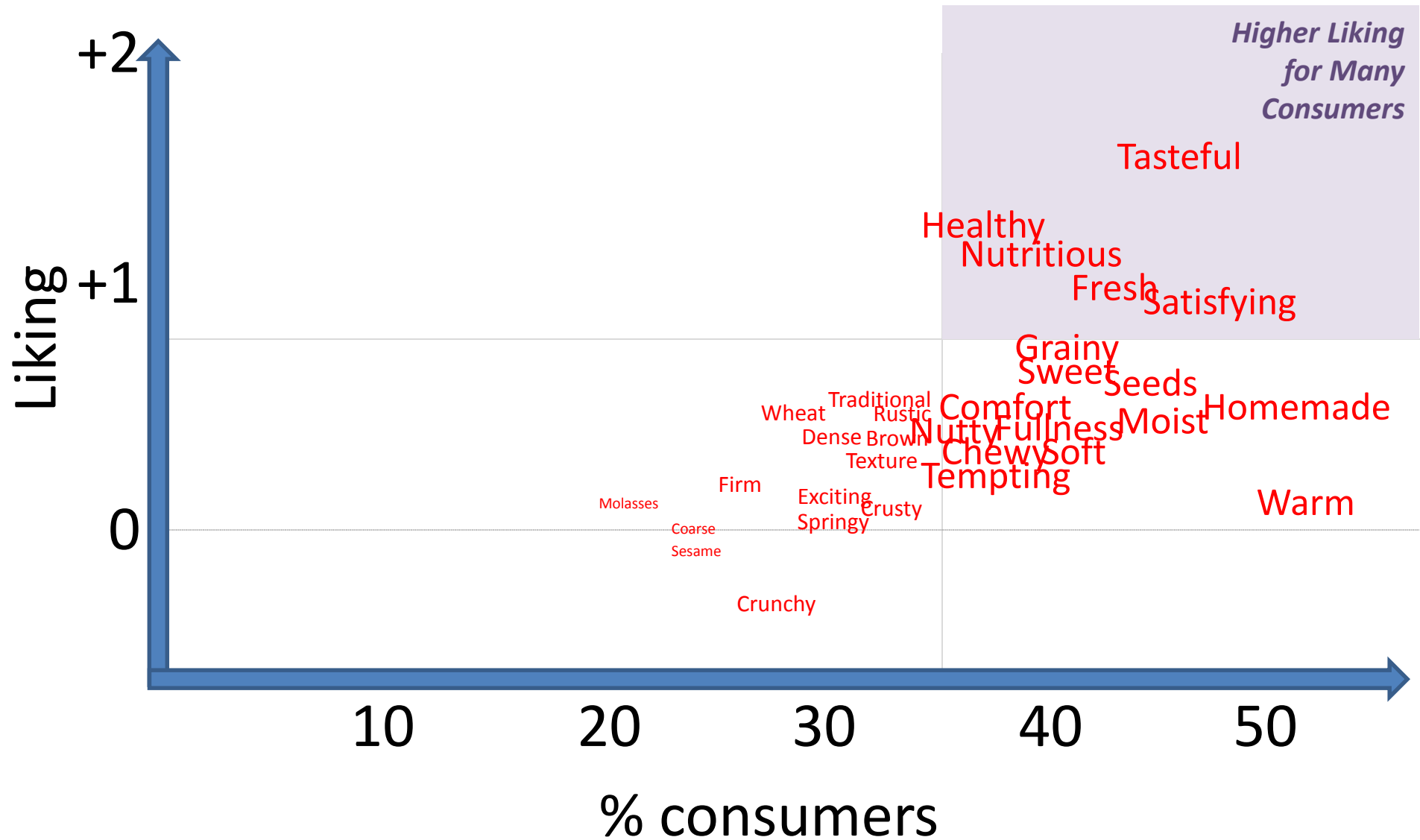


Comparison with Ideal

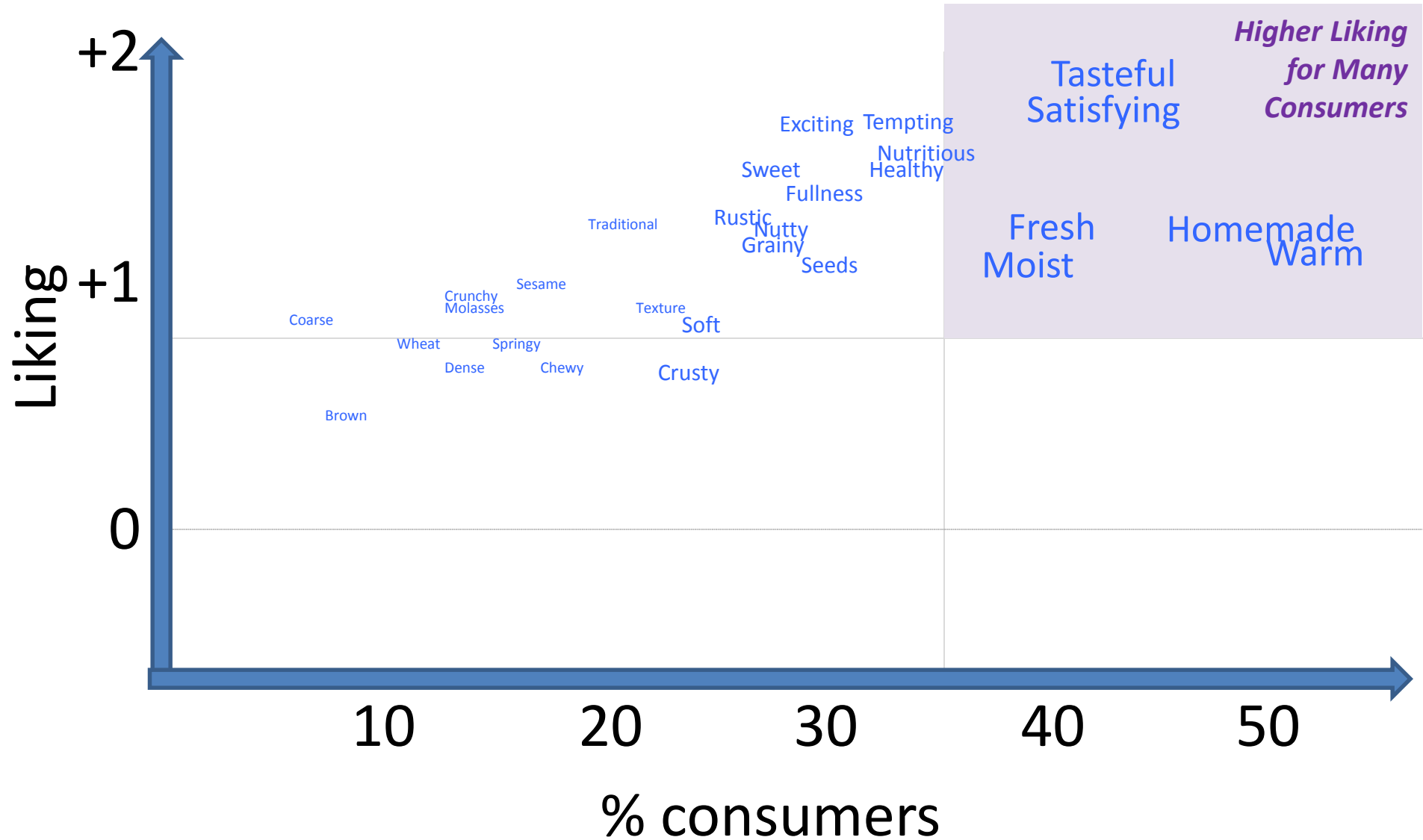
Problem: How to modify a product to delight consumers?

Solution 2: Use penalty analysis to identify attributes that drive the hedonic response

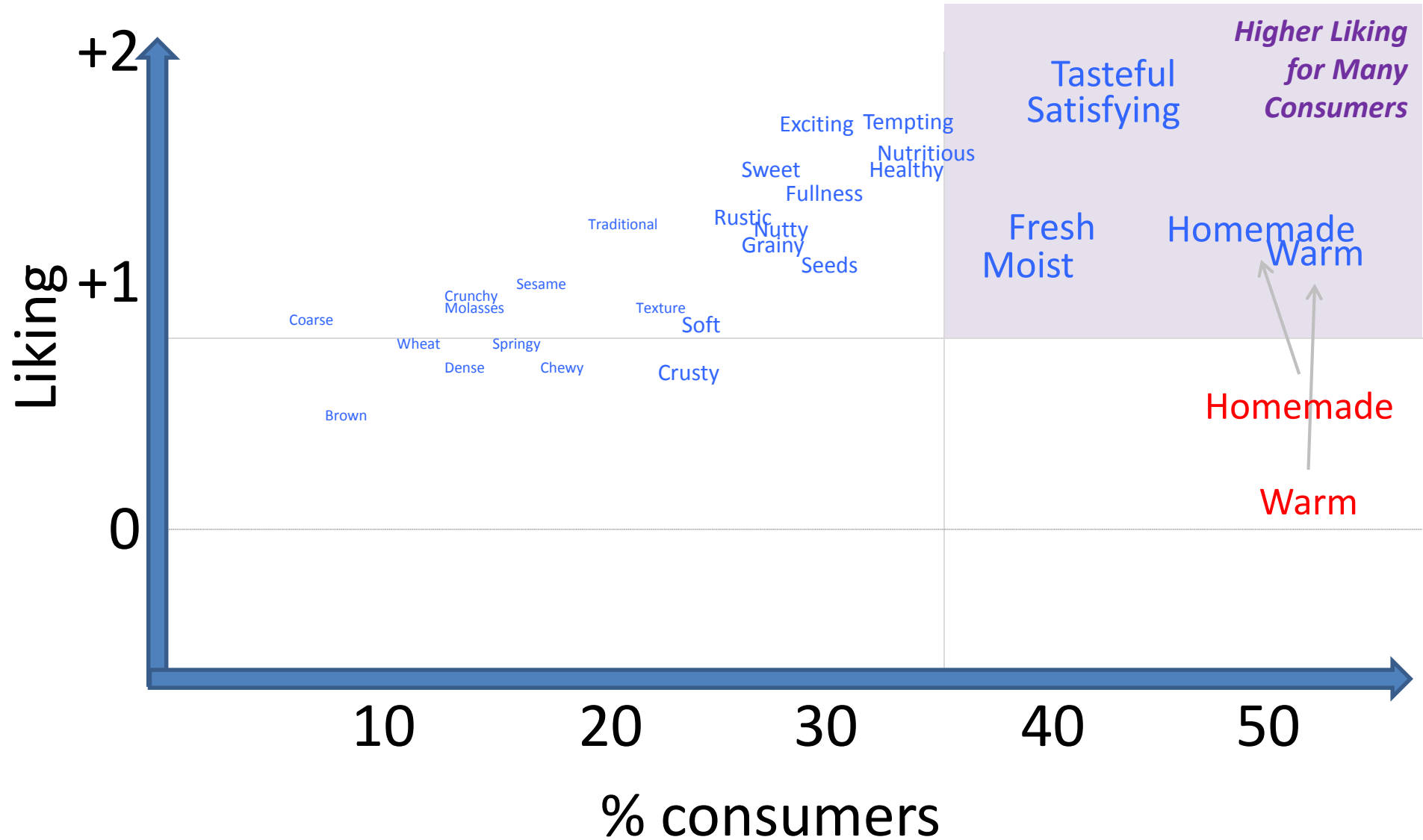
Penalty analysis based on general incongruence



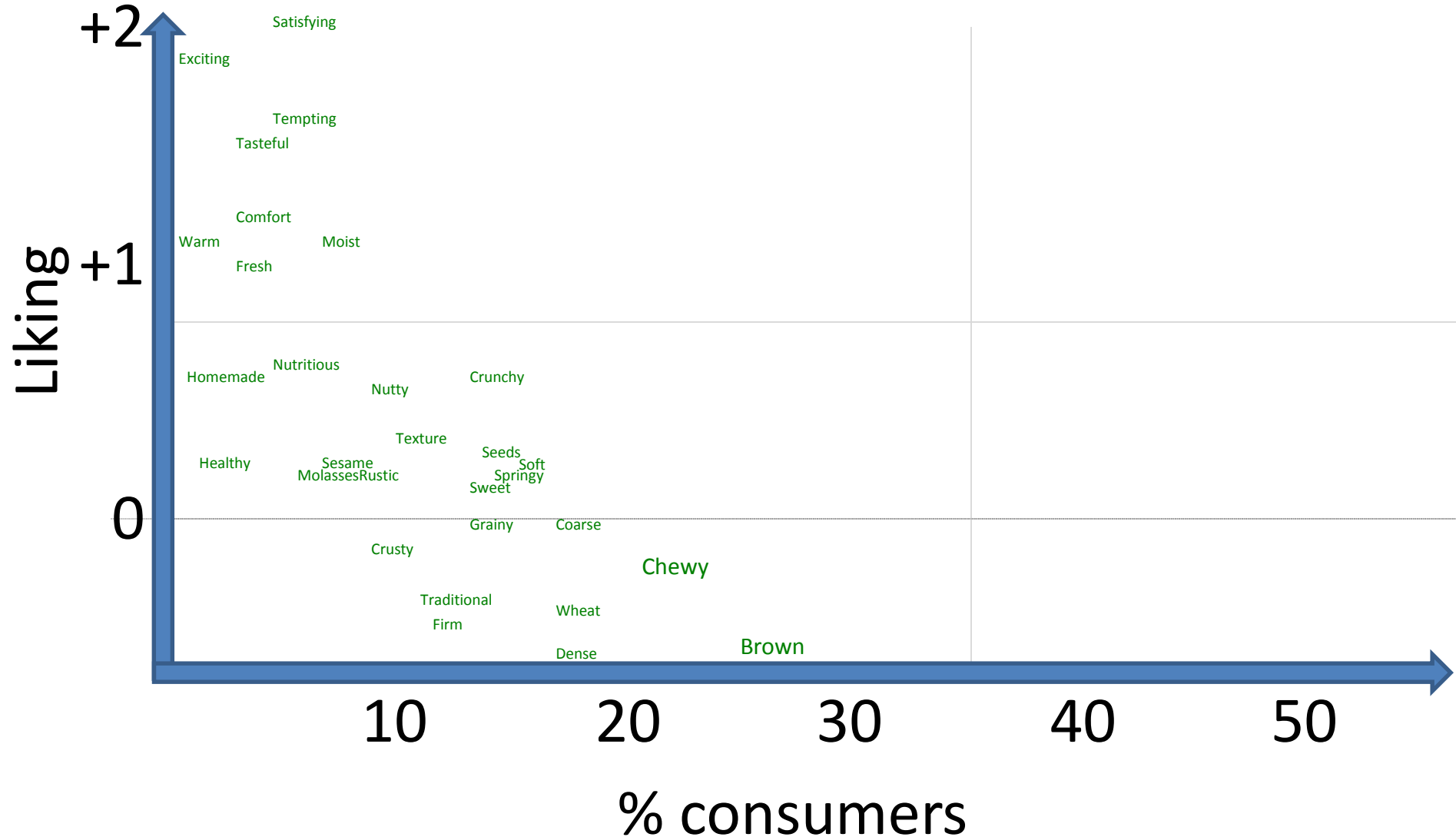
Penalty analysis based on “must have” attributes



Penalty analysis based on “must have” attributes



Penalty analysis based on “nice to have” attributes



Gain insights from CATA data

- Are products different?
- How do consumers perceive the products?
- What are the relationships between variables?
- What variables drive hedonic response?
- What variables should be the focus if attempting to improve a product?

Extensions

Methods discussed can also be applied to...

- other multivariate binary data
- response data other than liking

Thank you for your attention!

For more information see the following publication and references therein:

Meyners, M., Castura, J.C., Carr, B.T. (2013). **Existing and new approaches for the analysis of CATA data.** *Food Quality and Preference*, 30, 309-319.



Michael Meyners
Procter & Gamble



B. Thomas Carr
Carr Consulting



John C. Castura
Compusense Inc.

