

The Authenticity Test



A joint venture of CSIRO & the Victorian Government

*Investigating a possible new approach in
Sensory Research*

Ciarán Forde and Patrick O’Riordan

Sensory and Consumer Science

Food Science Australia



Emlyn Williams

Statistical Consulting Unit, ANU, Canberra

39th AIFST Conference, Adelaide

Tuesday July 11th, 2006

Background

- Traditional consumer test methods such as difference or preference testing are often criticised as being:

- Low emotional involvement
- Ecologically unrepresentative
- “Scoring or boring”
- Low predictive value - do not represent “real choice”



- Food knowledge is “implicit” – collected through unconscious experience

- **Hypothesis:** Maybe possible to use “indirect methods” that don’t draw on intentionally recalled experiences

Possible Approaches

■ Situation/Context:

- Situation Analysis
- Telling stories, images – creating a context



■ Behavioural Frequency:

- “do you like this cheese?” / “how often do you eat cheese?”

■ Indirect Questions:

- “Do you like...?” / “Have you ever...?”

■ Authenticity Methods (Mojet/Koster):

- *“We have invented a much cheaper way of brewing beer...”* (Mojet & Koster)
- *“We have found a cheaper supplier of milk...”*

Examples of Authenticity Tests

■ Example of milk

- Real differences: level of fat, cattle food, storage conditions
- Cover-up story about patriotic beliefs (Danish vs. German)
- Results demonstrated that Authenticity approach could find differences not found by descriptive analysis

(Wolf-Frandsen *et al* 2003 and 2005)



■ Test conducted in FSA with lettuce

- Cover up story: preservative spray
- No significant differences found:
 - Differences in the Samples were sub-threshold
 - HOWEVER – participants were really influenced by story



→ Sensitivity of tests **may** be improved

→ Possible implications for sensory testing with consumers (difference/preference)

Our Authenticity test...



- **ESN Task Group 2: Innovations in Sensory Research**
 - 3 part study
 - Conducted in Wageningen by A+F with Ketchup (Mojet, Poelman & Koster)
 - Lessons have been incorporated into the FSA approach
 - Third part will be conducted in INRA (Sulmont & Issanchou)

- **How do you evaluate the efficacy of the Authenticity Approach?**
 - Between groups comparison (strict recruitment criteria to ensure validity)
 - Control Groups conduct normal difference/preference test
 - Test Group conduct the same tests after receiving a psychological bias

- ➔ **Null Hypothesis:** That test sensitivity will be improved through a psychological bias

Our Authenticity Test...

1. Select a suitable product for the large scale experiment

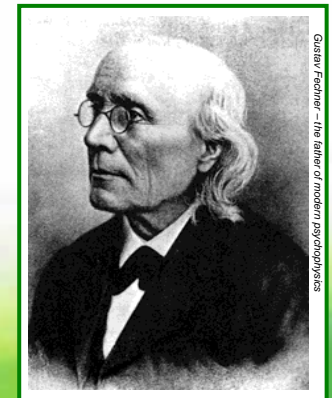
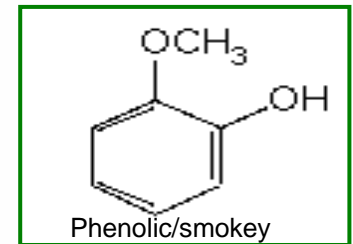
- High emotional involvement product ✓
- Easy modification of product ✓
- Originality ✓

2. Select a suitable stimulus

- Congruent – “*must fit well in the product*” ✓
- Easy to manipulate ✓
- Quantify flavour analytically ✓

3. Quantify a suitable level of the stimulus

- Psychophysics ✓
- Just Noticeable Difference* ✓
- To truly test the sensitivity of the Authenticity method – differences must be barely detectable ✓



How did we modify Vegemite...

■ *Once again: Criteria for selection of stimulus*

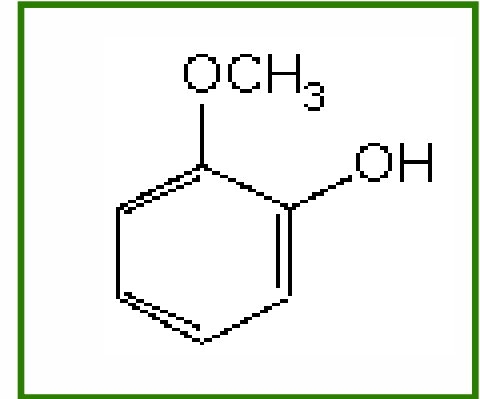
- Must be congruent
- Must be easy to incorporate
- Must be perceptible

■ *Different molecules have been tested*

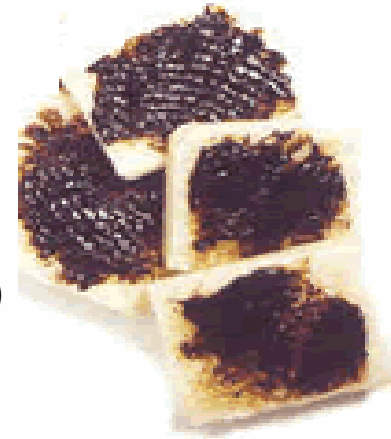
- Natural components: Salt, Citric acid, Sugar, MSG
- Flavour Compounds: Guaiacol, Cysteine, Methionine and Ethylbutyrate

■ *Guaiacol sensory attributes*

- Phenolic, medicinal odour
- Smoky taste
- Harsh aftertaste

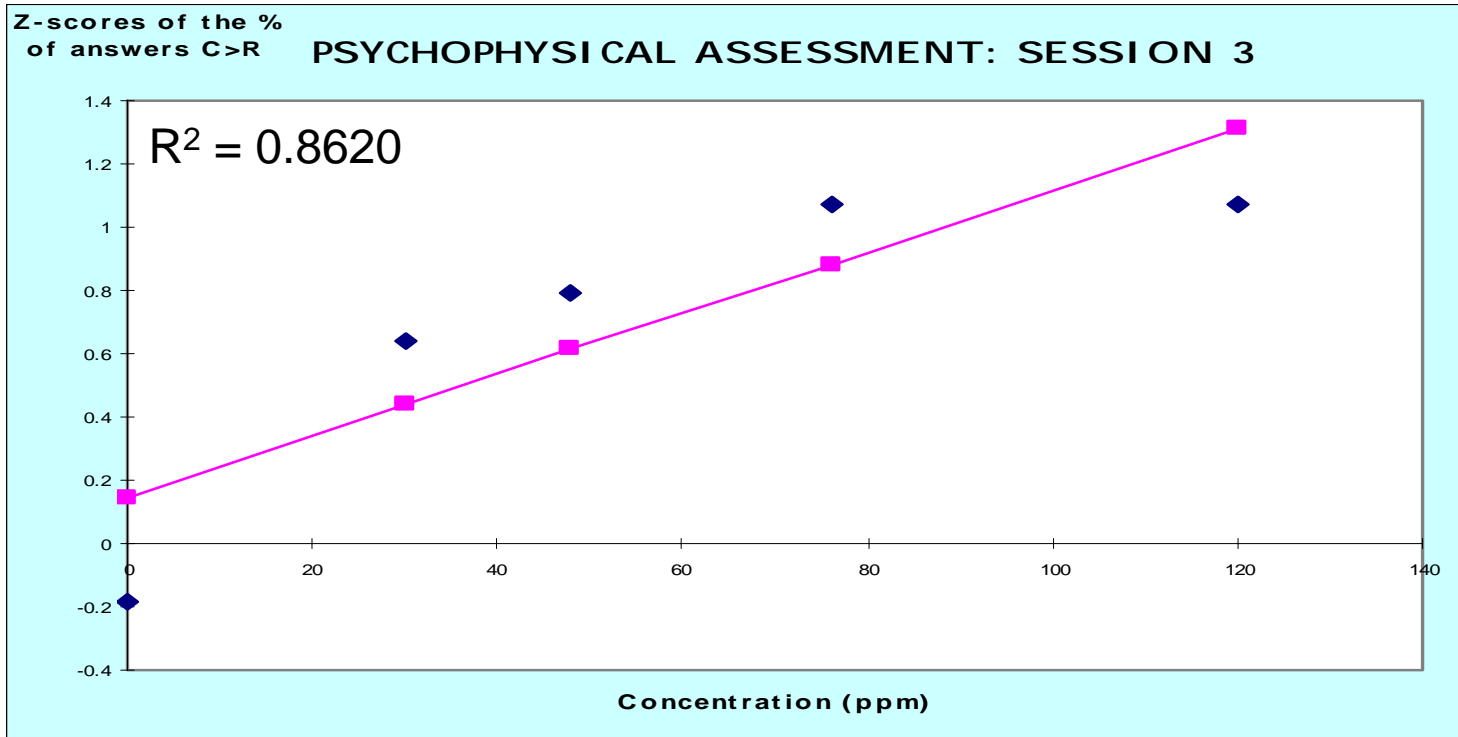


Psychophysical results



■ Method of Constant Stimuli

- Multiple paired comparison – spiked Vs. reference (modulus)
- $JND = (0.675 - 0.1469) / 0.0097$
- $JND = 54 \text{ ppm}$



Experimental approach...



■ The Triangle Test

- Subjects are asked to select the most different stimulus from among three, two of which are different

■ Paired Preference Test

- Subjects are presented with two samples and asked “which of the these samples do you prefer?”

■ Monadic Test (Same/Different test)

- Subjects are presented with samples one by one and asked “if this your normal sample?”
- This procedure can produce Signal Detection Measures
- Numerous presentations of both “targets” and “distractors” enables a measure of hits, misses, correct rejections and false alarms

Experimental design...



- **The Triangle Test – (Control Group, N=42, Test Group N=42)**
 - Subjects repeated the triangle test (x2) so the design had to balance
 - 6 possible triangle test combinations – 18 possible pairs
 - AAB, ABA, BAA, BBA, BAB, ABB
 - All pairs sampled equally across groups
 - **Paired Preference Test - (Control Group, N=42, Test Group N=42)**
 - Subjects completed three preference tests (AB, BA, AB)
 - All pairs of samples were presented an equal number of times across groups
 - **Monadic Test - (Same/Different test) - (Control Group, N=84, Test Group N=84)**
 - Samples were presented six times (3 normal and 3 spiked) in a random manner
- Throughout the test – all consumers received 12 samples (6 normal and 6 spiked)
 - The test presentation was balanced – 50% received the Triangle or PP first
– 50% received the Monadic First
 - Data was analysed using a General Linear Model

Our Experimental approach...

Triangle test

Triangle test 1 Triangle test 2

“Circle the odd sample?” “Circle the odd sample?”

Paired Preference

“Which sample do you prefer?”

Paired Preference 1 Paired Preference 2 Paired Preference 3

Monadic

“Is this your Vegemite?”

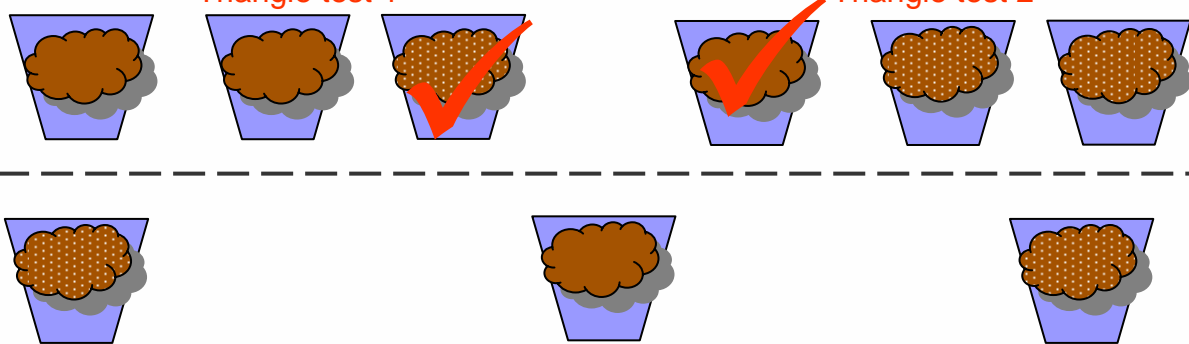
Same Same Same Same Same Same
Different Different Different Different Different Different

Our Experimental approach...

Triangle Test (Control group – No story)

Triangle test 1

Triangle test 2



Same Different Same Different Same Different

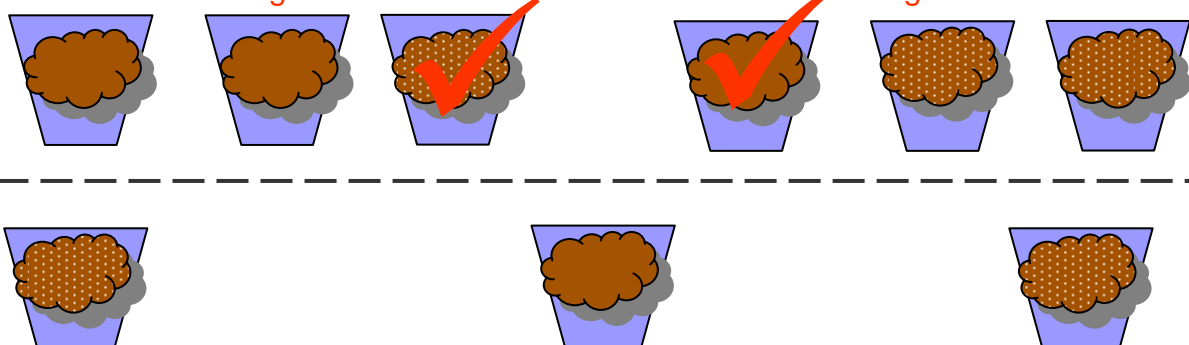
Detailed description: This diagram illustrates a triangle test for a control group without a story. It is divided into two rows by a dashed line. The top row shows three trials. In the first trial, three blue cups are shown: the first two contain a solid brown substance, and the third contains a dotted brown substance, with a red checkmark below it. In the second trial, the first two cups contain solid brown substance and the third contains dotted brown substance, with a red checkmark below the first cup. In the third trial, the first cup contains solid brown substance and the second and third contain dotted brown substance. The bottom row shows three trials. In the first trial, the first cup contains dotted brown substance and the second and third contain solid brown substance. In the second trial, all three cups contain solid brown substance. In the third trial, the first and second cups contain dotted brown substance and the third contains solid brown substance. Below each trial in the bottom row are labels 'Same' and 'Different' with checkboxes. The results are: Trial 1 (Same Different) and Trial 2 (Same Different).

Triangle Test (Authenticity Group – Story)

Cover Story

Triangle test 1

Triangle test 2



Same Different Same Different Same Different

Detailed description: This diagram illustrates a triangle test for an authenticity group with a story. It is divided into two rows by a dashed line. The top row shows three trials. In the first trial, three blue cups are shown: the first two contain a solid brown substance, and the third contains a dotted brown substance, with a red checkmark below it. In the second trial, the first two cups contain solid brown substance and the third contains dotted brown substance, with a red checkmark below the first cup. In the third trial, the first cup contains solid brown substance and the second and third contain dotted brown substance. The bottom row shows three trials. In the first trial, the first cup contains dotted brown substance and the second and third contain solid brown substance. In the second trial, all three cups contain solid brown substance. In the third trial, the first and second cups contain dotted brown substance and the third contains solid brown substance. Below each trial in the bottom row are labels 'Same' and 'Different' with checkboxes. The results are: Trial 1 (Same Different) and Trial 2 (Same Different).

Our Cover Story...

“In a recent survey - Vegemite was featured as the top Australian icon of all time and was described as *"A real true-blue Aussie food"*. It is thought that only real Australian's can distinguish the true taste of Vegemite.

We are not convinced that Australians can distinguish the true taste of Vegemite – in fact we are willing to demonstrate that YOU can't tell the difference!

To do this - we have sourced a product from overseas – which we think is very similar to Vegemite.

Let's see if YOU can tell the difference?”



FSA – Authenticity Test

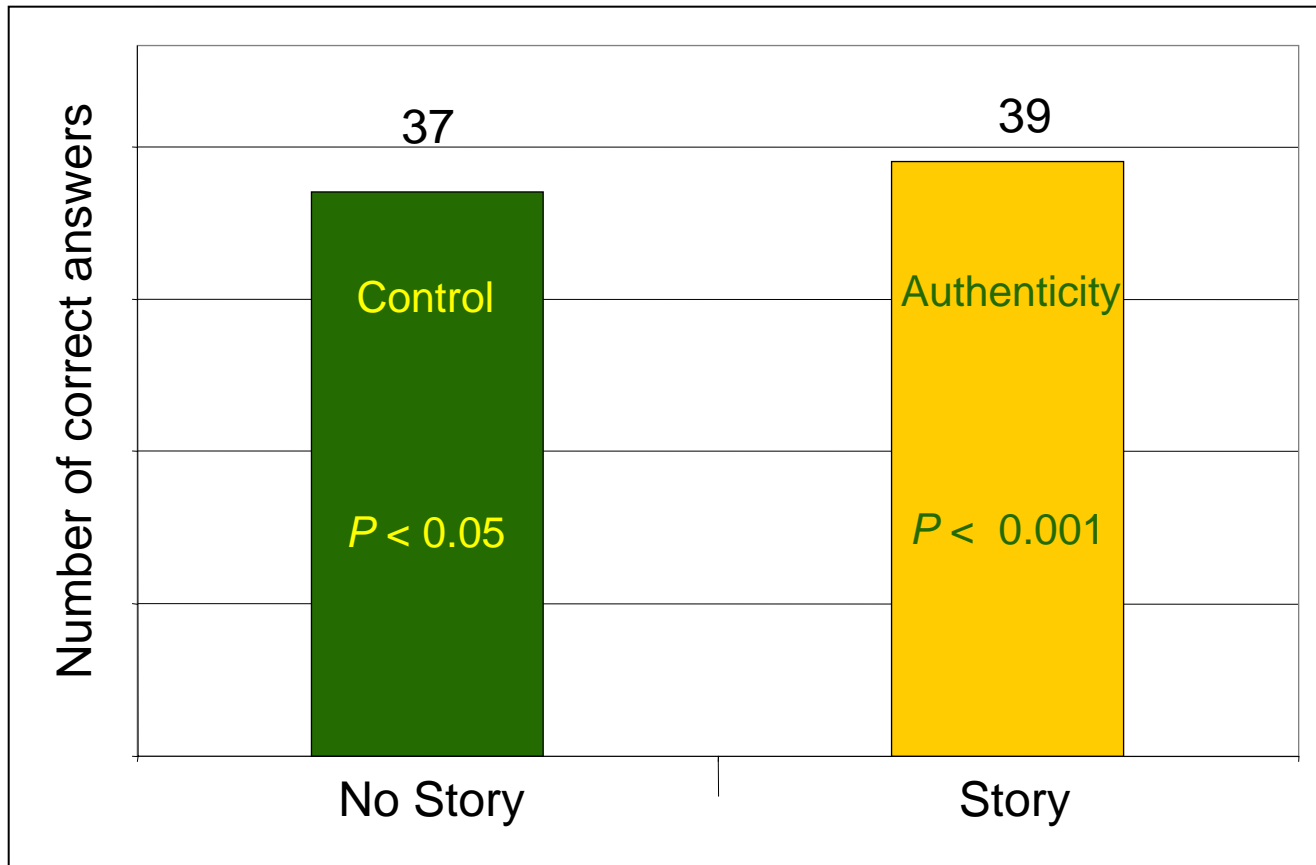


Results for today **Ideas for tomorrow**

Results: The Triangle Test

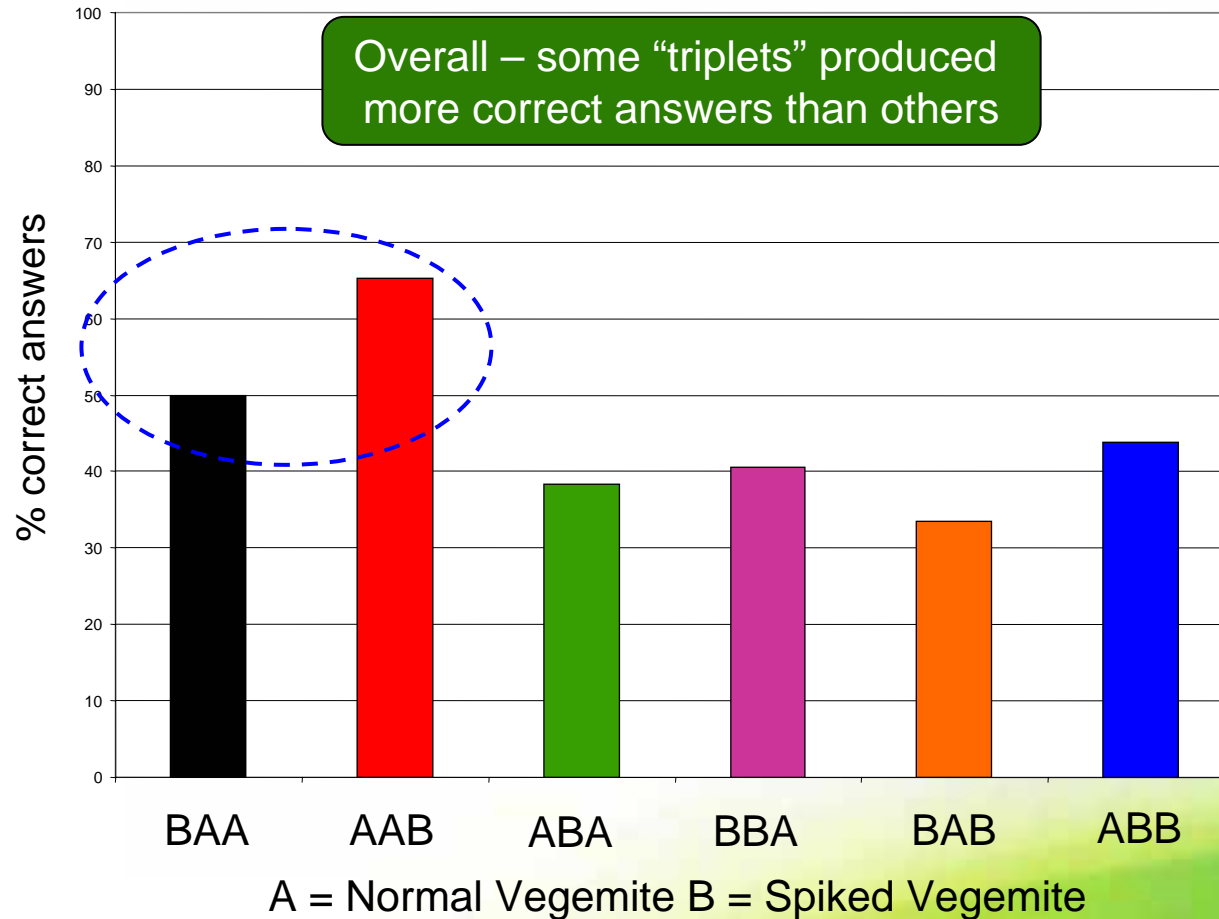
Triangle Test Results I

- When the results were cumulated across all assessors the spiked sample was found to be (just) significantly different to the control (un-spiked Vegemite)
- This proves the JND levels used were appropriate and that the difference was not too big
- *Was discrimination uniform across the triangle tests?*



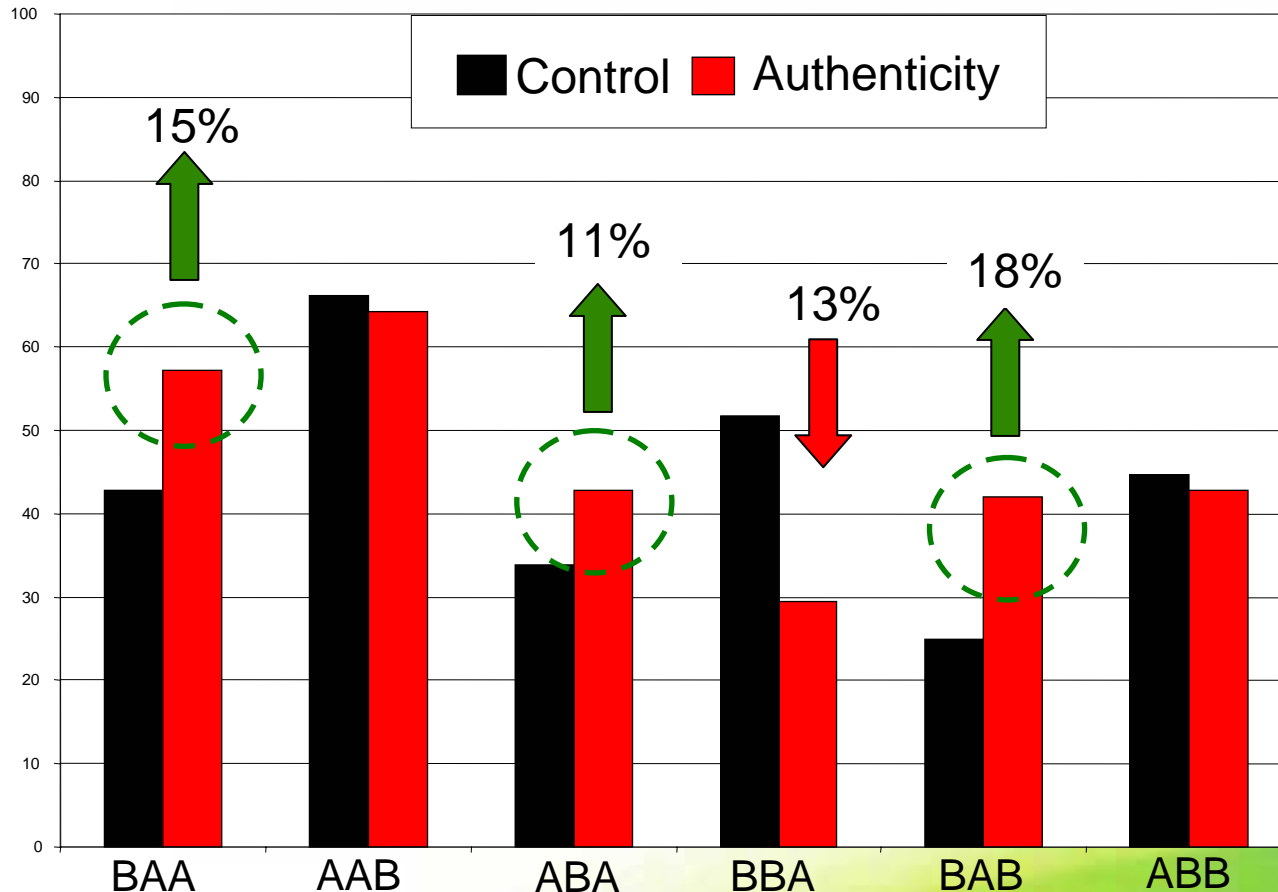
Triangle Test Results II

- *Differential response across the triangle tests presented*
- *Some pairings were easier to discriminate than others (BAA/ABB)*
- *What happened in the control and Authenticity groups?*



Triangle Test Results III

- Authenticity group were significantly **better** than Control in 3/6 tests
- Cover story reduced sensitivity for one test (BAB)
- *Why is this happening? What are the implications for difference tests in industry?*



Triangle Test Conclusions

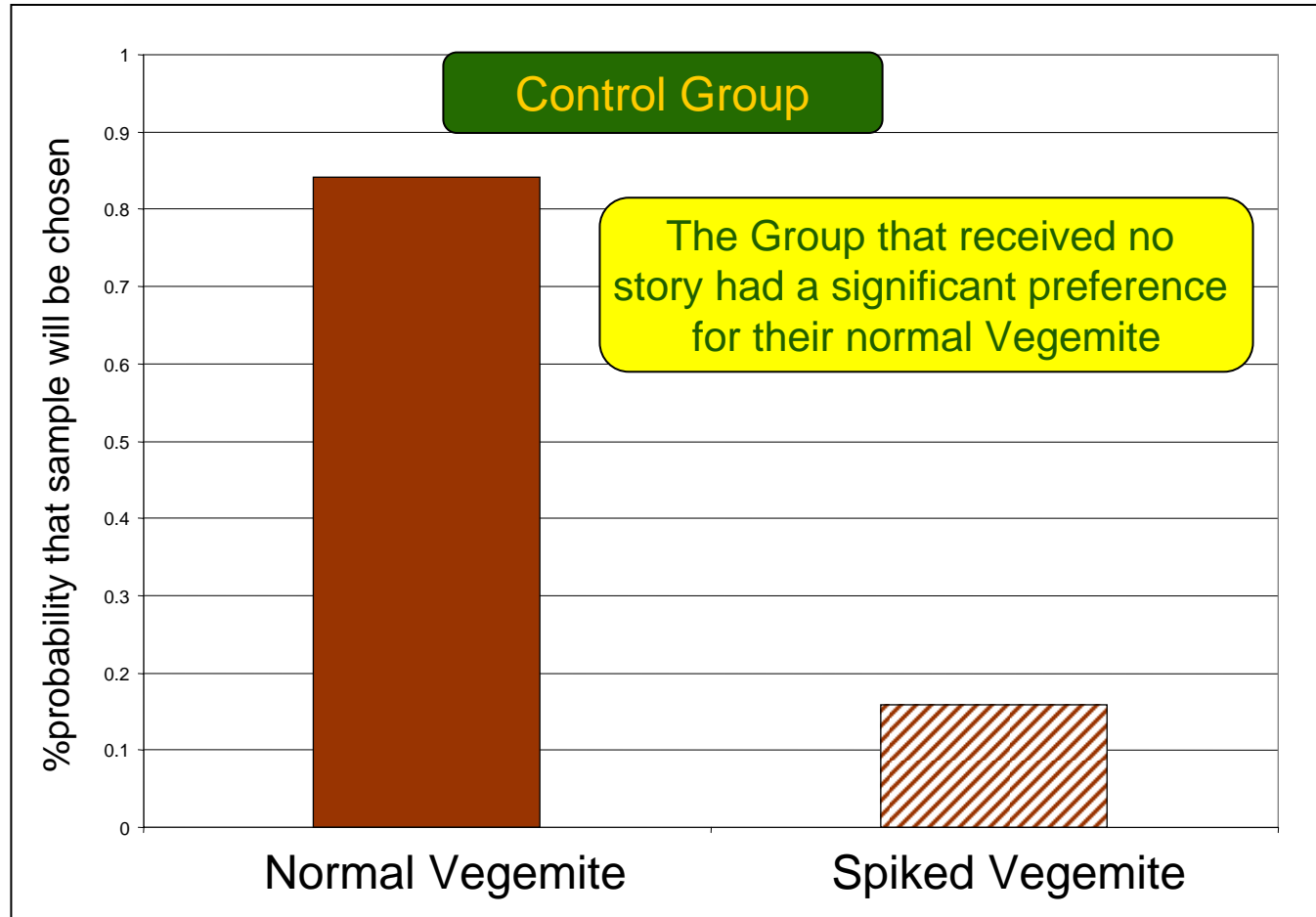
*The Australian Standard for Triangle tests (AS 2542.2.2) recommends we should cumulate responses across individual tests and compare totals to a table
However our study reveals a significant effect of Triplet (ABA vs. BBA) & would suggest we not total across triangle tests*

- **Using a cover story:**
 - *Improved the discrimination in half the tests*
 - *Had no effect on two of the triangle tests*
 - *Made one of the tests more difficult*
- *Why is there a differential response across the triangle tests?*
- *Why does a psychological bias influence the “BBA” test?*
- *What are the implications?*
- Further work will focus on understanding why there was a differential influence of the cover story across the different test triplets

Results: The Paired Preference Test

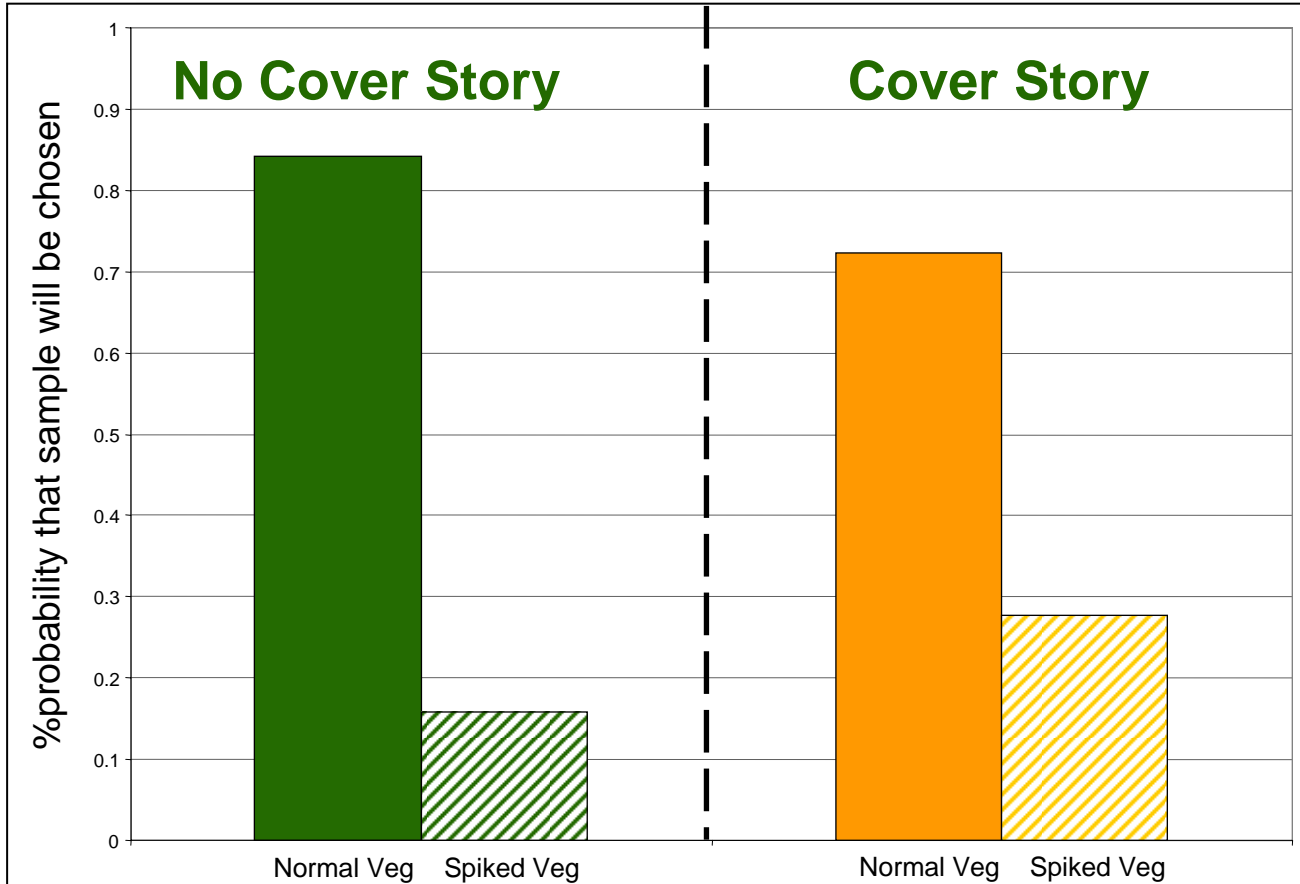
Paired Preference Test Results I

- We look at “discrimination ability” rather than preference *per se* (i.e.) using ability to discriminate the samples based on preference response as a measure of test discrimination power



Paired Preference Test Results II

- Consumers that received the cover story also showed a significant liking for their normal Vegemite
- However in comparison to the group that received no cover story, they were slightly less discriminating



Paired Preference test Conclusions

- *Overall – discrimination between normal and spiked Vegemite was strong....(Implications?)*
- *Consumers had no doubt which sample they preferred*
- *In the test group that received a cover story – there was a slight decrease in discrimination ability*
- *Why did the cover story (psychological bias) slightly reduce discrimination on this test?*

Liking and Cognition



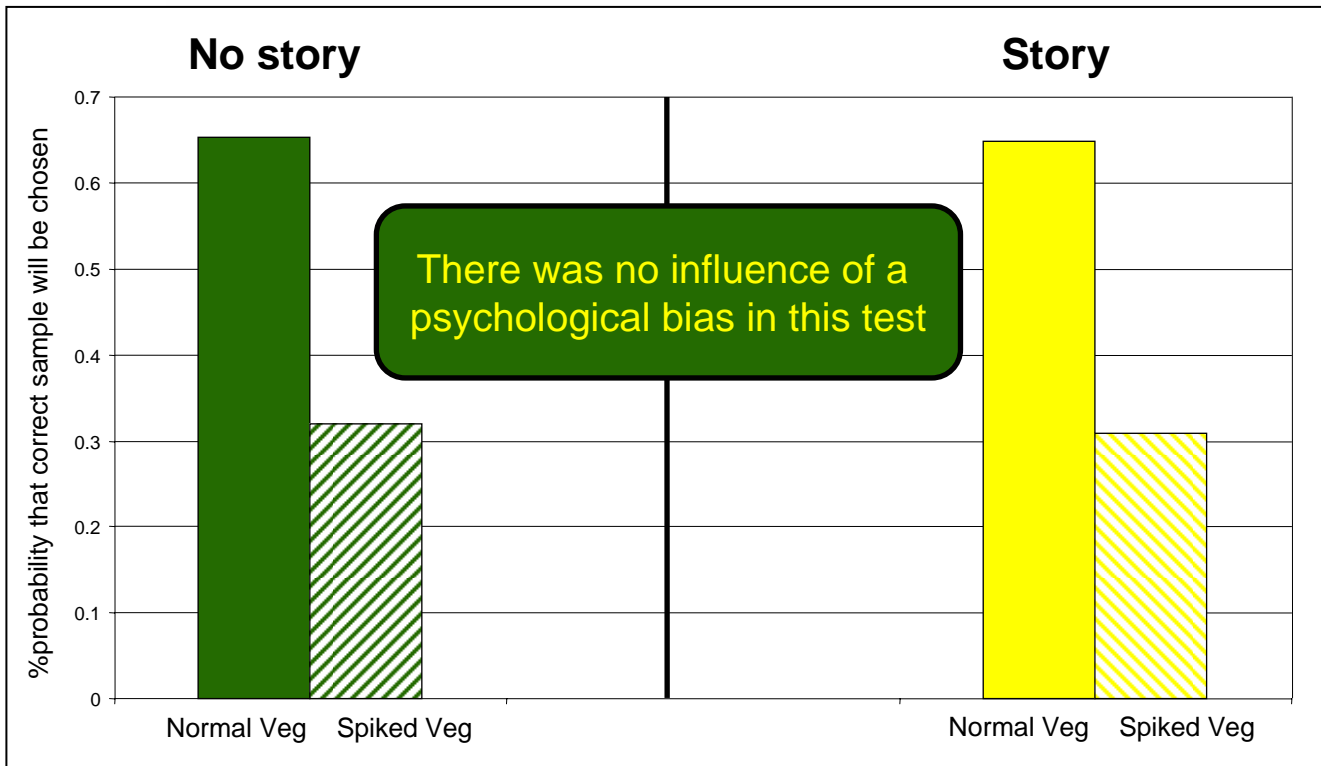
It may be that the application of a psychological bias interrupts a process that is inherently sensitive

- *“Affective reactions (liking) occur without cognitive encoding”* (Zajonc 1980) – so asking a consumer to consciously “think” about liking may bias the response
- Cover story aims to get consumers into an “emotional” rather than an “analytical” frame of mind
- **Emotional** = base decisions on an implicit reference (little/no conscious awareness)
- **Analytical** = leads to conscious attempts to retrieve info (destroys implicit memory)
(Trained panel results are based on conscious awareness)

Results: The Monadic Test

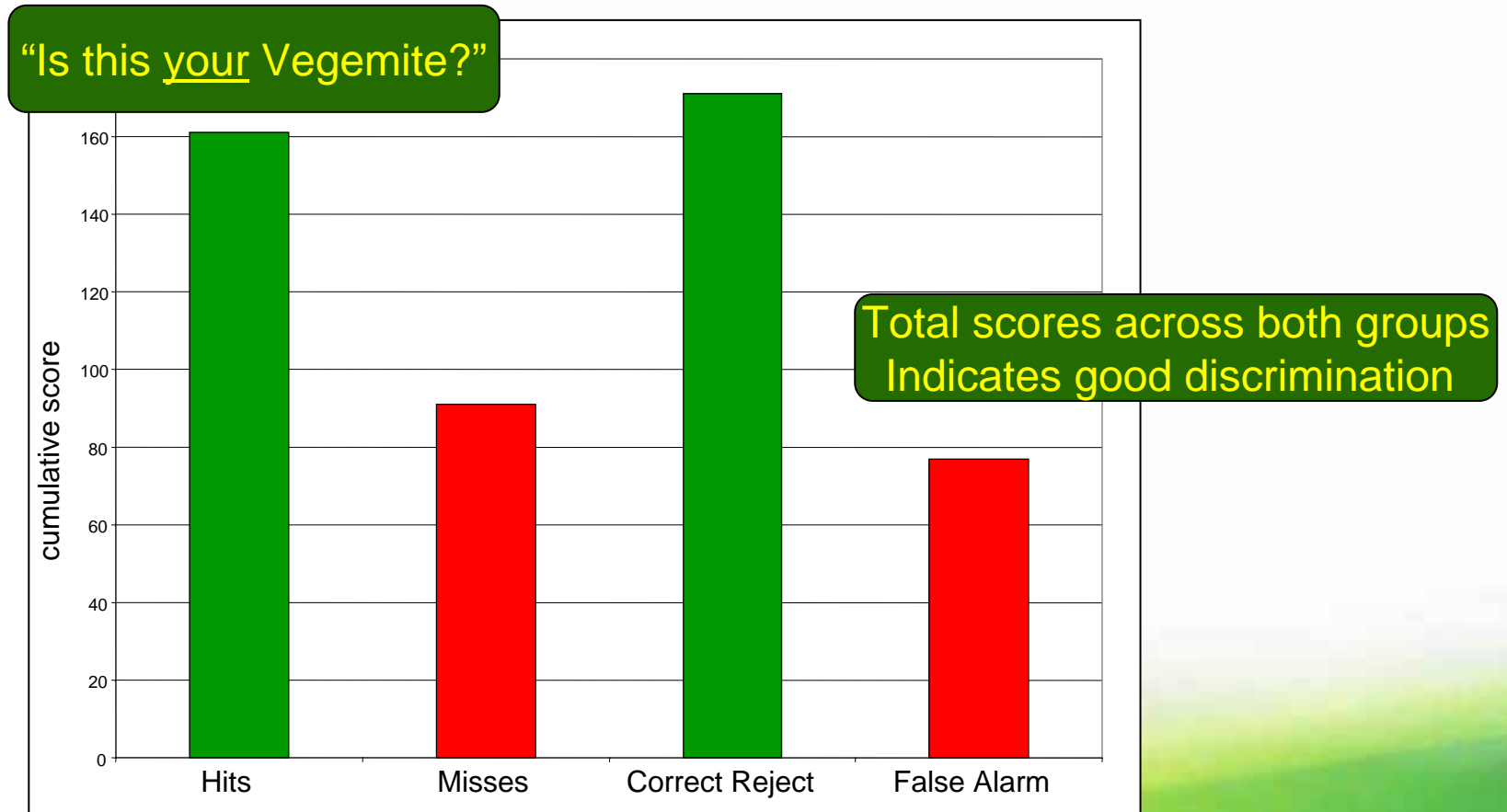
Monadic Test Results I

- Consumers picked their normal Vegemite correctly approx. 66% of the time (2:1)
- There was no difference in the discrimination between the consumers that received the cover story and those that did not

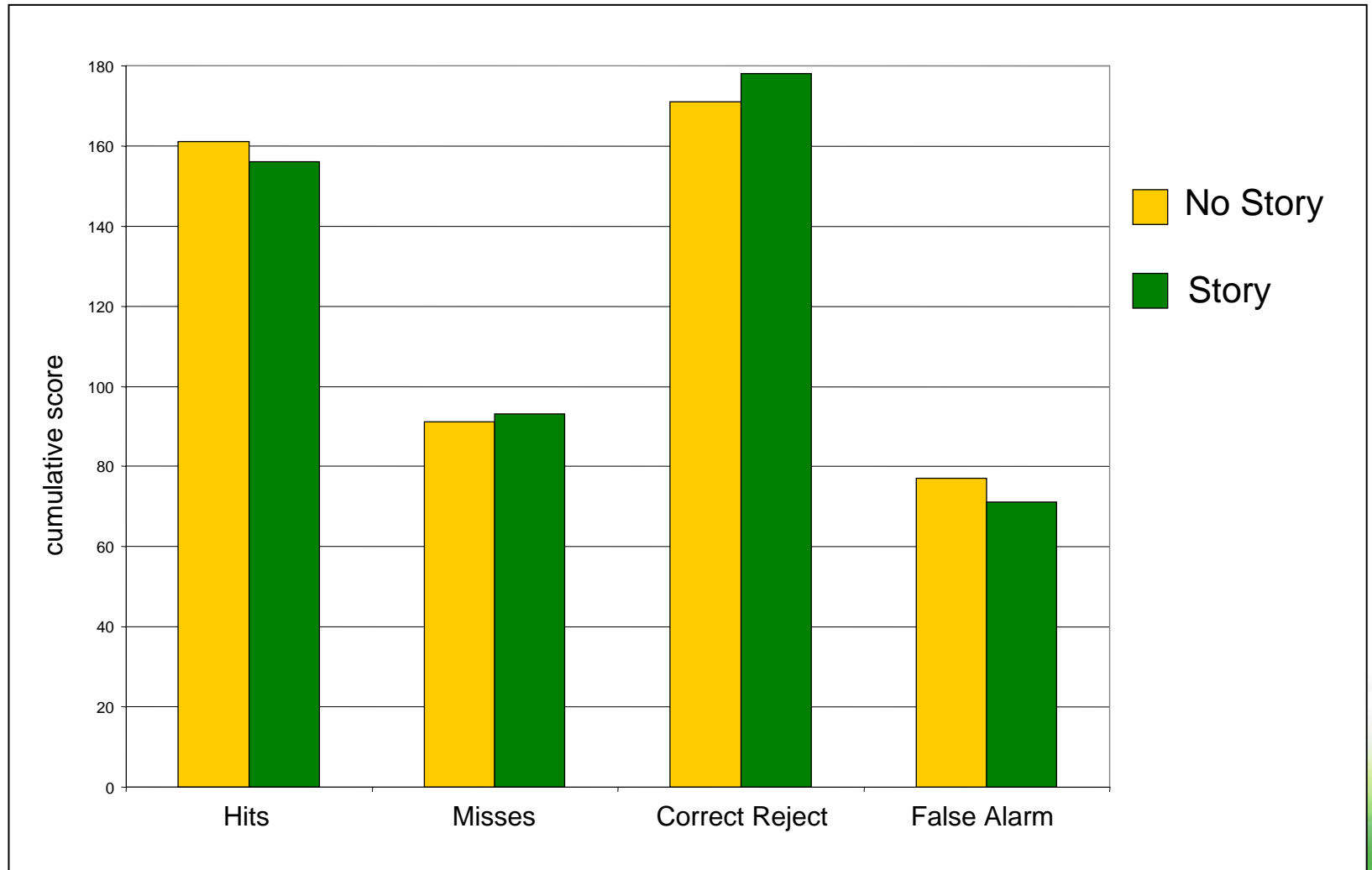


Monadic Test Results II

- The Monadic test is based on a Signal Detection Approach
- Each response produces 4 answers –
 - (i) Correct Acceptance
 - (ii) Incorrect Acceptance
 - (iii) Correct Rejection
 - (iv) Incorrect Rejection



Monadic Test Results III *(Signal Detection Results – Contingency Tables)*

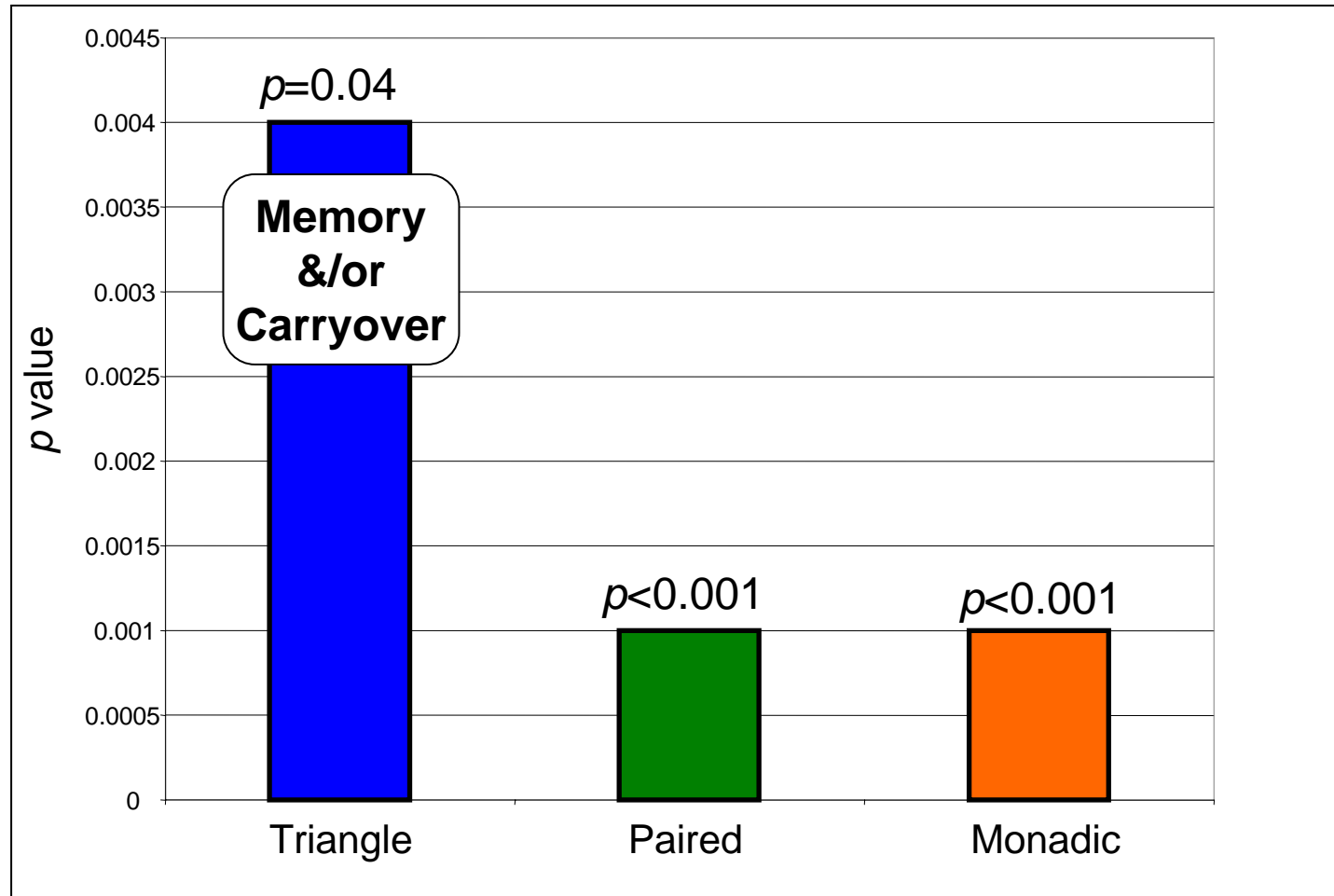


Monadic Test Conclusions

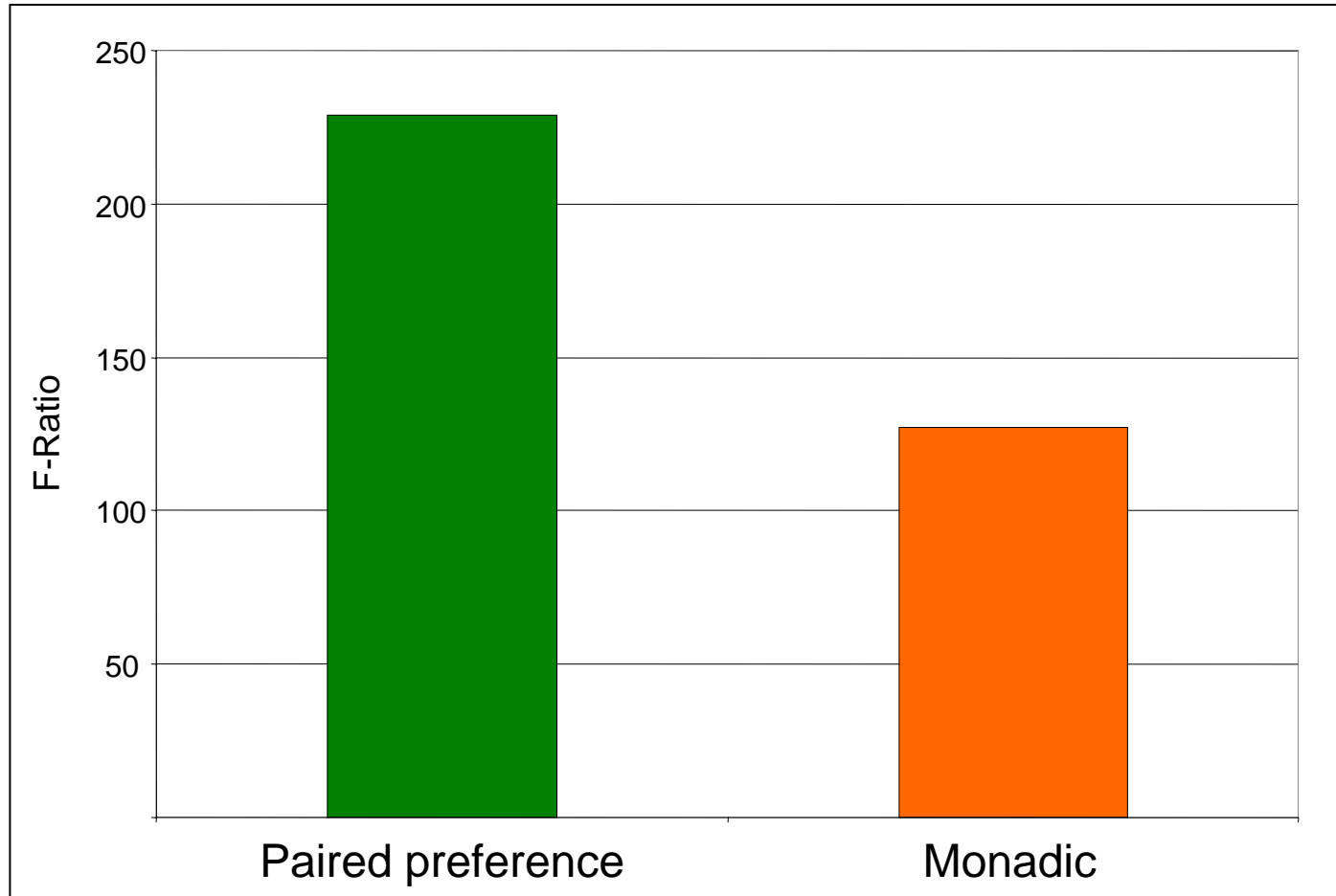
- *Discrimination between normal and spiked Vegemite was strong*
- *Consumers that received the cover story and those that did not had no doubt which sample was their Vegemite*
- *The cover story (psychological bias) had no influence on the discrimination between the two tests*

Results: Comparison across the 3 methods

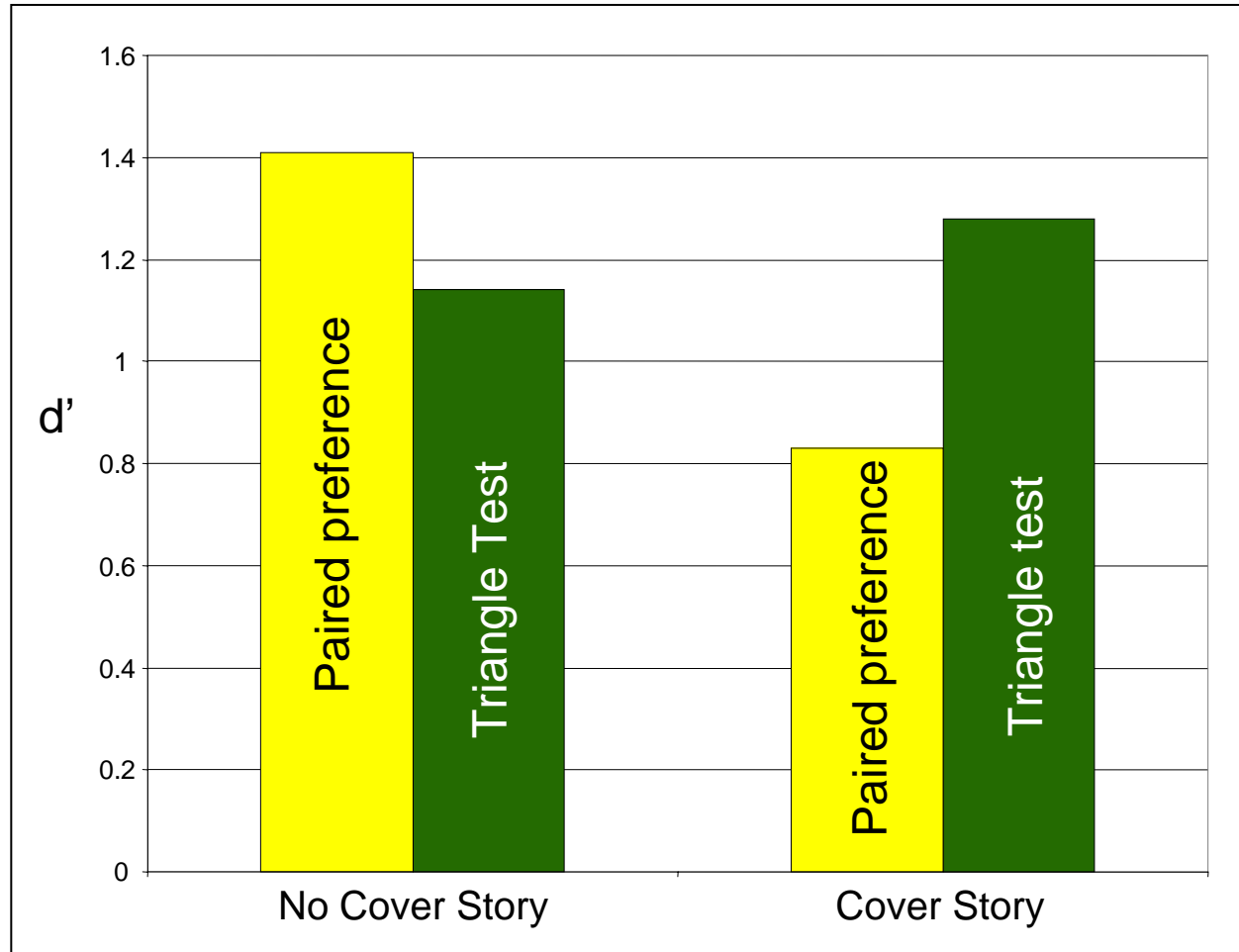
Discrimination Across Tests



Comparison Across Tests



Comparison of d' for PP & Triangle Test



Conclusions

- The **Paired Preference** technique was best at discriminating the subtle sensory differences
 - Implications (i.e. shelf life testing sensory analysis?)
- **Triangle Test** – had a differential response to the cover story
 - Some tests were easier than others
 - Cover story may have application in analytical sensory tests
- The **Monadic test** was not influenced by the cover story
 - Both groups were good at discriminating the sensory differences
 - Perhaps the story did engage consumers?



Possible fallacies...

- Stimulus: were the differences suitable? **Most likely yes...**
- Groups – was it valid to compare groups? **Most likely yes...**
- Perhaps the cover story did not engage the consumers enough

Further Questions...

- Why was there a differential response to the cover story?
- Why did the cover story reduce the accuracy of some responses?
- What application can the Authenticity method have in the future?

Next Steps...

- Learning's from our study will be fed into the next experiment which will be conducted in Dijon by INRA

Acknowledgments

Thanks to the FSA Sensory Team

Patrick O'Riordan

Stephanie Kirchhoff

Emlyn Williams

Niki Georgoulis

Celine Spinosi

Sophie Hunot

Ewa Orszulok

Georgie Giannakopoulos



ESN TG2

Jos Mojet (A+F)

Ep Koster (A+F)

Astrid Poelman (A+F)

Sylvie Issanchou (INRA)

Claire Sulmont (INRA)

David Lyon (Firmenich)

Henriette de Koch (Uni Pretoria)