SENSOMETRICS 2012
NEW SKIN FOR THE OLD CEREMONY
10th to 13th July 2012,
Agrocampus Ouest, Rennes, FRANCE

Operational power outweighs statistical power: Optimization of test performance of the triangle and duo-trio method

Hye-Seong Lee* and Min-A Kim
Ewha Womans University, South Korea

Danielle van Hout
Unilever R&D Vlaardingen, The Netherlands

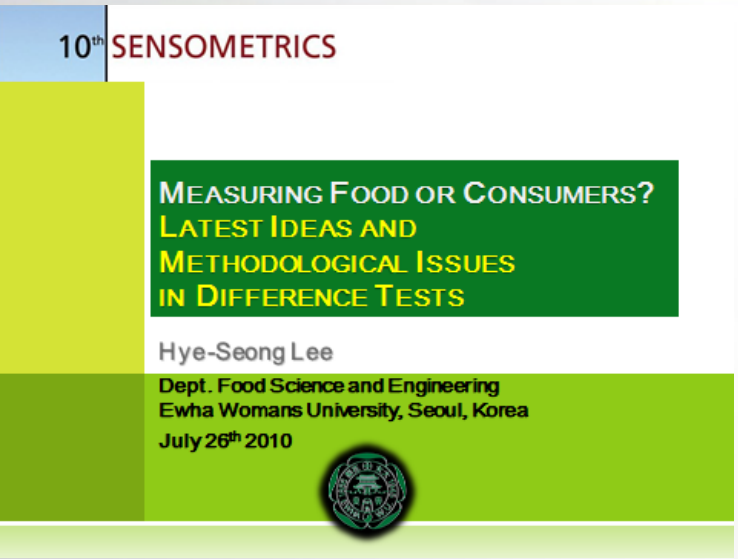
Ewha,
Where Change Begins

이화여자대학교
EWHA WOMANS UNIVERSITY

Unilever

Contents

Contents



- Sensory discrimination method that is **optimal** for measuring consumers' sensory discriminability?
- **Operational test power**

2. Experiment

- Triangle vs. 3 fixed-reference Duo-trio (DT)
- DT variants in terms of how reference was assigned

3. Results & Discussion

- Test performance in terms of d'
- Sequence effects using Cochran-Mantel-Haenszel test (fixing test order effects)
- Discussions for optimal test procedure

1. Introduction

Sensory Difference Test

Sensory Specification

- Trained panelists
- Sensory identification /discrimination
- Degree of difference between Foods

Consumer Discriminability

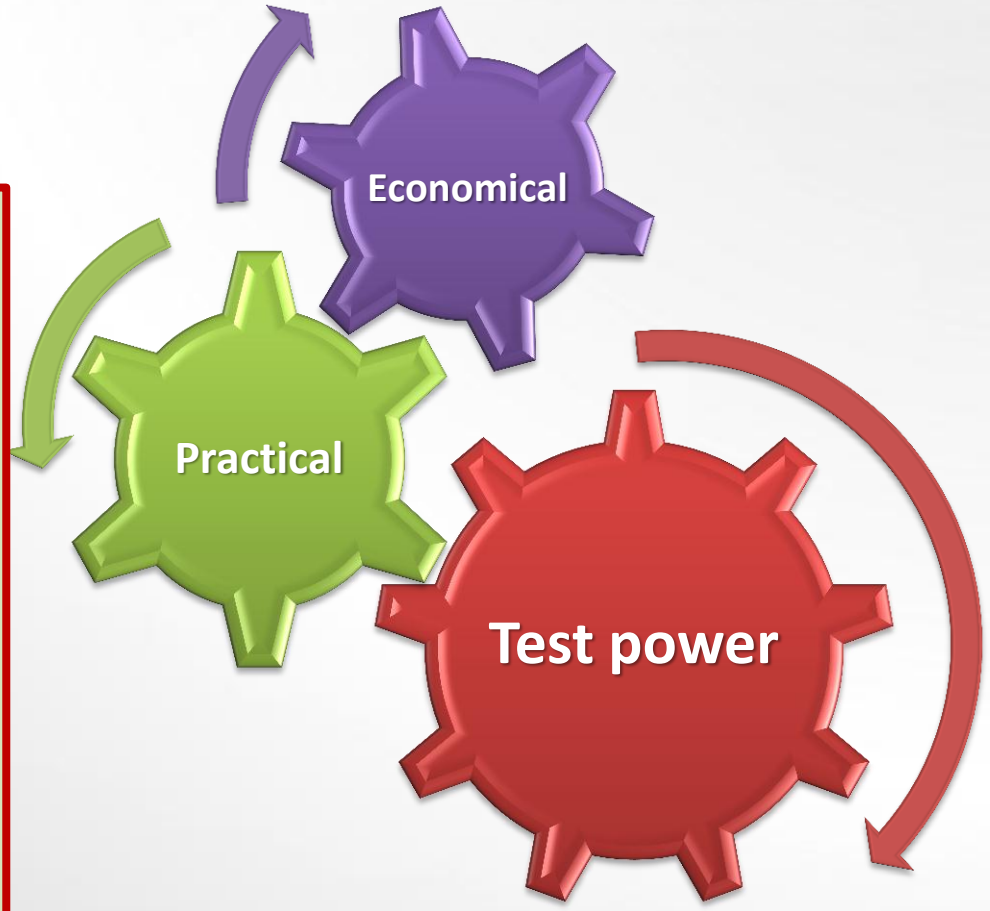
- Natural consumer perception and discriminability
- Strategic approach
- Quantitative index system is needed with pre-determined decision criteria

Consumer Discrimination Test

How to optimize the test method?

- The statistical power of the test method has been generally emphasized.

- Yet, it is also important that the test procedure should …
  1. **generate data reflecting the real perceptual discriminability**
  2. **be practically manageable in a consistent way**
  3. **not produce significant non-relevant sources of perceptual variables**



Power of Sensory Difference Test

Statistical power

- Probability to detect a true sensory difference

Operational power

- Operational probability to detect a true sensory difference (Bi, Lee, & O'Mahony, 2010; van Hout, Hautus, & Lee, 2011)

- Test method to be used needs to be determined by several factors, theoretical **statistical power** as well as various physiological and cognitive effects which influence the operational power.
- **Sequence effects** caused by **physiological and cognitive perceptual biases such as adaptation and memory effects** have been reported as important factors affecting this operational power (Kim & Lee, 2012; Lee, Chae, & Lee, 2009).

Sequence Effects & Test Power

- When sequence effects are significant, the statistical power for sensory discrimination tests cannot be validly predicted.

➢ **In the present study, formulating the test protocol that can be less affected by the sequence effects was considered as a way to improve operational power.**

Attribute Non-specified Discrimination Test

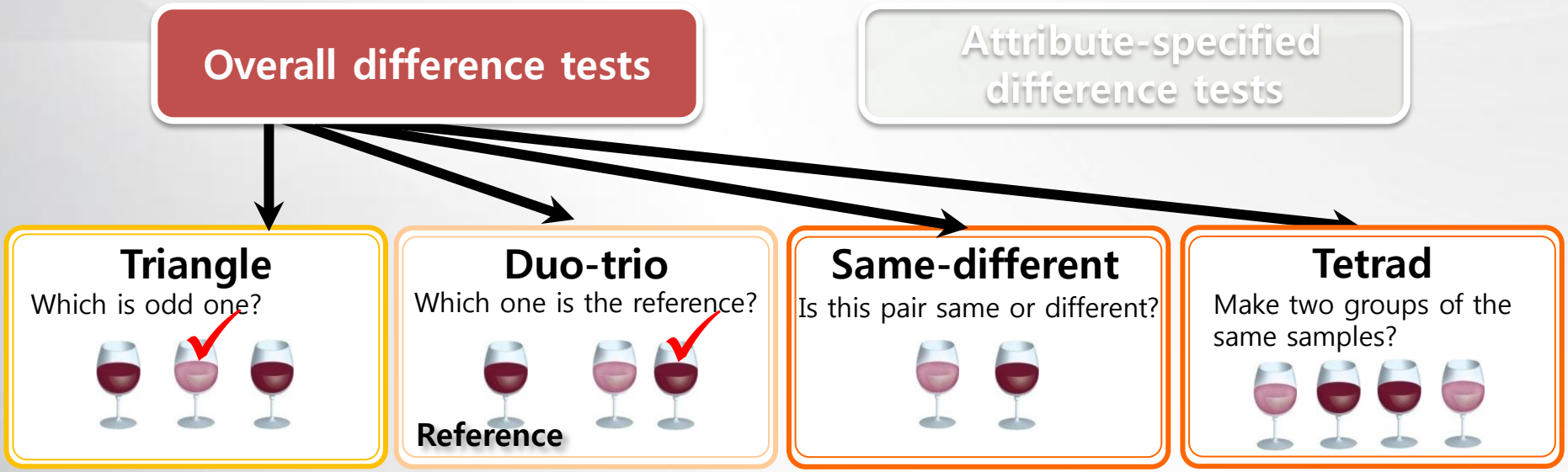
- Sensory difference test methods have been classified into:

Overall difference tests

Attribute-specified difference tests

Attribute Non-specified Discrimination Test

- Sensory difference test methods have been classified into:



- The test method for consumer discrimination test needs to involve consumers' natural attention and perception.
- The overall difference test where consumers do not need to selectively attend to particular attribute is generally recommended.

Design of the Overall Different Test

- Test design can be classified into:

Variable-reference design

Two samples are used in balanced design

Fixed-reference design

Constantly only one sample is used as reference or odd one

Design of the Overall Different Test

- Test design can be classified into:

Variable-reference design

Two samples are used in balanced design

Fixed-reference design

Constantly only one sample is used as reference or odd one

Triangle

Duo-trio

Design of the Overall Different Test

- Test design can be classified into:

Variable-reference design

Two samples are used in balanced design

Fixed-reference design

Constantly only one sample is used as reference or odd one

Triangle

- Both samples are presented as odd one.
- The position of the odd stimulus is balanced.
- Data from this triangle method are actually best summarized in a 6x3 matrix.

Design of the Overall Different Test

- Test design can be classified into:

Variable-reference design

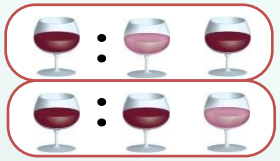
Two samples are used in balanced design

Fixed-reference design

Constantly only one sample is used as reference or odd one

Triangle

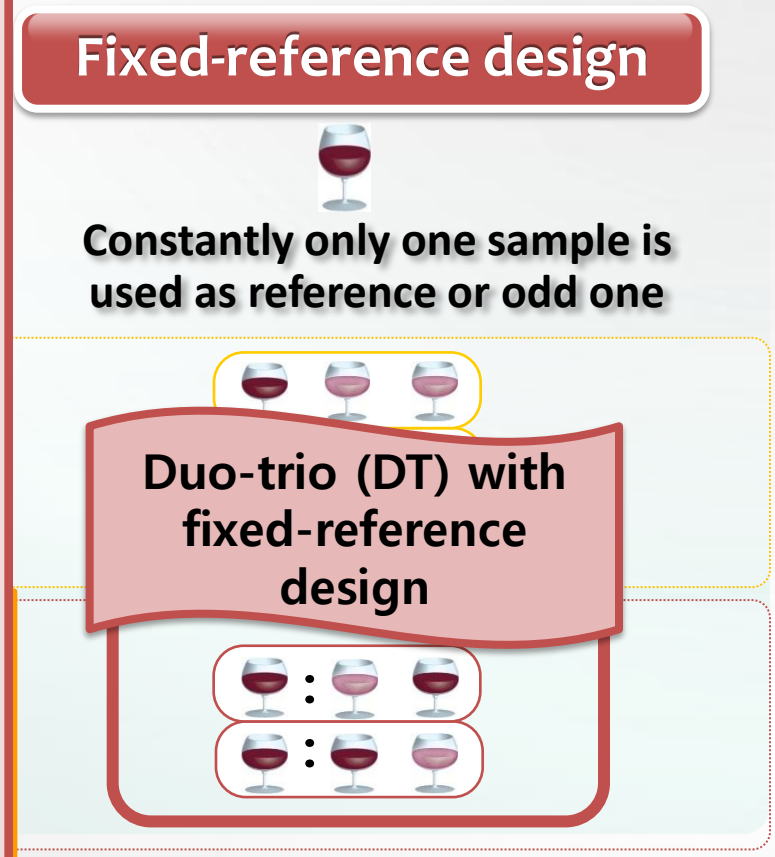
Duo-trio



Design of the Overall Different Test

DT with fixed (constant) reference design might be more suitable for studying consumers' discriminability.

1. Regarding sequence effects of difference tests using three stimuli, position of each stimulus was more important than the number of the stronger one (Lee, Chae, & Lee, 2009).
2. Compared to when the reference was balanced, DT was more sensitive when the stronger-reference OR preferred reference was used (Chae, Lee, & Lee, 2010; Kim & Lee, 2010; Kim & Lee, 2012).
3. Practically, consumers' discriminability is important for business objectives such as reformulation and cost reduction in the situations where an original control sample is available.



Objectives of the Present Experiment

The objective of this paper is to investigate operationally more powerful way of using the CONSUMER DISCRIMINATION method.

- To investigate the test performance of the **three different types of duo-trio method with fixed (constant) reference design in comparison to the (variable-reference) triangle method: discriminability & sequence effects**
- The three types of duo-trio consumer discrimination tests were designed to investigate **the effects of the brief familiarization (pre-viewing process) incorporating affective components** in the task.
- The samples varying its salt contents were tested attempting to simulate the situations of investigating consumers' discriminability between the original and sodium-reduced product.

2. Experiment

Experimental Variables

- **Investigated discrimination method:**
  - Duo-trio with a fixed (constant) reference design
  - Saltier sample as the constant reference

- **4 consumer discrimination test methods to be compared**

Condition 1	**Traditional triangle method**
Condition 2	**Duo-trio in a normal analytical way**
Condition 3	**Duo-trio with a brand image provided**
Condition 4	**Duo-trio providing the preferred one as reference** (pre-test of preference)

Hypotheses

Condition 1	**Traditional triangle method**
Condition 2	**Duo-trio in a normal analytical way**
Condition 3	**Duo-trio with a brand image provided**
Condition 4	**Duo-trio providing the preferred one as reference**

1) When discriminating between samples varying salt contents, the saltier reference is better remembered by consumers. Duo-trio test with saltier reference would be based on more stable perceptual dimension for discrimination, and thus performing better than the (variable-reference) triangle test: **Conditions 2-4 > Condition 1**

2) Positioning the preferred sample or a sample driving better attention and affection as the reference might improve consumers discrimination: **Conditions 3-4 > Condition 2**

Experimental Design

- Comparisons using an independent samples design with a triangle method as a control

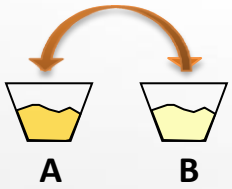
258 consumers

Pre-test: triangle method for sensitivity checking

	Group 1 (N=66)	Group 2 (N=64)	Group 3 (N=64)	Group 4 (N=64)
Main-test:	Condition 1: Traditional triangle method	Condition 2: Duo-trio method in a normal analytical way	Condition 3: Duo-trio method with a brand image provided	Condition 4: Duo-trio method providing the preferred one as reference

- **Stimuli**
  - Corn-soup varying salt contents

2 stimuli in comparisons:



Experimental Procedures

Main-test

- For all conditions, the total number of tastings required was the same.

Condition 1 **Traditional triangle method**

Q: Which sample is odd one?

Duo-trio method **Condition 2**

- Called by "the reference"
- Coded with "R"

Q: Which sample is the reference?

Condition 3 **Duo-trio method with a brand image provided**

- Called by "the best of a particular brand"
- Coded with the brand's logo

Q: Which sample is the best product which was tasted before?

Duo-trio method providing preferred one as reference **Condition 4**

- Called by "your choice"
- Coded with "C"

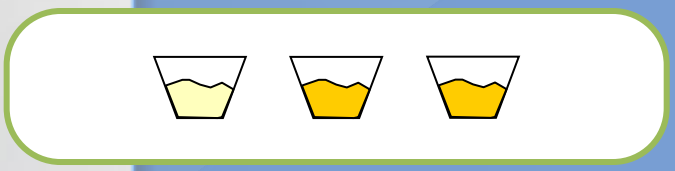
Q: Which sample is the preferred product which was chosen by you?

Experimental Procedures

Main-test

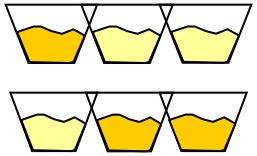
- For all conditions, the total number of tastings required was the same.

Condition 1 **Traditional triangle method**



Q: Which sample is odd one?

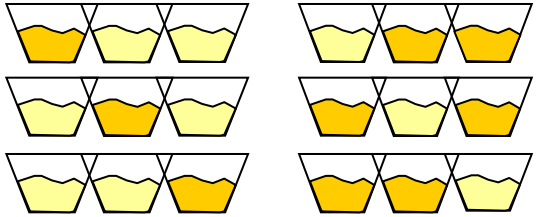
1st sub-session



Triangle method (2 out of 6 possible sequence)

2 min break

2nd sub-session



Triangle method (6 possible sequence)

Total : 24 no. of tastings for a session

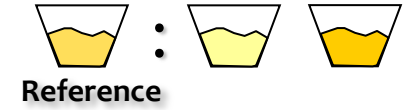
Experimental Procedures

Main-test

- For all conditions, the total number of tastings required was the same.

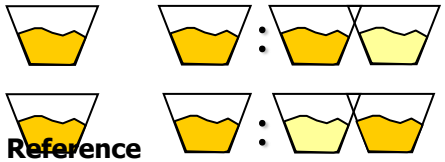
Duo-trio method — Condition 2

- Called by "the reference"
- Coded with "R"



Q: Which sample is the reference?

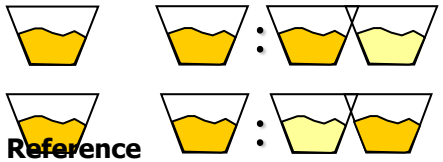
1st sub-session



Duo-trio method
(2 possible sequence)

2 min break

2nd sub-session



Duo-trio method
(2 possible sequence)

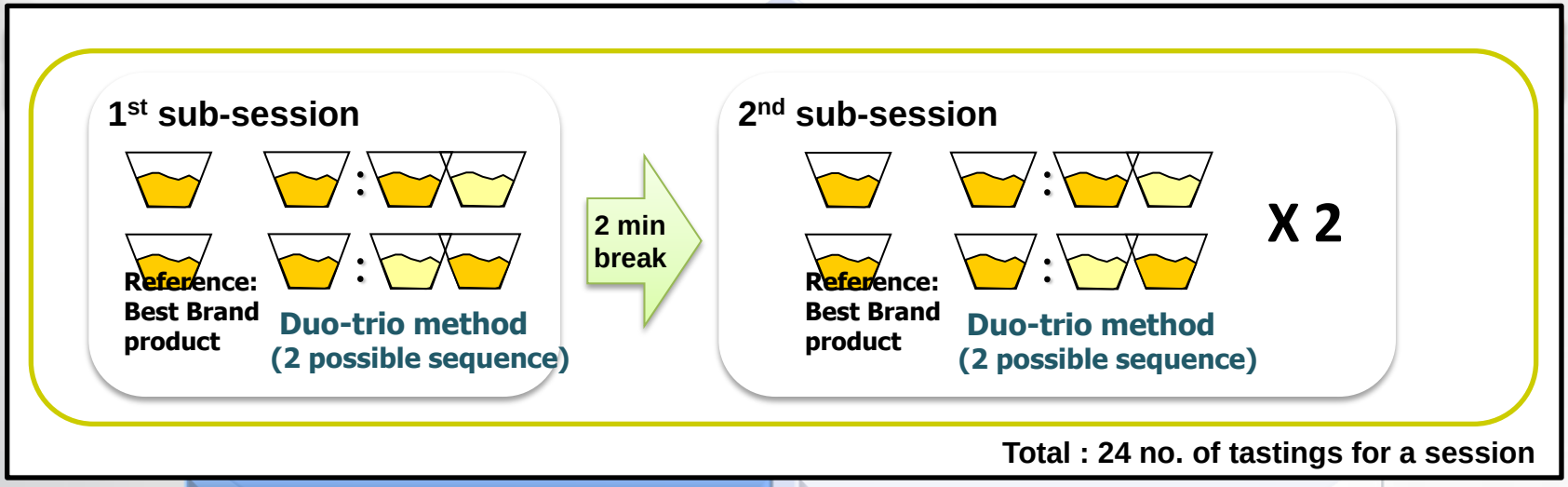
X 2

Total : 24 no. of tastings for a session

Experimental Procedures

Main-test

- For all conditions, the total number of tastings required was the same.



Condition 3 **Duo-trio method with a brand image provided**

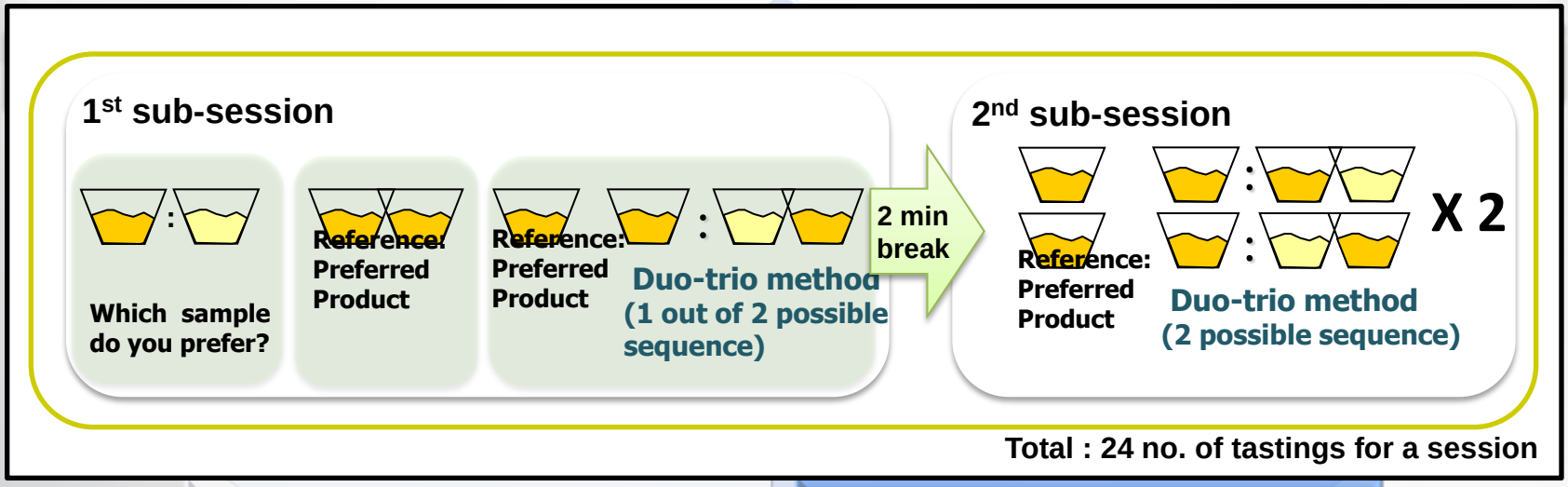
- Called by "the best of a particular brand"
- Coded with the brand's logo

Q: Which sample is the best product which was tasted before?

Experimental Procedures

Main-test

- For all conditions, the total number of tastings required was the same.



Duo-trio method providing preferred one as reference

Condition 4

- Called by "your choice"
- Coded with "C"

Q: Which sample is the preferred product which was chosen by you?

Data Analysis

Discriminability

- Test performances were compared in terms of ***d'* estimates** as well as $\mathbf{P_d}$ (probability of discriminators).
- For triangle method and three different forms of duo-trio method, the group mean *d'*s were obtained based on **'comparison of distances(COD) strategy'** using R-package sensR (Christensen & Brockhoff, 2011) freely available for the free statistical software package R (R Development Core Team, 2011)
- The group mean *d'*s were estimated based on **pooled data** using **a standard beta-binomial model** (Ennis & Bi, 1998).
- The significance tests among multiple *d'*s were determined based on the values of the *d'*s and variances of *d'*s (Marascuilo, 1970).

Data Analysis

Sequence (or position) effects

- To investigate whether the stimuli sequence presented in a test had a significant effect on the response variable affecting the performances, **Cochran-Mantel-Haenszel tests** were performed after fixing test order effect using the XLSTAT add-in for Microsoft Excel (ver. 2010 for Windows, XLSTAT, Addinsoft, Paris, France) (Mental, 1963).

3. Results & Discussion

Results

Comparison of test performances

Pre-test: Triangle method (Total 6 sets)

Consumers (n=258) were equally divided into 4 groups according to their sensitivity.

	Group 1 (N=66) Triangle	Group 2 (N=64) Triangle	Group 3 (N=64) Triangle	Group 4 (N=64) Triangle	Total (N=258)
Replication	6	6	6	6	**6**
Pc	0.54	0.54	0.54	0.54	**0.54**
Pd *(SE)*	0.31 (0.04)	0.31 (0.04)	0.30 (0.04)	0.31 (0.04)	**0.31 (0.02)**
d' *(SE)*	1.66 (0.13)^{a}	1.67 (0.13)^{a}	1.65 (0.13)^{a}	1.68 (0.13)^{a}	**1.67 (0.06)**

Main-test: Comparison of test performance among different test methods

	Condition 1 Triangle	Condition 2 Duo-trio	Condition 3 Duo-trio w/ brand	Condition 4 Duo-trio w/ preferred R	
Replication	8	6	6	5	
Pc	0.53	0.74	0.81	0.82	
Pd *(SE)*	0.29 (0.04)	0.49 (0.06)	0.62 (0.04)	0.63 (0.05)	
d' *(SE)*	1.61 (0.12)^{b}	2.00 (0.19)^{ab}	**2.43 (0.16)^{a}**	**2.48 (0.20)^{a}**	

Results (cont'd)

Comparison of test performances

Pre-test: Triangle method (Total 6 sets)

Consumers (n=258) were equally divided into 4 groups according to their sensitivity.

	Group 1 (N=66) Triangle	Group 2 (N=64) Triangle	Group 3 (N=64) Triangle	Group 4 (N=64) Triangle	Total (N=258)
Replication	6	6	6	6	**6**
Pc	0.54	0.54	0.54	0.54	**0.54**
Pd (*SE*)	0.31 (0.04)	0.31 (0.04)	0.30 (0.04)	0.31 (0.04)	**0.31 (0.02)**
d' (*SE*)	1.66 (0.13)[a]	1.67 (0.13)[a]	**1.65 (0.13)[b]**	**1.68 (0.13)[b]**	**1.67 (0.06)**

Main-test: Comparison of test performance among different test methods

	Condition 1 Triangle	Condition 2 Duo-trio	Condition 3 Duo-trio w/ brand	Condition 4 Duo-trio w/ preferred R
Replication	8	6	6	5
Pc	0.53	0.74	0.81	0.82
Pd (*SE*)	0.29 (0.04)	0.49 (0.06)	0.62 (0.04)	0.63 (0.05)
d' (*SE*)	1.61 (0.12)[a]	2.00 (0.19)[a]	**2.43 (0.16)[a]**	**2.48 (0.20)[a]**

Results (cont'd)

Comparison of performances between the first two & later tests

Main-test: Comparison of test performance among different test methods

		Condition 1 Triangle	Condition 2 Duo-trio	Condition 3 Duo-trio w/ brand	Condition 4 Duo-trio w/ preferred R
1st sub-session	Replication	2	2	2	1
	Pc	0.47	0.76	0.81	0.83
	Pd (*SE*)	0.20 (0.07)	0.52 (0.09)	0.63 (0.07)	0.66 (0.09)
	d' (*SE*)	1.30 (0.24)^{b}	**2.07 (0.27)^{a}**	**2.45 (0.26)^{a}**	**2.57 (0.37)^{a}**
2nd sub-session	Replication	**6**	**4**	**4**	**4**
	Pc	0.55	0.74	0.81	0.81
	Pd (*SE*)	0.32 (0.04)	0.48 (0.06)	0.62 (0.05)	0.63 (0.06)
	d' (*SE*)	1.71 (0.13)^{b}	1.95 (0.20)^{ab}	**2.42 (0.20)^{a}**	**2.45 (0.21)^{a}**

For triangle, in the first 2 tests the discriminability was much lower, but in the later repetitions, it has been improved.

Results (cont'd)

Examination of sequence effects

Considering individual difference as random variable, the effect of the variations in stimuli sequence was examined based on Cochran-Mantel-Haenszel test (fixing test order effects)

Experimental session	Group Protocol		Test order								Total
			1st	2nd	3rd	4th	5th	6th	7th	8th	
Pre-test	**1 ~ 4 Triangle**	$\chi 2$	41.14	23.98	23.99	22.80	14.48	12.75			
	(N=258)	***p***	< 0.01	0.01	0.01	0.01	0.15	0.24			**< 0.01**
Main-test	**1 Triangle**	$\chi 2$	7.73	13.26	12.70	18.89	5.62	4.90	14.06	5.20	
	(N=66)	***p***	0.66	0.21	0.24	0.04	0.85	0.90	0.17	0.88	**< 0.01**
	2	χ^2	0.02	0.03	0.07	0.03	0.07	0.04			
	(N=64)	***p***	1.00	1.00	1.00	1.00	1.00	1.00			**1.00**
	3	χ^2	1.12	4.93	0.34	0.44	1.46	0.73			
	(N=64)	***p***	0.33	0.04	0.73	0.69	0.26	0.55			**0.44**
	4 (pref. A)	χ^2	0.93	1.25	0.93	1.29	3.62				
	(N=38)	***p***	0.45	0.38	0.44	0.43	0.11				**0.26**
	4 (pref. B)	χ^2	0.01	0.05	0.00	0.03	0.87				
	(N=26)	***p***	1.00	1.00	1.00	1.00	0.65				**0.82**

Results (cont'd)

Examination of sequence effects

Considering individual difference as random variable, the effect of the variations in stimuli sequence was examined based on Cochran-Mantel-Haenszel test (fixing test order effects)

Experimental session	Group Protocol		Test order								Total
			1st	2nd	3rd	4th	5th	6th	7th	8th	
Pre-test	1 ~ 4 Triangle	$\chi 2$	41.14	23.98	23.99	22.80	14.48	12.75			
	(N=258)	*p*	< 0.01	0.01	0.01	0.01	0.15	0.24			**< 0.01**
Main-test	1 Triangle	$\chi 2$	7.73	13.26	12.70	18.89	5.62	4.90	14.06	5.20	
	(N=66)	*p*	0.66	0.21	0.24	0.04	0.85	0.90	0.17	0.88	**< 0.01**
	2	χ^2	0.02	0.03	0.07	0.03	0.07	0.04			
	(N=64)	*p*									**1.00**
	3	χ^2									
	(N=64)	*p*									**0.44**
	4 (pref. A)	χ^2	0.93	1.25	0.93	1.29	3.62				
	(N=38)	*p*	0.45	0.38	0.44	0.43	0.11				**0.26**
	4 (pref. B)	χ^2	0.01	0.05	0.00	0.03	0.87				
	(N=26)	*p*	1.00	1.00	1.00	1.00	0.65				**0.82**

Significant sequence effect was observed only in the triangle method (p-value < 0.01)

Results (cont'd)

Comparison of test performances using identical sequences

Pre-test

	Group 2 (N=64) Triangle	Group 3 (N=64) Triangle	Group 4: pref. A (N=38) Triangle	Group 4: pref. B (N=26) Triangle
Stimuli presentation	<A-A-B>, <A-B-A>	<A-A-B>, <A-B-A>	<A-A-B>, <A-B-A>	<B-A-B>, <B-B-A>
d' (SE)	2.14 (0.22)	1.83 (0.22)	2.21 (0.29)	2.05 (0.40)

Main-test

	Condition 2 Duo-trio	Condition 3 Duo-trio w/ brand	Condition 4 Duo-trio w/ preferred R	Condition 4 Duo-trio w/ preferred R
Stimuli presentation	<A: A B>, <A: B A>	<A: A B>, <A: B A>	<A: A B>, <A: B A>	<B: A B>, <B: B A>
d' (SE)	2.00 (0.19)	2.43 (0.16)	2.48 (0.27)	2.47 (0.29)

Results (cont'd)

Comparison of test performances using identical sequences

Pre-test

	Group 2 (N=64) Triangle	Group 3 (N=64) Triangle	Group 4: pref. A (N=38) Triangle	Group 4: pref. B (N=26) Triangle
Stimuli presentation	<A-A-B>, <A-B-A>	<A-A-B>, <A-B-A>	<A-A-B>, <A-B-A>	<B-A-B>, <B-B-A>
d' (SE)	2.14 (0.22)	1.83 (0.22)	2.21 (0.29)	2.05 (0.40)

Main-test

	Condition 2 Duo-trio	Condition 3 Duo-trio w/ brand	Condition 4 Duo-trio w/ preferred R	Condition 4 Duo-trio w/ preferred R
Stimuli presentation	<A: A B>, <A: B A>	<A: A B>, <A: B A>	<A: A B>, <A: B A>	<B: A B>, <B: B A>
d' (SE)	2.00 (0.19)	2.43 (0.16)	2.48 (0.27)	2.47 (0.29)

No significant difference between triangle and duo-trio

Results (cont'd)

Comparison of test performances using identical sequences

Pre-test

	Group 2 (N=64) Triangle	Group 3 (N=64) Triangle	Group 4: pref. A (N=38) Triangle	Group 4: pref. B (N=26) Triangle	All Group (N=258) Triangle
Stimuli presentation	<A-A-B>, <A-B-A>	<A-A-B>, <A-B-A>	<A-A-B>, <A-B-A>	<B-A-B>, <B-B-A>	**All sequences**
d' (SE)	2.14 (0.22)	1.83 (0.22)	2.21 (0.29)	2.05 (0.40)	**1.67 (0.06)**

Main-test

	Condition 2 Duo-trio	**Condition 3 Duo-trio w/ brand**	**Condition 4 Duo-trio w/ preferred R**	**Condition 4 Duo-trio w/ preferred R**
Stimuli presentation	<A: A B>, <A: B A>	<A: A B>, <A: B A>	<A: A B>, <A: B A>	<B: A B>, <B: B A>
d' (SE)	2.00 (0.19)	2.43 (0.16)	2.48 (0.27)	2.47 (0.29)

The results suggest that in condition 3, the higher attention might be the reason for improving the performance, while in condition 4, the test sequence itself was favorable or optimized.

Discussions

- In the present study, the performance of **the fixed-reference duo-trio (DT) method** was investigated as a consumer discrimination method.
- This method has generally found to be superior to the (variable-reference) triangle method due to the favourable sequence and memory advantage.

Discriminability of DT

- The discriminability of this fixed-reference duo-trio method, improved with brand encoding of the reference or engaging consumers' preference within the pre-viewing phase.
- Such modifications seem to be important to induce more natural consumers' perception because in normal life situations, branded products are often consumed and consumers develop preference to their familiar products.

Discussions (cont'd)

- Yet, in the present study, not real brand loyal consumers were tested.
  - It can be hypothesized that for loyal consumers who are more familiar to, or have stronger affects towards one original sample (reference), familiarization engaging consumers' involvements and affective state of mind could also induce more efficient form of the decision strategy used for the test.
  - This will also lead to improvement of test power.
  - Therefore, as a follow-up study, the effects of these test methods should be further investigated using different groups of consumers having different degree of familiarity to the products.

Discussions (cont'd)

Sequence (or position) effects

- Significant sequence effects were found only in triangle method. **Variability introduced by test sequences can be confounded with test order effects and individual differences when each consumer performs 2-3 tests in a session.**
- Results indicated that on average, **the duo-trio with fixed-reference had the operationally favorable stimulus sequences.** These sequences also tended to show higher discriminability in triangle method as well.
- These results suggest that **in order to optimize the test power, the duo-trio with saltier and/or preferred reference should be recommended utilizing the fixed-reference design and a scheme to stabilize the memory of the reference,** rather than randomizing all the possible test sequences of the method.

Acknowledgements

Unilever R&D
Sensation, perception & Behavior

Our Lab. members with Dr. Anna Thomas, Danielle van Hout, Dr. Kevin Blot, & Dr. JM Dessirier (Unilever R&D)

Thank you !

감사합니다.

Ewha,
Where Change Begins