Sensometrics has got to move on

Hal MacFie

AGROSTAT, Lausanne, March 21 to 24 2016

Overview

- Sensometrics needs impact
- Discrimination testing for behaviour
- Diversity and the non-fitting
- Decision Making
- The way forward

Most Downloaded Food Quality and Preference Articles

The most downloaded articles from ScienceDirect in the last 90 days.

The multisensory perception of flavor: Assessing the influence of color cues on flavor discrimination responses

Massimiliano Zampini[a, b, c], Daniel Sanabria[c, d], Nicola Phillips[e], Charles Spence[c]

Show more

How are organic food prices affecting consumer behaviour? A review

Manika Rödiger, Ulrich Hamm

Consumer attitudes and purchase intentions in relation to organic foods in Taiwan: Moderating effects of food-related personality traits

Mei-Fang Chen

In fact we have to get to number 11 before we get to anything remotely methodological and there are no Sensometrics papers in the top 30.

11. Consumer research in the early stages of new product development: a critical review of methods and techniques ↗

April 2005

Ellen van Kleef | Hans C.M. van Trijp | Pieternel Luning

And yet :

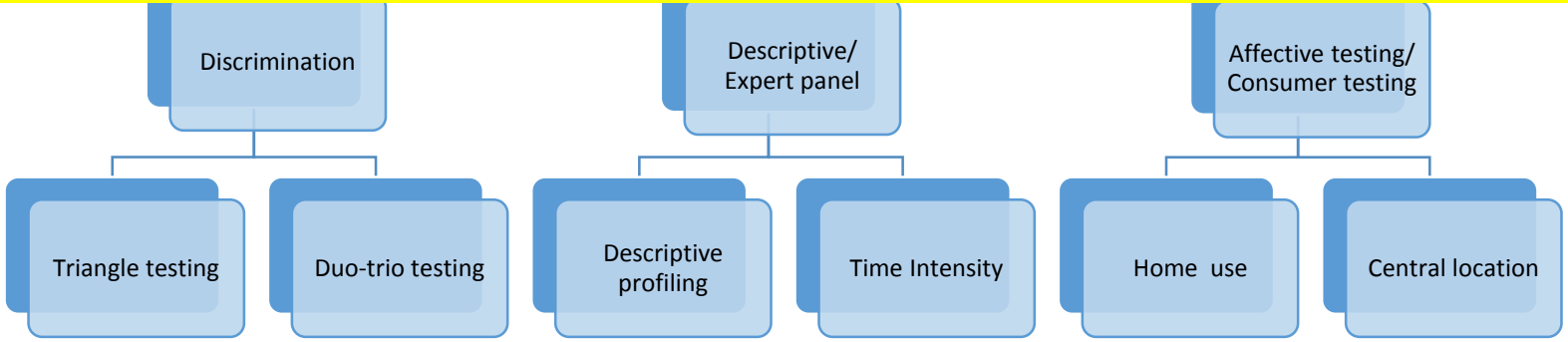
I believe we are often

- Measuring the wrong variables
- Using the wrong tests
- Using the wrong models

- Big opportunity for Sensometricians to show leadership and go to the top of the impact scoreboard.

Sensory Science –traditional view

CONSUMER BEHAVIOUR



Are these methods suitable for 21st century testing?

Proposition- To be impactful

- Sensometrics must use Statistical Theory and Tools to explain and model Consumer behaviour rather than maximising power

Acknowledge John Ennis and colleagues for their excellent research

Food Quality and Preference 32 (2014) 77–82

Contents lists available at SciVerse ScienceDirect

Food Quality and Preference

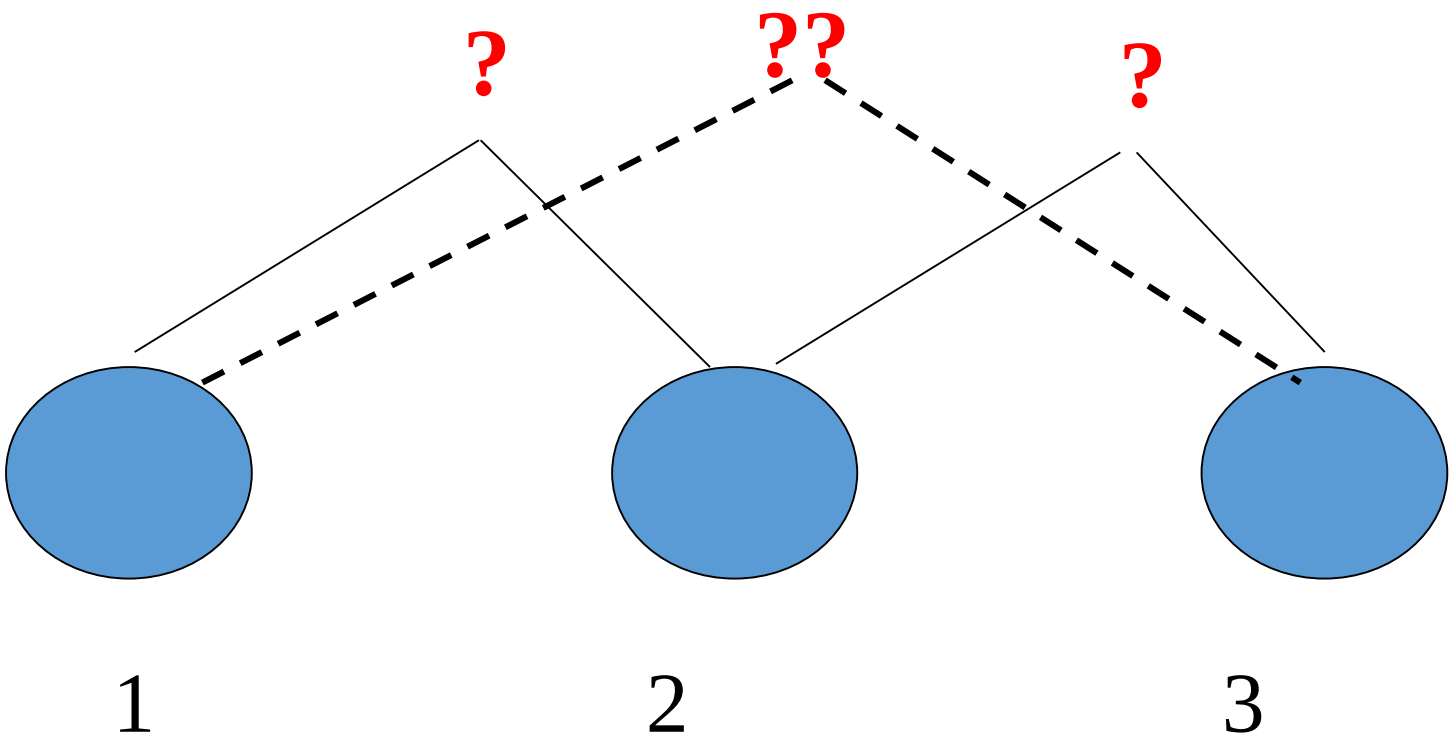
journal homepage: www.elsevier.com/locate/foodqual

Transitioning from proportion of discriminators to a more meaningful measure of sensory difference

Virginie Jesionka [a], Benoît Rousseau [b], John M. Ennis [c,*]

[a] SKIM Group, Weena 695-B6.036, Rotterdam, The Netherlands
[b] The Institute for Perception, 2306 Anza Avenue, Davis, CA 95616, USA
[c] The Institute for Perception, 7629 Hull Street Road, Richmond, VA 23235, USA

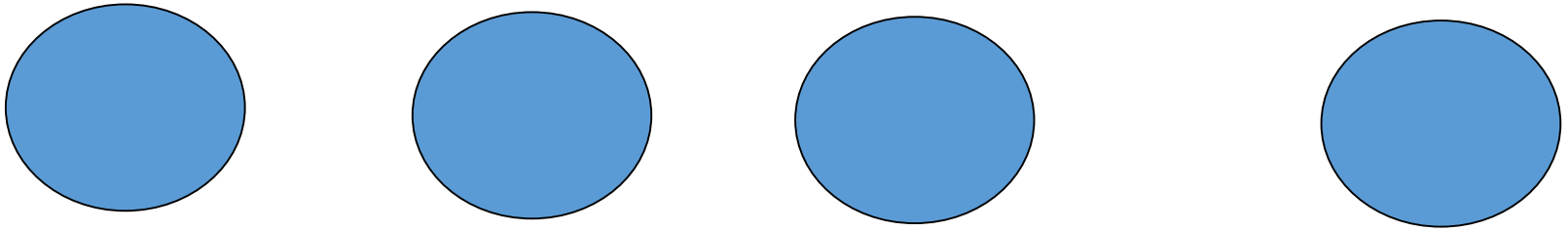
We are all familiar with the Triangle Test



Which One is Different?

Used throughout the Food, Beverage and Personal Products industry to see if people can notice a difference

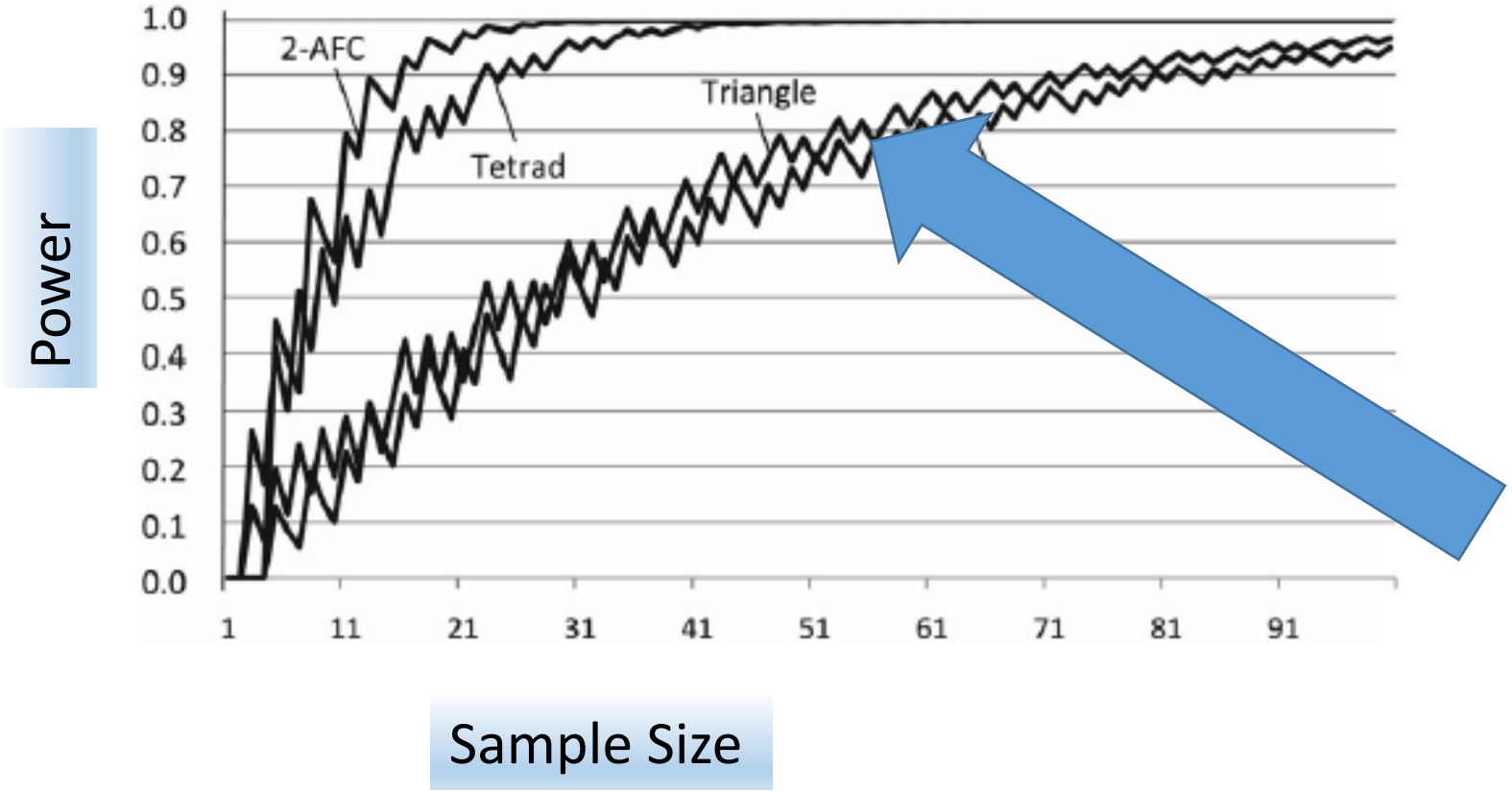
The Tetrad test proposed by Ennis



1 2 3 4

Pick the two pairs of samples that are most similar to each other

Comparing the power of different tests at a delta of 1.5



Power as a function of sample size for the 2AFC, Duo Trio, Triangle and Tetrad for a delta of 1.5 (fairly easily discriminable)

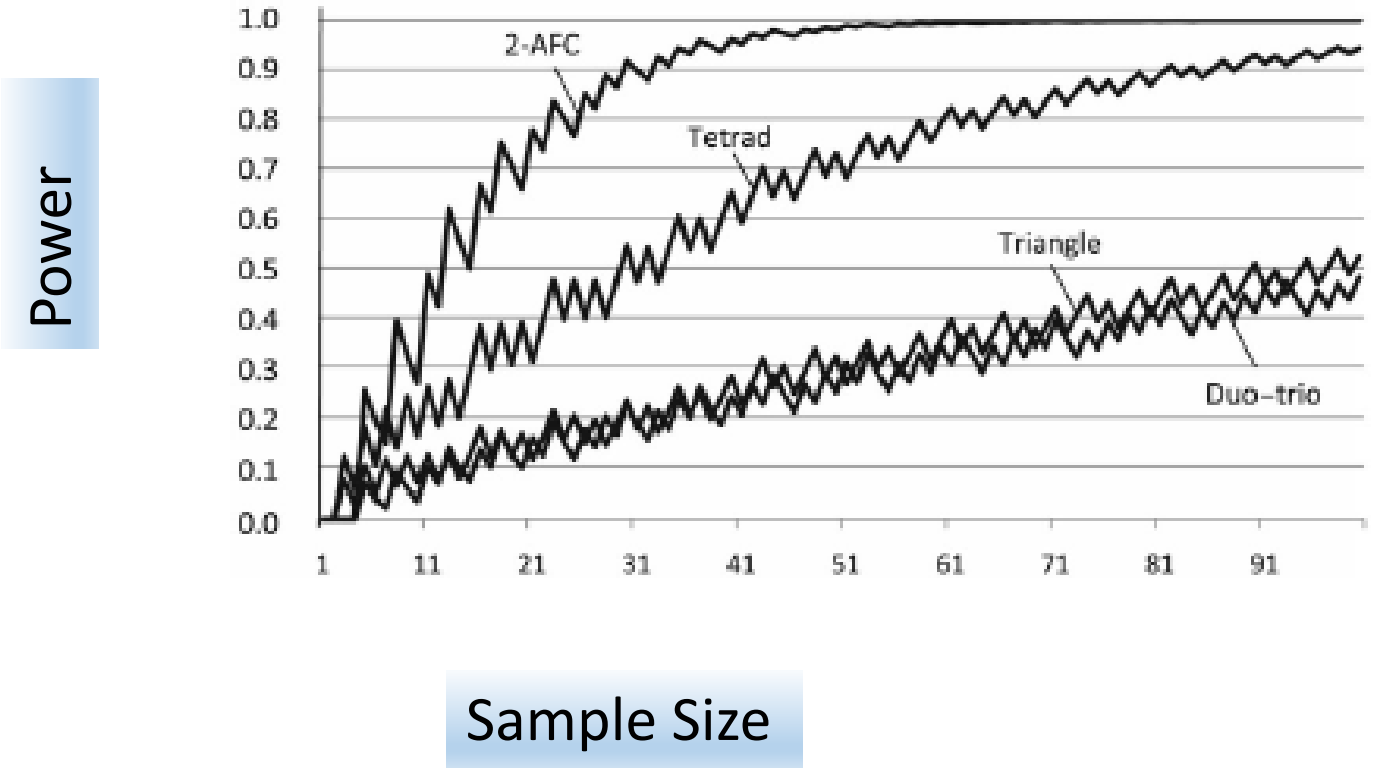
Delta is an index of signal to noise so at a Delta of 1 the signal is about the same as the noise

So for a power of 0.8 you can replace a Triangle test with 60 people by a Tetrad test using 15 people

Big saving!

General Mills have adopted Tetrad instead of Triangle

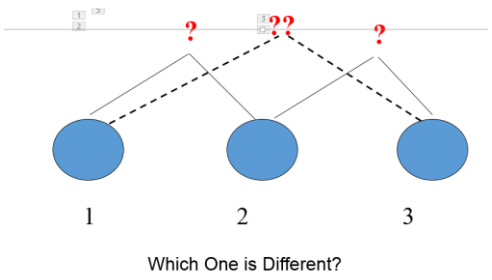
Comparing the power of different tests at a delta of 1.0



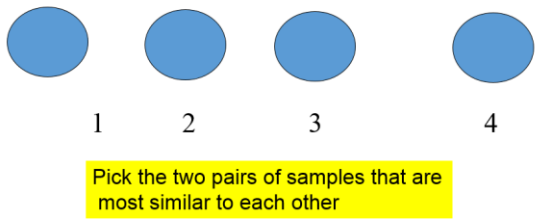
The Tetrad performs much better than the Triangle test for very noisy studies

Power as a function of sample size for the 2AFC, Duo Trio, Triangle and Tetrad for a delta of 1.0 (sensory difference equals noise)

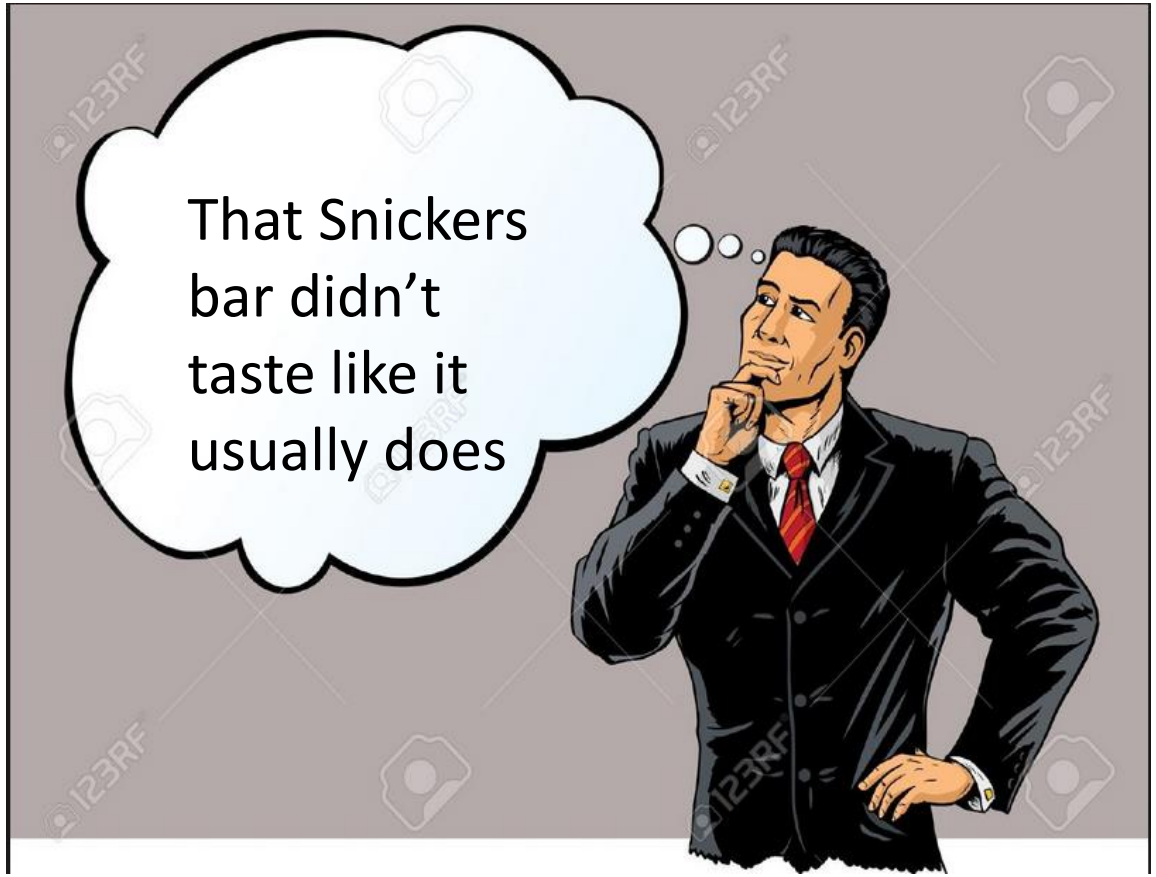
And yet what are we trying to measure here?



And yet what are we trying to measure here?



And yet what are we trying to measure here?



A not A
Is this product Snickers or not?

A not A testing

- Please taste this product and tell if you think it is your usual Snickers bar or not

Reminder A not A testing

Here is a normal Snickers bar please taste it
Now please taste this product and tell if you think it is your usual Snickers bar or not

Testing for formulation change

- Behaviours being measured
  - Consume product and notice difference
  - Consume product repeatedly and then notice difference
  - Consume product repeatedly and then change buying behaviour, sometimes without conscious recognition that difference has been perceived

- Testing Protocol
  - A not A without tasting the reference
  - Reminder A not A where the reference is given first

We have successfully applied this approach to a commercial shelf-life study

Compared on the basis of number of samples used –A not A reminder can beat Tetrad

Contents lists available at ScienceDirect

Food Quality and Preference

journal homepage: www.elsevier.com/locate/foodqual

Sensory discrimination by consumers of multiple stimuli from a reference: Stimulus configuration in A-Not AR and constant-ref. duo-trio superior to triangle and unspecified tetrad?

Yu-Na Jeong [a,1], Bi-A Kang [a,1], Min-Ju Jeong [a], Min-Jeong Song [a], Michael J. Hautus [b], Hye-Seong Lee [a,*]

[a] *Department of Food Science and Technology, College of Engineering, Ewha Womans University, Seoul 120-750, South Korea*
[b] *School of Psychology, The University of Auckland, New Zealand*

Reminder methods work better than Triangle and Tetrad on this basis

Results

Comparison of test performances

Comparison of ratio of estimated values of *d'* for the three methods to the Triangle method.

	Subjects	Stimuli pair	Group 1 (N=60) A-Not AR using τ-strategy	Group2 (N=60) Constant Ref. DTF/ 2-AFCR using COD	Group3 (N=60) Tetrad
Session 2	LSS	T_0vsT_{1+2+}	1.55	1.80	1.25
	HSS	T_0vsT_{1+2+}	1.29	1.08	0.79
	Pooled	T_0vsT_{1+2+}	1.33	1.30	0.89
			Better than Triangle		
			1.49	1.46	

÷

Ratio of estimated values of *d'*
for the reminder methods to the tetrad method
was close to 1.5

Behaviour versus Power

- Since A not A task closely models the thought patterns we are trying to measure we can easily extend the testing to measure responses in different environments
- For example in a bar during a session where a reformulated drink might be presented straight after an original prototype

Snickers Story

- Chocolate removed to save money
- More removed as triangles showed no perceived difference
- After 6 months sales started to dip

Question

- If the reformulation changes are small how does the perception change with time
- What is the updating process between expectation and actual?
- **<u>Needed: repeated assessment trials</u>**

The effect of a conscious change

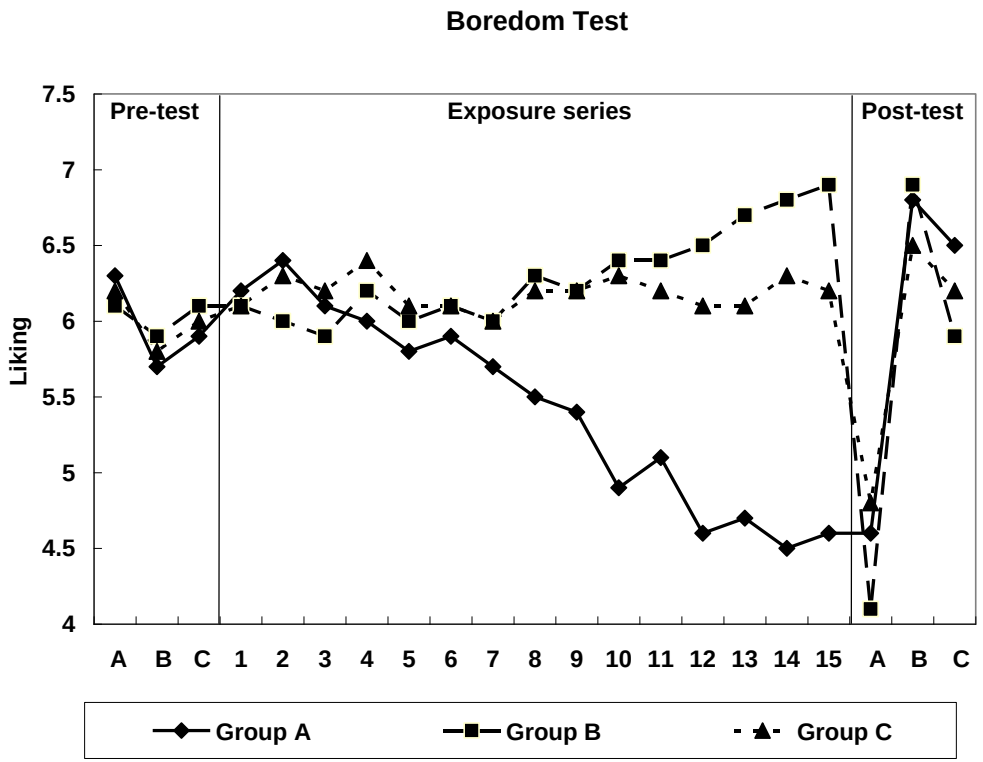
Rick Mattes Am J Clin Nutr. *1993* Mar; 57(3):373-81. Fat preference and adherence to a reduced-fat diet

- Respondents were asked to switch to a low fat diet for health reasons
- For the group that did not use modified fat products, Hedonic liking for full fat foods dropped across a 12 week period and this was maintained for a further 12 week follow up

Question

- What is the updating process between liking and expectation
- **<u>Need – repeated assessment trials</u>**

Boredom Test Example



The stimulus assessed first by subgroup C was not well discriminated at first

Points so far

- Select the test that Measures behaviour above power
- Repeated assessments needed to understand decision making after reformulation

Complete the sentence….

Complete the sentence….

Complete the sentence….

Complete the sentence….

Diversity

- Koster notes that we must acknowledge diversity of response

ELSEVIER

Food Quality and Preference

Volume 47, Part A, January 2016, Pages 54–63

Sensometric 2014: Data That Works In The City That Works

Segmentation of consumers in preference studies while setting aside atypical or irrelevant consumers

E. Vigneau[a, b], E.M. Qannari[a, b], B. Navez[c], V. Cottet[d]

Show more

doi:10.1016/j.foodqual.2015.02.008

Get rights and content

- Vigneau and others propose a “noise” (Idiosyncratic) cluster of all those who don’t fit in any of the main cluster

Diversity

- Typically this noise cluster can have 20% to 30% of respondents in it
- Not a very satisfactory model of consumer behaviour
- Who are these people?
- Are they poor sensory discriminators?
- How much do they vary on repeated sampling
- How do they make choices in practice?
- **Needed: Repeated assessment studies on the non-fitting!**
- **Needed: Sensory sensitivity testing on the non-fitting**

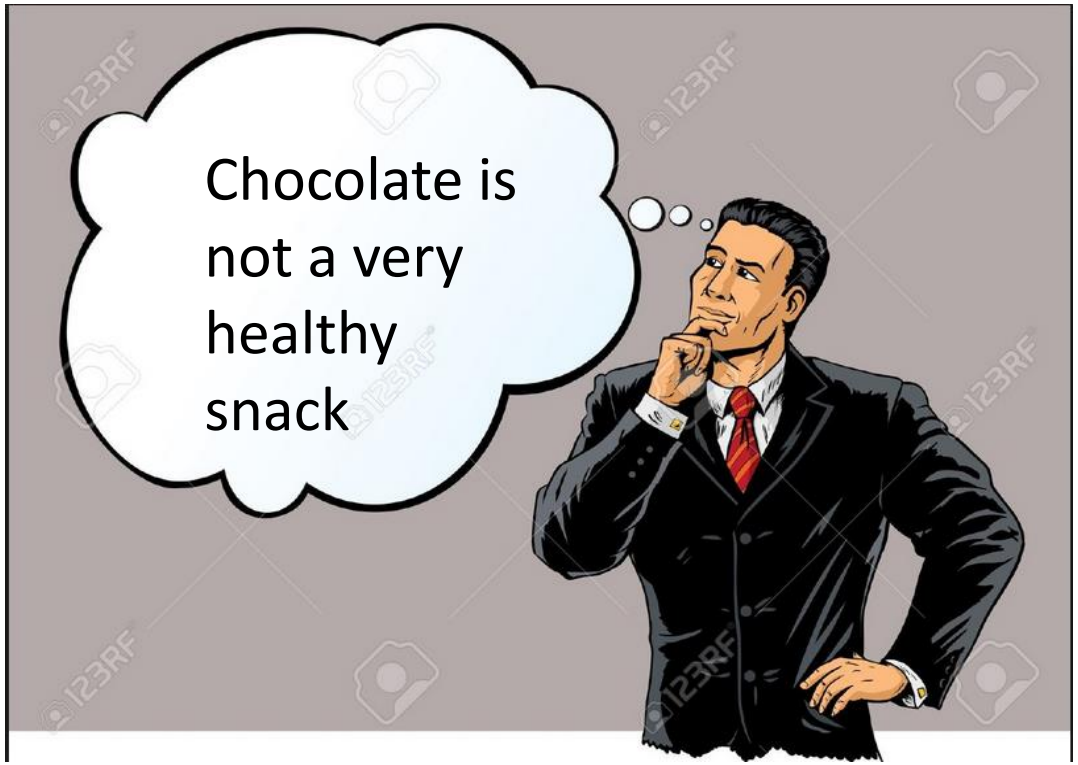
Decision making models from Psychology and Economics

Dual processing

- From Plato to Freud, many have remarked that humans seem to have more than one mind.
- Today, detailed 'dual-process' models are being tested by psychologists and neuroscientists:

- Since the 1970s dual-process theories have been developed to explain various aspects of human psychology...
- One of the processes is characterized as fast, effortless, automatic, non-conscious, inflexible, heavily contextualized, and undemanding of working memory
- The other as slow, effortful, controlled, conscious, flexible, decontextualized, and demanding of working memory.

Decision making – Two minds - Dual Processing



One process is slow, effortful, controlled, conscious, flexible, decontextualized, and demanding of working memory.

One of the processes is characterized as fast, effortless, automatic, non-conscious, inflexible, heavily contextualized, and undemanding of working memory

Most popular Dual processing theory Kahneman and Tversky

- Daniel Kahneman provided further interpretation by differentiating the two styles of processing more, calling them intuition and reasoning in 2003.
- Intuition (or system 1), similar to associative reasoning, was determined to be fast and automatic, **usually with strong emotional bonds included in the reasoning process.**
- Kahneman said that this kind of reasoning was based on formed habits and very difficult to change or manipulate.
- Reasoning (or system 2) was slower and much more **changeable**, being subject to conscious judgments and attitudes

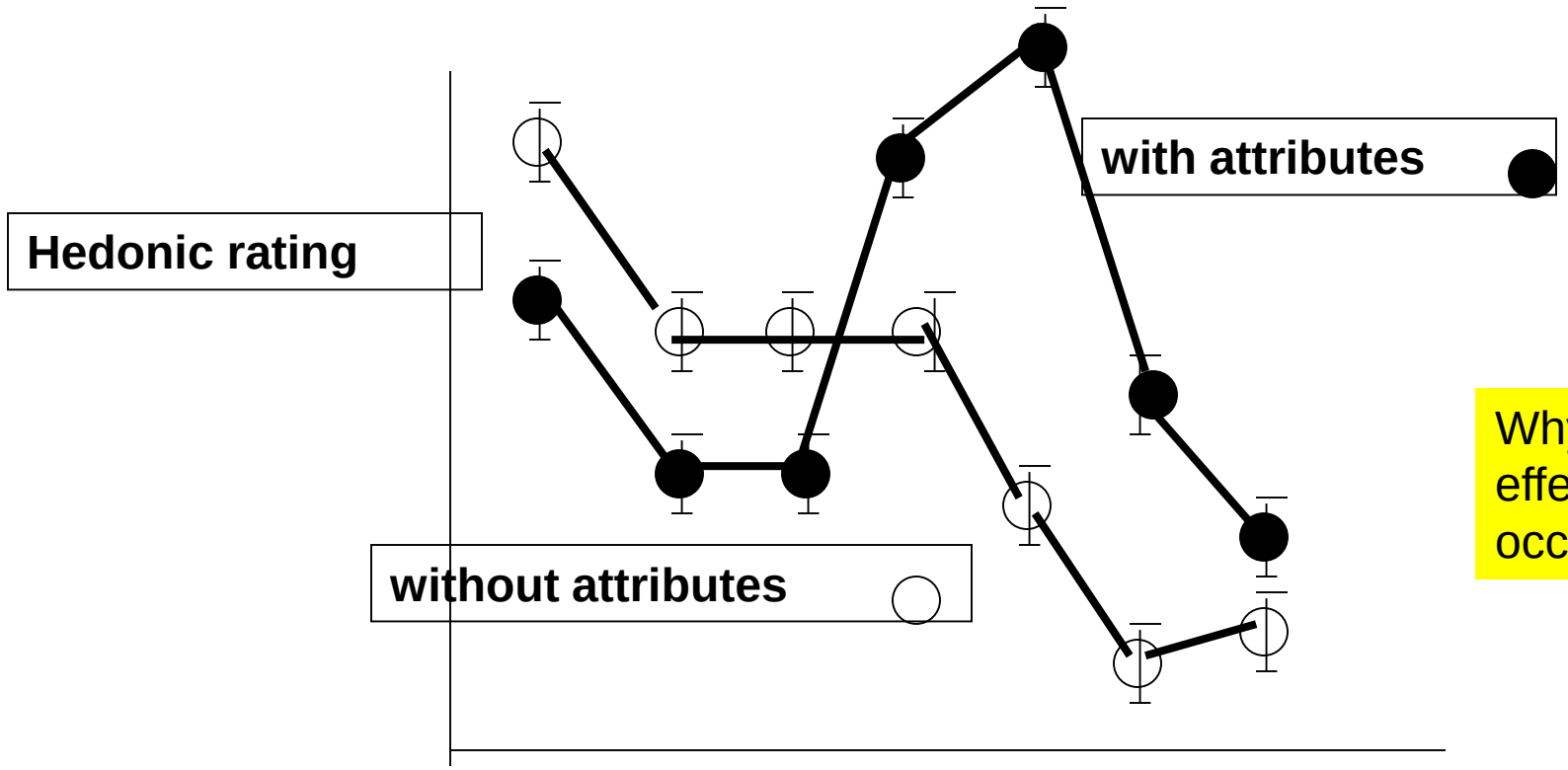
The key point is that system 1 is low effort and rapid and is used very often in decision making.

System 2 gets called in when system 1 cannot cope

System 1	System 2
Unconscious reasoning	Conscious reasoning
Implicit	Explicit
Automatic	Controlled
Low Effort	High Effort
Large capacity	Small capacity
Rapid	Slow
Default Process	Inhibitory
Associative	Rule based
Contextualized	Abstract
Domain Specific	Domain General
Evolutionarily Old	Evolutionarily recent
Nonverbal	Linked to language
Includes recognition, perception, orientation	Includes rule following, comparisons, weighing of options
Modular Cognition	Fluid Intelligence
Independent of working memory	Limited by working memory capacity
Non-Logical	Logical
Parallel	Serial

Earthy, P. J., MacFie, H. J. H. and Hedderley, D. (1997) Effect of question order on sensory perception and preference in central location trials. *J. Sens. Stud*., **12**(3), 215-237.

Does the addition of attribute questions alter the hedonic ratings?



Why has this effect occurred?

milk%	100	75	75	50	25	25	0
dark%	0	25	25	50	75	75	100
sugar gms	9	18	0	9	0	18	9

Implications

- Elaborate and long questionnaires may cause us to measure system 2 (conscious) when the true behaviour is system 1 (automatic)
- A respondent may use both system 1 and system 2 decision making in different contexts or from one tasting to another
- In the long term system 2 conscious decision making may dominate (reduced fat example)
- The effect of repeated consumption is not understood
- **Needed: repeated assessment trials**

ELSEVIER

Available online at www.sciencedirect.com
ScienceDirect
Food Quality and Preference 20 (2009) 70–82

Food Quality and Preference
www.elsevier.com/locate/foodqual

Diversity in the determinants of food choice: A psychological perspective

E.P. Köster *

- Eating, Drinking and Food choices are among the most frequent human behaviours
- Intuitive thinking and decision making are the rule rather than the exception in these behaviours
- Many of our measurement methods do not acknowledge this
- Even when asked a simple question such as “Do you like this? They detect nuances of taste that they never noticed and system 2 goes on to alert
- Situational analysis, observational methods and memory and expectation research will give more insightful results

Currently many groups and researchers are starting to explore how to measure unconscious processing
: Facial coding, Weighed intake, Speed of intake
Important to remember that conscious processing is also important

Decision Making – Heuristics

Heuristics are strategies we use to deal with problems

- When confronted with a difficult question, people may answer an easier one instead and are often unaware of the substitution.

- A person who is asked "What proportion of long-distance relationships break up within a year?" may answer as if she had been asked "Do instances of failed long-distance relationships come readily to mind?" <u>This would be an application of the availability heuristic.</u>

- A professor who has heard a candidate's job talk and now considers the question "How likely is it that this candidate could be tenured in our department?" may answer the much easier question: "How impressive was the talk?" <u>This would be an example of one form of the representativeness heuristic.</u>

Consumer Science Heuristics

The question: “How does repeated consumer experience of consuming or using a new or modified product influence repurchase”

has been substituted by

“How does the sensory experience of 10 people influence declared liking after one tasting/use by consumers”

Does liking predict behaviour?

HOW WELL DOES THE 9-POINT HEDONIC SCALE PREDICT PURCHASE FREQUENCY?

M. ROSAS-NEXTICAPA[1], O. ANGULO[1] and M. O'MAHONY[2,3]

Journal of Sensory Studies **20** (2005) 313–331. *All Rights Reserved.*
© Copyright 2005, Blackwell Publishing

- 109 consumers assessed 3 strawberry flavoured yogurts using the 9 point hedonic scale
- They were then monitored for a year to determine whether their ratings had any predictive value for their purchase behaviour
- Highest rated yoghurts tended to be purchased
- Negligible correspondence between rank order of rating and purchase frequency

Rating, Ranking, Take-Away, Meal context

Food Quality and Preference 27 (2013) 63–71

Contents lists available at SciVerse ScienceDirect

Food Quality and Preference

journal homepage: www.elsevier.com/locate/foodqual

A comparison of two new take-away strategies and their relation to rating and ranking of extrinsic properties of dry cured ham

Tormod Næs [a,c,*], Erminio Monteleone [b], Anne Segtnan [a], Margrethe Hersleth [a]

- Test 1 – 32 consumers ranked 6 packages of cured ham for meal context 1 and then meal context 2
- They then rated prob buying each package
- Take away – each consumer was given 25 euros and asked to buy two packages of ham with priority 1 and priority 2 and to indicate which kind of meal
- Test 2 - 120 consumers

For ranking and rating, if you took it away then it was very likely to have been rated as either the best or the next best
Similar results for test 2

Implications

- We do not need a complex quadratic model linking liking to sensory
- Just use Pick a winner task eg Hedonext

POSTER 326

COMPARISON OF HEDONIC SCORES WITH THE NEW HEDONEXT CONSUMER APPROACH: A CASE STUDY ON SELECTIVE SKIN CARE PRODUCTS

Wantz, Nathalie[1] – Loescher, Eleonore[1] – Blumenthal, David[2] – Sieffermann, Jean-Marc[2] – Gazano, Germaine[1]

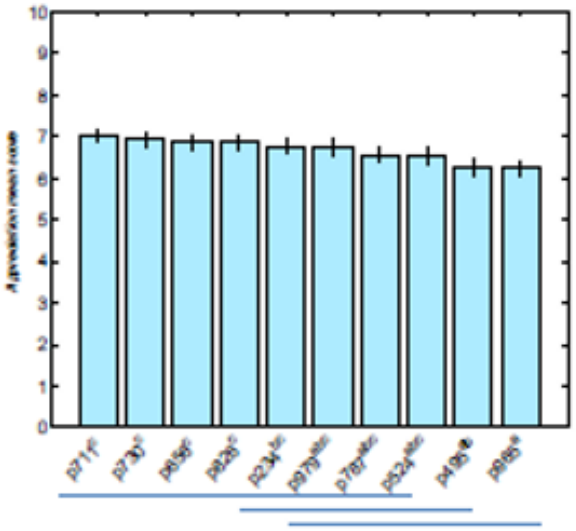
[1] LVMH Recherche F-45804 Saint-Jean de Braye CEDEX, France
[2] Laboratoire Perception Sensorielle et Sensométrie, AgroParisTech-Massy, 1 avenue des Olympiades, F-91744, Massy, France

Methodology

- 250 women assessed the products monadically at home giving a global hedonic score on a 0 to 10 scale
- 60 women were asked, after each product test, if they were satisfied enough with the current product. If so they were rewarded with a month’s supply and their trial was terminated.

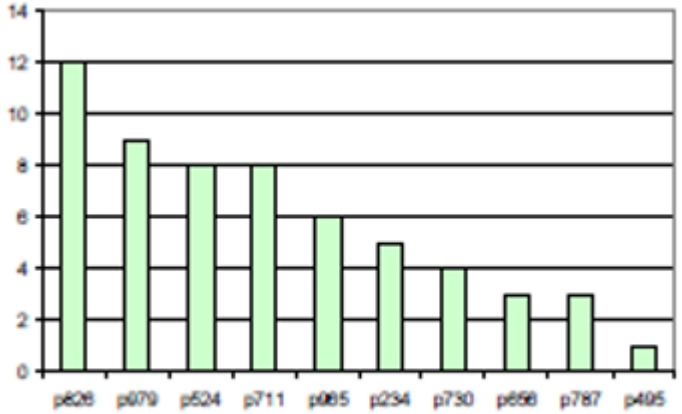
Comparing Average Liking with Hedonext

Average liking scores
No clear winner



Graph 1: Mean score of appreciation and significant differences between products

Hedonext
Clear winner on a genuine behaviour



Graph 2: Number of times each product was chosen

Marketing Proposition

- 1. Top 2/3 box is a better indicator of likely purchase behaviour than Mean Liking
- 2. Products will only be bought by those that score top 2/3 box

Yoghurts

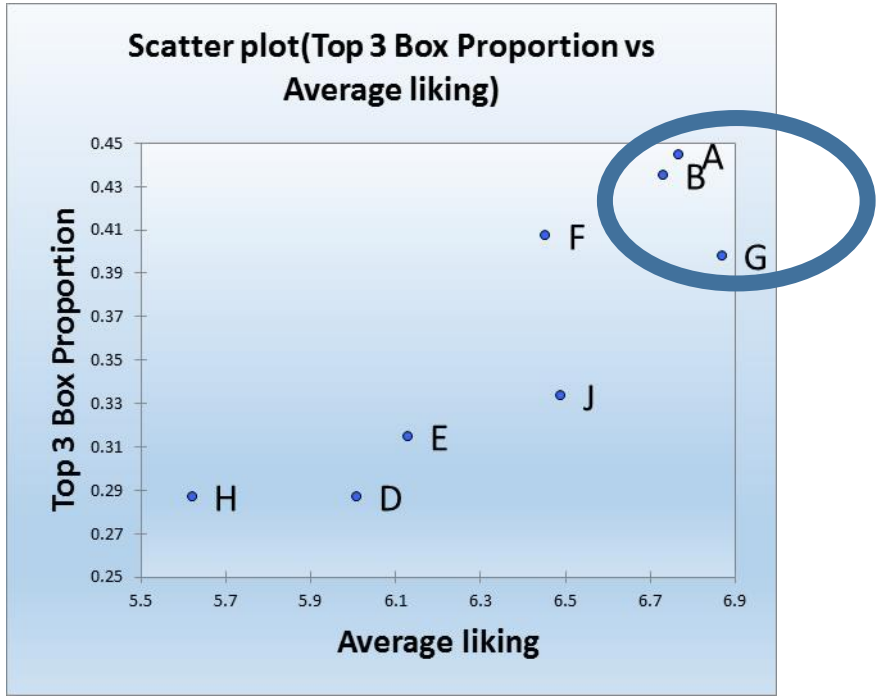
Anova of Liking shows more discrimination than k proportions test on top 3 box

Sample	Average liking	Groups			
G	6.9	A			
A	6.8	A			
B	6.7	A			
J	6.5	A	B		
F	6.5	A	B		
E	6.1		B	C	
D	6.0			C	D
H	5.6				D

Sample	Top 3 Box Proportion	Groups
A	0.44	A
B	0.44	A
F	0.41	A
G	0.40	A
J	0.33	A
E	0.31	A
D	0.29	A
H	0.29	A

Yoghurts

Plotting Top 3 Box proportion versus Mean Liking indicates that A and B will be selected by marketing because of the higher Top 3 box ratings.
Sample G will be recommended by the product developer working on select the best from Average liking



Visions of Cognition

- Superhuman
  - Unbounded Rationality
  - Optimization under Constraints
- Bounded Rationality

General findings

People tend to stop searching for cues and make a decision after the first discriminating cue: one reason decision making

People tend to learn which cues are most useful, and search for those first: Take The Best

Preference mapping versus Heuristic

SENSORY

LIKING

TAKE THE BEST

Points so far

- Select the test that Measures behaviour above power
- Repeated assessments needed to understand decision making after reformulation
- Dual processing may explain variation in response
- Heuristic choices may not be modelled by regression
- Need to model behaviour

Which US city has more inhabitants,
San Diego or San Antonio?

Which US city has more inhabitants, San Diego or San Antonio?

Americans:	*Germans:*
62%	*100%*
correct	*correct*

The Recognition Heuristic

When deciding which of two objects is greater on some criterion, if one object is recognized and the other is not, then infer that the recognized object has the higher value.

(e.g., if just San Diego is recognized, choose it)

We know Brand to be a crucial determinant of selection behaviour but 90% of our trials are done on blind products

Sensory researchers are starting to use more complex models with choice as the response

Food Quality and Preference 45 (2015) 140–148

Contents lists available at ScienceDirect

Food Quality and Preference

journal homepage: www.elsevier.com/locate/foodqual

What reported food-evoked emotions may add: A model to predict consumer food choice

Swetlana Gutjar [a,b,*], Jelle R. Dalenberg [a,d], Cees de Graaf [a,b], René A. de Wijk [a,c], Aikaterini Palascha [b], Remco J. Renken [a,d], Gerry Jager [b]

[a] Top Institute Food and Nutrition, P.O. Box 557, NL-6700 AN Wageningen, The Netherlands
[b] Division of Human Nutrition, Wageningen University, P.O. Box 8129, NL-6700 EV Wageningen, The Netherlands
[c] Food and Biobased Research, Consumer Science & Intelligent Systems, NL-6700 AA Wageningen, The Netherlands
[d] Neuroimaging Center Groningen, University Medical Center Groningen, Antonius Deusinglaan 2, 9713 AW Groningen, The Netherlands

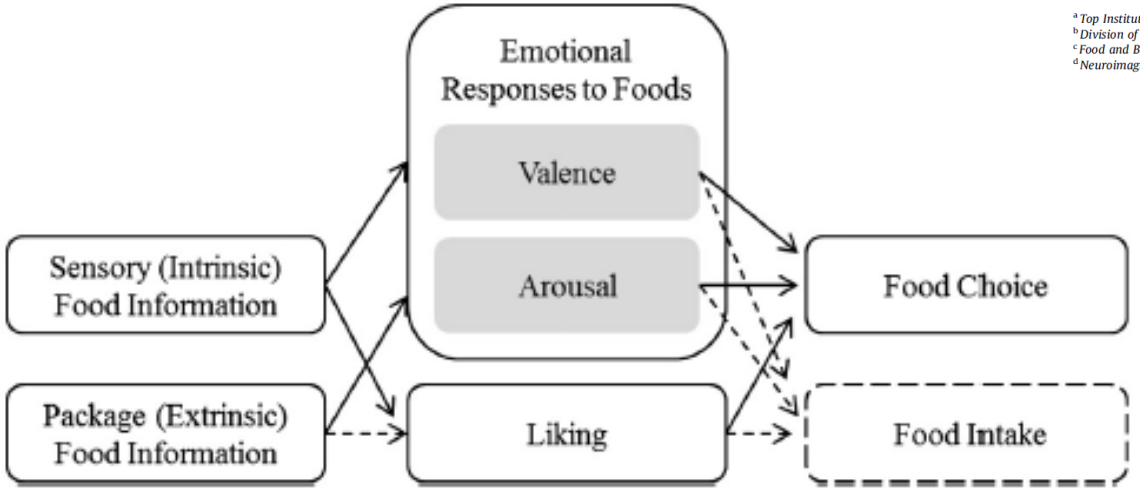


Fig. 1. Flow chart of the proposed model to illustrate how food-evoked emotional responses and liking contribute to food choice. In this model, sensory and packaging information are antecedent to liking and emotional responses (which can be decomposed into a valence and arousal dimension) to foods, subsequently both influence food choice and intake. Solid lines demonstrate the relationships investigated in this study. Broken lines and a broken line box are added to complete the model but were not investigated in the current study.

The modelling of rank1 explains 40% of choices

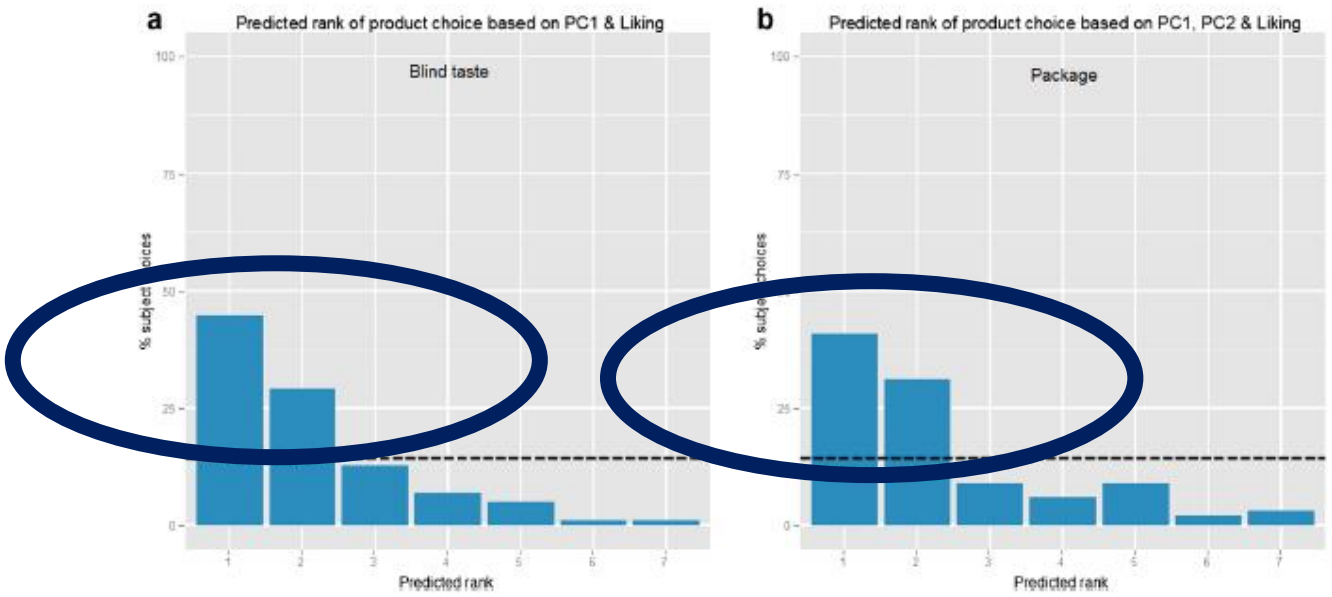


Fig. 4. Result of the LOOCV predictions. This figure shows the prediction outcomes of the LOOCV using different subsets of the predictors liking, valence (PC1) and arousal (PC2) per choice context (blind taste and package). In the figure we show the percentage of product choices per predicted rank (e.g. blind product choice for 44.7% of the participants was (correctly) predicted as rank 1 by the model when using liking & valence (PC1) as predictors). The dashed line indicates how the model would perform on chance level (14.7%).

Better model needed

Charles Spence (Psychology, Univ Oxford)

- What is the absolute basis of Sensory Perception?

- ATTENTION

TWO COFFEES are presented that do differ perceptually (A not A) if pay attention

GOOD MOOD OR BUSY ELSEWHERE

NORMAL
EXPECTATION SPACE

System 1 unconscious is not on alert, System 2 is not called in and is on other business

BAD MOOD OR PAYING ATTENTION

NORMAL EXPECTATION

System 2 is more critical and brings attention to bear on the sensory properties of the product
The Normal range is much smaller

<u>Experimental Conditions</u>

Repeated measurement

Natural behaviour

Psychographic Variables

SUB CONSCIOUS PERCEPTION HEDONIC AND EMOTION RESPONSE

CONSCIOUS PERCEPTION HEDONIC AND EMOTION RESPONSE

CONTEXT CUES

HEURISTIC STRATEGY

BRAND CUES

LABEL CUES

SENSORY SENSITIVITY

Answer the question:

“How does repeated consumer experience of consuming or using a new or modified product influence repurchase”

Not

“How does the sensory experience of 10 people influence declared liking after one tasting/use by consumers”

With a 30% bundle of irrelevant consumers!