Measuring Food or Consumers?
Latest Ideas and Methodological Issues in Difference Tests

Hye-Seong Lee
Dept. Food Science and Engineering
Ewha Womans University, Seoul, Korea
July 26th 2010
**SENSORY DIFFERENCE TESTS**

Various paradigms of difference tests can be used for assessment of

1. Sensory differences between confusable food samples in the analytical sensory evaluation of food

2. Consumers discriminability between samples

<table>
<thead>
<tr>
<th>Test protocol</th>
<th>Sample presentation &amp; Instruction</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-Not A</td>
<td>Is this ‘A’ or not?</td>
</tr>
<tr>
<td>2-AFC</td>
<td>Which one is ‘A’?</td>
</tr>
<tr>
<td>3-AFC</td>
<td>Which one is stronger?</td>
</tr>
<tr>
<td>Duo-trio</td>
<td>Which one is the reference?</td>
</tr>
<tr>
<td>Triangle</td>
<td>Which is odd one?</td>
</tr>
<tr>
<td>Dual-pair</td>
<td>Which pair is the same pair?</td>
</tr>
<tr>
<td>Same-different</td>
<td>Is this pair same or different?</td>
</tr>
</tbody>
</table>
SENSORY DIFFERENCE TESTS

(1) Sensory differences between confusible food samples in the analytical sensory evaluation of food specifications, using methods that are sensitive, having proper power, and reliable.

(2) Consumers discriminability between samples in the consumption context should be close to the normal situation. What else? How should the methodology differ then?

Test protocol  
A-Not A

Sample presentation & Instruction

Is this ‘A’ or not?

Which one is ‘A’?

Which one is stronger?

Which is the reference?

Which pair is the same pair?

Which pair is different?

Objective sensory specification, using equally sensitive panel.
UNDERSTANDING VARIABLES OF SENSORY DIFFERENCE TESTS

“What are subjects or consumers actually doing during performing the difference test procedures?”

Understanding this will help us...

(1) developing accurate modeling to various difference tests

(2) selecting and applying appropriate sensory difference tests according to the different purposes of experiments
PERCEPTION AND DECISION PROCESS IN DIFFERENCE TESTS
PERCEPTION AND JUDGMENT FOR DIFFERENCE TESTS

Decision process to make the judgment using the obtained information

Perceptual process to get the sensory information

Decision

Perception

Attention

Response

Tastings of Food Samples

Sampling in sequence with temporal intervals
### Factors influencing perception

<table>
<thead>
<tr>
<th>Cognitive perception strategy</th>
<th>How and what dimension subjects attend to: Affecting the nature of information</th>
</tr>
</thead>
</table>

**Function of Sensory System for Food**

- Affecting the clarity of information getting into the brain

**Tastings of Food Samples**
## Decision Process

<table>
<thead>
<tr>
<th>Factors affecting proportion of correct responses ($P_c$)</th>
<th>Cognitive decision strategy</th>
<th>Response bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjects' way to compare samples to choose a response</td>
<td></td>
<td>Subjects favoring one response over the other</td>
</tr>
</tbody>
</table>

### Tastings of Food Samples

**Decision**

**Perceptual representations of samples**

**Response**
FACTORS & MODELS EXPLAINING VARIABILITY IN DIFFERENCE TEST PERFORMANCES

1. Effects of involving hedonic state of mind on consumer discrimination

Factors influencing perception
- Cognitive perception strategy
- Function of Sensory System for Food

Factors affecting proportion of correct responses ($P_c$)
- Cognitive decision strategy
- Response bias

2. Effects of test designs & Thurstonian Modeling/Signal Detection Theory

3. Effects of order of samples presentation on a test & Sequential Perception Analysis
EFFECTS OF PERCEPTION PROCESS ON CONSUMER DISCRIMINATION
# Perception Process

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Test Influences</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Analytic (selective)</strong></td>
<td>• Previous training</td>
<td>• Prescott, Johnstone &amp; Francis, 2004</td>
</tr>
<tr>
<td>paying selective attention to specified attribute</td>
<td>• Nature of the instruction</td>
<td>• Le Berre, et al., 2008</td>
</tr>
<tr>
<td><strong>Synthetic (holistic)</strong></td>
<td>• Nature and degree of familiarization procedure</td>
<td>• Prescott &amp; Murphy, 2009</td>
</tr>
<tr>
<td>paying unitary, global attention to the overall food flavor</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Analytic (not subjective)</strong></td>
<td>• Test design</td>
<td>• Chae, Lee &amp; Lee, 2010</td>
</tr>
<tr>
<td>tests that do not consider the affective/hedonic states of the subjects</td>
<td>• Nature of the instruction</td>
<td>• Mojet &amp; Köster, 1986</td>
</tr>
<tr>
<td><strong>Affective (unitary, holistic)</strong></td>
<td>• Nature and degree of familiarization procedure</td>
<td>• Frandsen et al. 2003, 2007</td>
</tr>
<tr>
<td>Tests influenced by the mind set and subjective feelings of the individual</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Regarding food perceptions, depending on how you give your attentions to, there could be two distinctive perception processes:

- **Analytical (and selective) approach**, which a trained sensory panel would normally apply when evaluating food.

- **Affective (and synthetic) approach**, which naive consumers would normally apply when consuming food.

It was reported that affective/hedonic approach (different mode of attention, incompatible with analytical approach) promote synthetic perception (Prescott, Johnstone and Francis, 2004).
It has been reported that affective approach (i.e. authenticity test) is more sensitive than analytical approach to discriminate subtle differences in foods.

These results suggest that for the products that consumers have high emotional involvements, the affective concept (foreign vs. national) help to apply the synthetic perception process and define the perceptual variable.

<table>
<thead>
<tr>
<th>References</th>
<th>Food Sample</th>
<th>Analytical test</th>
<th>Sensitivity</th>
<th>Affective test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mojet &amp; Köster 1986</td>
<td>Beer</td>
<td>-</td>
<td></td>
<td>Authenticity test</td>
</tr>
<tr>
<td>Kjearulff. 2002</td>
<td>Milk</td>
<td>Multiple A-not A test</td>
<td>〈</td>
<td>Authenticity test</td>
</tr>
<tr>
<td>Frandsen et al. 2003</td>
<td>Milk</td>
<td>Descriptive analysis</td>
<td>〈</td>
<td>Authenticity test</td>
</tr>
<tr>
<td>Frandsen et al. 2007</td>
<td>Milk</td>
<td>Same-different test</td>
<td>〈</td>
<td>Authenticity test</td>
</tr>
</tbody>
</table>
Objective I: investigating the effects of the affective familiarization on the consumers’ discriminability in comparison with the same discrimination test in an analytical mode

EXPERIMENTAL PROCEDURE

- Stimuli: four commercial confusable milk products (A~D)
- Subjects: 100 female milk consumers (age range 22±2 years)
- Consumer performed 4 sessions of same-different tests and each session was preceded by one of the two different familiarization procedures

**Af.F:**
Affective familiarization procedure

- Four sets of 10 point rank-ratings for ‘liking’ and three affective and integrated attributes (‘freshness’, ‘well-being’, and ‘off-flavor’)

**An.F:**
Analytical familiarization procedure

- Four sets of 10 point rank-ratings for similarity to each of the four milk products

RESULTS: SAME-DIFFERENT RATINGS

- SD/Af.F resulted in higher discriminability than SD/An.F.

RESULTS: SAME-DIFFERENT RATINGS

- SD/Af.F resulted in higher discriminability than SD/An.F.

Drawing consumers’ attention in affective modes enhanced consumers' perceptual discriminability.

For measuring more natural consumers’ discriminability, it might be essential to trigger the affective mode of perception and allow the consumers’ natural, synthetic perception process to occur.
RESULTS: NOT HOMOGENEOUS RESPONSE DISTRIBUTIONS FOR SAME PRODUCTS

- χ²-tests for the same pairs revealed that consumers’ familiarity to the tested products were not all the same. Consumers developed better familiarity (or memory) to ‘A’ and ‘B’ than ‘C’ and ‘D’.

- These suggest that there might be an interaction between the affective perception and familiarity (or memory) to the products.

<table>
<thead>
<tr>
<th>Sample pair</th>
<th>SD/Af.F</th>
<th></th>
<th></th>
<th></th>
<th>Least liked,</th>
<th>Least fresh,</th>
<th>Most liked,</th>
<th>Most fresh</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Same</td>
<td>Different</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sure</td>
<td>Not sure</td>
<td>Don’t know</td>
<td>But guess</td>
<td>Don’t know</td>
<td>Not sure</td>
<td>Sure</td>
</tr>
<tr>
<td>AA</td>
<td>46</td>
<td>18</td>
<td>8</td>
<td>4</td>
<td>11</td>
<td>13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BB</td>
<td>39</td>
<td>36</td>
<td>6</td>
<td>4</td>
<td>9</td>
<td>6</td>
<td>31.25</td>
<td>0.01</td>
</tr>
<tr>
<td>CC</td>
<td>26</td>
<td>29</td>
<td>7</td>
<td>13</td>
<td>12</td>
<td>13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DD</td>
<td>28</td>
<td>27</td>
<td>11</td>
<td>4</td>
<td>17</td>
<td>13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>34.75</td>
<td>27.50</td>
<td>8.00</td>
<td>6.25</td>
<td>12.25</td>
<td>11.25</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
RESULTS: CONSUMERS SEGMENTATION

To check the effects of inter-consumers’ different familiarity and criteria in discriminability, a hierarchical cluster analysis was performed on frequency distributions elicited by the six response categories for the same product pairs (23 in Group1, 31 in Group2).

### Table 2. Results of ROC analyses on the pooled data across each class of subjects

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Sample pair</th>
<th>SD/Af.F</th>
<th>SD/An.F</th>
<th>$\chi^2$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>54F</td>
<td>A-B</td>
<td>2.27±0.10</td>
<td>1.81±0.11</td>
<td>9.28</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>A-C</td>
<td>1.80±0.11</td>
<td>1.79±0.11</td>
<td>0.01</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>A-D</td>
<td>1.79±0.11</td>
<td>1.57±0.11</td>
<td>1.35</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>B-C</td>
<td>1.77±0.10</td>
<td>1.66±0.11</td>
<td>0.48</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>B-D</td>
<td>2.06±0.10</td>
<td>1.71±0.11</td>
<td>5.87</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>C-D</td>
<td>1.08±0.13</td>
<td>1.10±0.13</td>
<td>0.02</td>
<td>0.88</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>1.80</td>
<td>1.61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>46F</td>
<td>A-B</td>
<td>1.39±0.12</td>
<td>1.44±0.12</td>
<td>0.08</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>A-C</td>
<td>1.18±0.13</td>
<td>1.03±0.14</td>
<td>0.59</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>A-D</td>
<td>1.25±0.13</td>
<td>1.32±0.12</td>
<td>0.15</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td>B-C</td>
<td>1.23±0.13</td>
<td>1.12±0.13</td>
<td>0.36</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>B-D</td>
<td>1.14±0.13</td>
<td>1.35±0.12</td>
<td>1.39</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>C-D</td>
<td>0.60±0.20</td>
<td>0.67±0.20</td>
<td>0.05</td>
<td>0.82</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>1.13</td>
<td>1.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t$-test</td>
<td></td>
<td>$p = &lt;0.01$</td>
<td>$p = &lt;0.01$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**STUDY II: BALANCED MODE DUO-Trio**
*(KIM & LEE, IN PREP.)*

Comparison of performance between different reference

- Affects/preference towards certain sample could induce more stable memory representations of those samples and may allow a more efficient form of information processing.
**SENSORY DIFFERENCE TESTS**

- Various paradigms of difference tests can be used for assessment of sensory differences between confusible food samples in the analytical sensory evaluation of food.

**Test protocol** | **Sample presentation & Instruction**
--- | ---
A-Not A | Is this ‘A’ or not?
2-AFC | Which one is ‘A’?
3-AFC | Which is the reference?
Duo-trio | Which pair is the same pair?
Triangle | Is this pair same or different?
Same-different | Subjective affects may play a role. Segments can exist.

**Methods should involve consumers’ natural attention and perception.**

**Not just general processing strategy, but the consumers’ affects towards the samples may also need to be understood.**
FACTORs & MODELS EXPLAINING VARIABILITY IN DIFFERENCE TEST PERFORMANCES

1. Effects of involving hedonic state of mind on consumer discrimination
2. Effects of test designs & Thurstonian Modeling/Signal Detection Theory
3. Effects of order of sample presentation in a test & Sequential Perception Analysis

Factors influencing perception
- Cognitive perception strategy
- Function of Sensory System for Food

Factors affecting proportion of correct responses ($P_c$)
- Cognitive decision strategy
- Response bias
Currently the most advanced psychometric approach to modeling various sensory difference tests accounting for differences in decision process
1. Assuming variation in sensory perception

Each sample presentation yields a value of a perceptual decision variable.

Repeated presentations do not always lead to the same results, but generate a distribution of perceptual values.

For two confusable samples, the average difference of the perceptual value represents the index of absolute sensitivity difference or discriminability.
2. Accommodation of the decision strategy used in the process of judgment

- Triangle
- Which is odd one?

- 3-AFC
- Which one is stronger?

- Guessing probability is 0.333 for both.

- TM & SDT model predicts the probability of proportion of correct responses to be different accounting for their cognitive decision strategy: Triangle < 3-AFC

- Thus TM & SDT model computes the absolute distance measure, \( d' \) (sensory difference, discriminability), independent of test procedures used for the discriminations.
Perceptual modeling based on one-dimensional perceptual space and possible cognitive decision strategy in theory.

**Triangle:** ‘Comparison of distances’ (COD) strategy
- Group together the two most similar.

**3-AFC:** ‘Skimming’ strategy
- More efficient strategy, resulting in higher $P(c)$ than triangle.
- “skim off” or choose the highest intensity.
THURSTONIAN PROBABILISTIC MODELING
BASED ON ONE-DIMENSIONAL PERCEPTUAL SPACE

**Triangle**

Perceptual presentation & Decision rule

\[ P(c) = P (|x_2-x_1|<|y-x_1| \text{ and } |x_2-x_1|<|y-x_2|) \]

Psychometric function

\[ P_c = 2 \int_0^\infty \left[ \Phi \left( -u\sqrt{3} + \delta \sqrt{\frac{2}{3}} \right) + \Phi \left( -u\sqrt{3} - \delta \sqrt{\frac{2}{3}} \right) \right] \exp \left( -\frac{u^2}{2} \right) / \sqrt{2\pi} du \]

**3-AFC**

Perceptual presentation & Decision rule

\[ P(c) = P ( y > x_1 \text{ and } y > x_2 ) \]

Psychometric function

\[ P_c = \int_{-\infty}^{\infty} \Phi^2(u)\Phi(u - \delta)du \]
**THURSTONIAN PROBABILISTIC MODELING BASED ON ONE-DIMENSIONAL PERCEPTUAL SPACE**

### Duo-Trio

**Perceptual presentation & Decision rule**

- \[ P(c) = P(|x_2 - x_1| < |y - x_1|) \]

**Psychometric function**

\[ P_c = 1 - \Phi \left( \frac{\delta}{\sqrt{2}} \right) - \Phi \left( \frac{\delta}{\sqrt{6}} \right) + 2\Phi \left( \frac{\delta}{\sqrt{2}} \right) \Phi \left( \frac{\delta}{\sqrt{6}} \right) \]

### 2-AFC

**Perceptual presentation & Decision rule**

- \[ P(c) = P(y > x) \]

**Psychometric function**

\[ P_c = \Phi \left( \frac{\delta}{\sqrt{2}} \right) \]
## Cognitive Decision Strategy

<table>
<thead>
<tr>
<th>Types</th>
<th>Terminology used in Psychology</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Based on one-dimensional Thurstonian model</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Skimming strategy</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Comparison of distances (COD) strategy</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Based on decision space (SDT)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| **β-decision strategy** | Subject sets a β-criterion at some level of sensory information and makes judgment based on where the information from the samples in a test falls about this criterion | • **Independent observation rule**  
• **Optimal decision rule** | • Green & Swets 1966  
• Hautus, van Hout & Lee, 2009 |
| **τ-decision strategy** | Subject sets a criterion difference (the τ-criterion) that is compared to the difference in sensory information that arises from two or more samples. | • **Differencing strategy**  
• **Sensory difference decision rule** | • Macmillan & Creelman 2005 |
Possible cognitive decision strategy in theory

**τ-decision strategy**

- Difference distribution
- "Same" or "Different"
- One-dimensional decision space

First sample (interval 1)

Second sample (interval 2)

"Same" or "Different"

More efficient strategy, resulting in higher $P(c)$
<table>
<thead>
<tr>
<th>Cognitive Decision Strategy (Cont.)</th>
<th>Types</th>
<th>Relevant Tests (in theory)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Based on One Dimensional Thurstonian Model</td>
<td><strong>Skimming strategy</strong></td>
<td>• m-AFC ($m &gt; 2$)</td>
</tr>
</tbody>
</table>
| | **Comparison of distances strategy** | • Triangle  
• Duo-trio |
| Based on Decision Space (SDT) | **β-decision strategy** | • A-Not A  
• 2-AFC  
• Same-different  
• Dual-Pair  
• Triangle  
• Duo-trio |
| | **τ-decision strategy** | • Same-different  
• Dual-Pair  
• 2-AFC  
• Duo-trio |

**Skimming strategy**

- m-AFC ($m > 2$)

**Comparison of distances strategy**

- Triangle
- Duo-trio

**β-decision strategy**

Subject sets a β-criterion at some level of sensory information and makes judgment based on where the information from the samples in a test falls about this criterion

- A-Not A
- 2-AFC
- Same-different
- Dual-Pair
- Triangle
- Duo-trio

**τ-decision strategy**

Subject sets a criterion difference (the τ-criterion) that is compared to the difference in sensory information that arises from two or more samples.

- Same-different
- Dual-Pair
- 2-AFC
- Duo-trio
STUDY I REVISITED

AFFECTIVE SAME-DIFFERENT DISCRIMINATION TESTS FOR ASSESSING CONSUMER DISCRIMINABILITY BETWEEN MILKS WITH SUBTLE DIFFERENCES

Hypothesis:
When consumers have been exposed to the products and involved in making judgments about individual products, they might then just try to identify the products using a $\beta$-criterion rather than using the commonly assumed $\tau$-criterion.

Objective 2: exploring the effects of the previous task on the cognitive decision strategies used in the same-different tests.

Chae, Lee, & Lee (2009) Food Quality & Preference 18, 920-928
Signal Detection Theory

- The shape of ROC analysis
- Investigation of $d'$ estimates in comparison with the standard detection method such as A-Not A (yes-no) and 2-AFC.

Based on the $\chi^2$-goodness of fit test, for all the data from each session, a model assuming β-strategy gave the better fitting.

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Sample pair</th>
<th>$\tau$-Strategy</th>
<th>$\beta$-Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$d'$</td>
<td>$\chi^2$</td>
</tr>
<tr>
<td>SD/Af.F</td>
<td>A-B</td>
<td>2.46±0.07</td>
<td>14.84</td>
</tr>
<tr>
<td></td>
<td>A-C</td>
<td>1.99±0.08</td>
<td>20.65</td>
</tr>
<tr>
<td></td>
<td>A-D</td>
<td>1.90±0.08</td>
<td>13.61</td>
</tr>
<tr>
<td></td>
<td>B-C</td>
<td>1.99±0.08</td>
<td>14.90</td>
</tr>
<tr>
<td></td>
<td>B-D</td>
<td>1.92±0.08</td>
<td>24.41</td>
</tr>
<tr>
<td></td>
<td>C-D</td>
<td>0.82±0.11</td>
<td>5.95</td>
</tr>
<tr>
<td>Mean</td>
<td>1.85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD/An.F</td>
<td>A-B</td>
<td>1.89±0.08</td>
<td>24.47</td>
</tr>
<tr>
<td></td>
<td>A-C</td>
<td>1.69±0.08</td>
<td>4.91</td>
</tr>
<tr>
<td></td>
<td>A-D</td>
<td>1.67±0.09</td>
<td>10.23</td>
</tr>
<tr>
<td></td>
<td>B-C</td>
<td>1.43±0.08</td>
<td>15.00</td>
</tr>
<tr>
<td></td>
<td>B-D</td>
<td>1.80±0.08</td>
<td>9.26</td>
</tr>
<tr>
<td></td>
<td>C-D</td>
<td>0.87±0.10</td>
<td>5.53</td>
</tr>
<tr>
<td>Mean</td>
<td>1.56</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Based on the $\chi^2$-goodness of fit test, for all the data from each session, a model assuming $\beta$-strategy gave the better fitting.

- This suggests that when complex food products are compared and the products are pre-viewed, consumers can use more efficient (optimal) decision strategy than the commonly assumed $\tau$-strategy.

- The nature of the product category (milk) might have been an influence on the decision strategy used in the same-different test.

- Therefore, when analyzing the same-different tests, checking the model’s validity and justifying the decision strategy is needed for accurate $d'$ computation.

THE IMPORTANCE OF KNOWING WHAT THE TEST DESIGN IS
In food science literature, for many commonly used difference tests, there are no agreed standard format. As long as the general procedure is followed, the method is given its name.

- "Lack of agreement on terminology"
- "Difficulty in deciding among competing models for the same test design"
- Without valid modeling, data like $P_{(c)}$ can be misinterpreted.

There are many different versions of the A-Not A test
(Meilgaard, Civille & Carr, 1991; Lawless & Heymann, 1996; Lee, van Hout & O’Mahony, 2007)

Familiarization before the tests to describe the dimension

- Only ‘A’
- Both ‘A’ & ‘Not A’

- Is this ‘A’ or not?
- Reminder during tests
- Sometimes available, other times not
- The different methods have the potential to change the cognitive decision strategy being used.
- If there were changes in decision strategy, comparisons of the discrimination indices between methods would be problematical.
A-Not A: SDT

A-Not A (yes-no)

Is this ‘A’ or not?

β- strategy

A-Not A Reminder

Reminder ‘A’

Decisionally separable boundary

β- strategy

τ- strategy

Comparison to a fixed product difference as a criterion

Hautus, van Hout, & Lee (2009) Food Quality & Preference 20, 222-229
A-Not A: SDT

A-Not A (yes-no)

Is this ‘A’ or not?

A-Not A Reminder

Reminder ‘A’

Decisionally separable boundary

β- strategy

Sometimes, it was called as “single-reference same-different test” (Rousseu et al., 1999)

Comparison to a fixed product difference as a criterion

Hautus, van Hout, & Lee (2009) Food Quality & Preference 20, 222-229
There are many different versions of the duo-trio test

**Balanced reference mode**
- Variable between ‘A’ & ‘Not A’

**Constant reference mode**
- Constantly ‘A’

Reference

Reference

Reference

Reference
There are many different versions of the duo-trio test


**Balanced reference mode**
- Variable between ‘A’ & ‘Not A’

**Constant reference mode**
- Constantly ‘A’

**2-AFC Reminder**
- Decisionally separable boundary

**2-AFC τ-/β- strategy**
- Comparison to a fixed product difference as a criterion

**2-AFCR τ- strategy**
- Fixed design
PROBABILISTIC MODELING (SDT)

2-AFC

Perceptual presentation & Decision rule

\[ P(c) = P(y > x) \]

Psychometric function

\[ P_c = \Phi\left(\frac{\delta}{\sqrt{2}}\right) \]

2-AFCR

Perceptual presentation & Decision rule

\[ P(c) = P(y_{x_1} > x_{2-x_1}) \]

Psychometric function

\[ P_c = \Phi\left(\frac{\delta}{\sqrt{2}}\right) \]

Hautus, van Hout, & Lee (2009) Food Quality & Preference 20, 222-229
Probabilistic Modeling (SDT)

2-AFC

Perceptual presentation & Decision rule

• $P(c) = P(y > x)$

Is this model valid for food discriminations?

It is an empirical question. More experiments are needed.

2-AFCR

Psychometric function

$P_c = \Phi(\delta / \sqrt{2})$

Hautus, van Hout, & Lee (2009) Food Quality & Preference 20, 222-229
# Classical Classification: What Property of Food Are We Comparing?

<table>
<thead>
<tr>
<th>Test protocol</th>
<th>Sample presentation &amp; Instruction</th>
<th>Attribute or specified difference test</th>
<th>Overall difference test</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-Not A</td>
<td>Is this ‘A’ or not?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-AFC</td>
<td>Which one is ‘A’?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-AFC</td>
<td>Which one is stronger?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duo-trio</td>
<td>Which one is the reference?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Triangle</td>
<td>Which is odd one?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dual-pair</td>
<td>Which pair is the same pair?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Same-different</td>
<td>Is this pair same or different?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(e.g. Meilgaard et al., 1999; Bi, 2006)
**CLASSICAL CLASSIFICATION:**

**WHAT PROPERTY OF FOOD ARE WE COMPARING?**

<table>
<thead>
<tr>
<th>Presentation &amp; Instruction</th>
<th>A-Not A</th>
<th>2-AFC</th>
<th>3-AFC</th>
<th>Duo-trio</th>
<th>Triangle</th>
<th>Dual-pair</th>
<th>Same-different</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-Not A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-AFC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-AFC</td>
<td>Which one is ‘A'?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duo-trio</td>
<td>Which one is the reference?</td>
<td></td>
<td></td>
<td>Reference</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Triangle</td>
<td>Which is odd one?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dual-pair</td>
<td>Which pair is the same pair?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Same-different</td>
<td>Is this pair same or different?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Authenticity test: Affective A-Not A

Is this ‘Foreign’ or not?

Conceptual values describing the synthetic sensory perception can also be used to define the decision space for consumer discrimination tests (e.g. Meilgaard et al., 1999; Bi, 2006)
**CLASSICAL CLASSIFICATION:**
WHAT PROPERTY OF FOOD ARE WE COMPARING?

<table>
<thead>
<tr>
<th>Test protocol</th>
<th>Sample presentation &amp; Instruction</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-Not A</td>
<td>Is this ‘A’ or not?</td>
</tr>
<tr>
<td>2-AFC</td>
<td>Which one is ‘A’?</td>
</tr>
<tr>
<td>3-AFC</td>
<td>Which one is stronger?</td>
</tr>
<tr>
<td>Duo-trio</td>
<td>Which one is the reference?</td>
</tr>
<tr>
<td>Triangle</td>
<td>Which is odd one?</td>
</tr>
<tr>
<td>Dual-pair</td>
<td>Which pair is the same pair?</td>
</tr>
<tr>
<td>Same-different</td>
<td>Is this pair same or different?</td>
</tr>
</tbody>
</table>

With appropriate **familiarization** to induce a synthetic perception, **m-AFC** can also be used as identification test based on overall sensory perception just like A-Not A test. (e.g. Meilgaard et al., 1999; Bi, 2006)
**Proposed Classification:**

**Do we have ‘a perception’ that can be referenced?**

<table>
<thead>
<tr>
<th>Test protocol</th>
<th>Sample presentation &amp; Instruction</th>
<th>Identification test with a fixed reference (reminder)</th>
<th>Classification test with a variable reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-Not A</td>
<td>Is this ‘A’ or not?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-AFC</td>
<td>Which one is ‘A’?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-AFC</td>
<td>Which one is ‘A’?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duo-trio</td>
<td>Which one is the reference?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Triangle</td>
<td>Which is odd one?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dual-pair</td>
<td>Which pair is the same pair?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Same-different</td>
<td>Is this pair same or different?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Proposed Classification:**

*Do we have 'a perception' that can be referenced?*

<table>
<thead>
<tr>
<th>Test protocol</th>
<th>Sample presentation &amp; Instruction</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-Not A</td>
<td>Is this 'A' or not?</td>
</tr>
<tr>
<td>2-AFC</td>
<td>Which one is 'A'?</td>
</tr>
<tr>
<td>3-AFC</td>
<td>Which one is 'A'?</td>
</tr>
<tr>
<td>Duo-trio</td>
<td>Which one is the reference?</td>
</tr>
<tr>
<td>Triangle</td>
<td></td>
</tr>
<tr>
<td>Dual-pair</td>
<td></td>
</tr>
<tr>
<td>Same-different</td>
<td></td>
</tr>
</tbody>
</table>

Identification test with a fixed reference (reminder)

When more familiar, or preferred sample is known, **applying a fixed reference discrimination design such as “2-AFC reminder” identification** rather than a variable reference design might be more suitable.
**PROPOSED CLASSIFICATION:**

Do we have ‘a perception’ that can be referenced?

<table>
<thead>
<tr>
<th>Test protocol</th>
<th>Sample presentation &amp; Instruction</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-Not A</td>
<td></td>
</tr>
<tr>
<td>2-AFC</td>
<td></td>
</tr>
<tr>
<td>3-AFC</td>
<td></td>
</tr>
<tr>
<td>Duo-trio</td>
<td></td>
</tr>
<tr>
<td>Triangle</td>
<td></td>
</tr>
<tr>
<td>Dual-pair</td>
<td></td>
</tr>
<tr>
<td>Same-different</td>
<td></td>
</tr>
</tbody>
</table>

More research on consumer discrimination test design are being conducted!

May be, I can tell you more coming years!!!
FACTORS & MODELS EXPLAINING VARIABILITY IN DIFFERENCE TEST PERFORMANCES

1. Effects of involving hedonic state of mind on discrimination

Factors influencing perception
- Cognitive perception strategy
- Function of Sensory System for Food

Factors affecting proportion of correct responses ($P_c$)
- Cognitive decision strategy
- Response bias

2. Effects of test designs & Thurstonian Modeling/Signal Detection Theory

3. Effects of order of sample presentation on a test & Sequential Perception Analysis
THE EFFECTS OF THE ORDER OF SAMPLES PRESENTED ON A TEST ON THE TEST PERFORMANCE

Achieving accuracy & efficiency in sensory data analysis
Way we need to study the effects of the order of samples presented on a test?

- **From a theoretical perspective**, we can learn more about the functioning of the sensory system and how it integrates information over time.
  - Learning particularly with regard to the 'complex' senses involved in flavor perception.
- **From a practical perspective**, we can learn how to optimize our sensory difference tests to best take advantage of 'favorable' orders of presentation.
  - Increasing the sensitivity of those tests to effectively measure perceptual differences.
<table>
<thead>
<tr>
<th>Factors Affecting the functioning of the sensory system for flavor perception</th>
<th>Relevant Factors</th>
<th>Test Influences</th>
<th>References</th>
</tr>
</thead>
</table>
| **Fatigue**  
physical and mental disfunction |  | • Number of samples  
• Inter-stimulus interval and rinses | • Lee & O’Manony, 2007a,b |
| **Adaptation**  
physiological desensitization |  | • Number of samples  
• Sequence of sample presentation | • Lee & O’Manony, 2007a,b  
• Lee, chae & Lee, 2009  
• O’Manony, 1974, 1979 |
| **Contrast**  
Physiological and psychological sensitization |  | • Sequence of sample presentation | • Lee & O’Manony, 2007a,b  
• Lee, 2008  
• Lee, chae & Lee, 2009 |
**Effects of Order of Tasting**

- **Adaptation effects**
  - the more intense the more desensitizing
    - O’Mahony & Odbert, 1987
    - O’Mahony and Goldstein, 1987
  - The larger number of tasting the more desensitizing
    - Lau, O’Mahony & Rousseu, 2004

- **Contrast effects**
  - the more contrasting the more recognizable
    - Lee & O’Mahony, 2007
    - Lee, Chae & Lee, 2009
    - Dessirier, Siffermann & O’Mahony, 1999
A four distribution Thurstonian model accounting for order effects from one prior stimulus, based on Tedja et al. (1994)’s data.

The test sensitivity was compared by considering the confusability between perceptual distributions.
The perceptual model was further modified based on bipolar synthetic perceptual dimension. This model accounted for not only physiological adaptation affected by the one previous sample but also cognitive contrasts between comparison samples that are tasted in sequence.
Sources of Response Bias
## Response Bias

<table>
<thead>
<tr>
<th>Factors lowering proportion of correct responses</th>
<th>Relevant Factors</th>
<th>Relevant Test Design</th>
<th>References</th>
</tr>
</thead>
</table>
| **Criterion variation**                         | • A-Not A        |                      | • Rousseu & O’Manony, 2002  
|                                                 | • Same-different |                      | • Lee and O’Mahony, 2004    |
| **Position of presentation**                    | • Tests having multiple sample presentations |                      | • Lee & O’Manony, 2007a,b  
|                                                 |                  |                      | • Nisbett & Wilson, 1977    |
| **Time order error (memory decay)**             | • Temporal test (When samples are separated by time rather than space)  
|                                                 | • 2-AFC          |                      | • Lee, chae & Lee, 2009    
|                                                 | • 3-AFC          |                      | • Wilson and Tanner 1961   
|                                                 | • Same-different |                      | • Berliner and Durlach 1973|
|                                                 |                  |                      | • Cubero et al.,1995       
|                                                 |                  |                      | • Avancini de Almeida et al., 1999 |
IMPERFECT MEMORY (TIME-ORDER ERROR)

- The greater the inter-stimulus interval the greater the bias
  - Cubero et al., 1995
  - Avancini de Almeida et al., 1999
    - same-different test

- The more recently tasting the more intense
  - Lee O’Mahony, 2007
    - With inter-stimulus rinsing
  - Lee, Chae & Lee, 2009
    - With water-inter-stimulus rinsing

- The more difficult the task to deal with the more error
  - Rousseu & O’Mahony, 1997
  - Dessirier & O’Mahony 1999
  - Lau, O’Mahony & Rousseu, 2004
A new model incorporating memory bias as well as adaptation and cognitive contrasts.

This predicted the position effects caused by the order of sample presentation in a 3-AFC using a skimming strategy.
A new model incorporating memory bias as well as adaptation and cognitive contrasts predicted the position effects caused by the order of sample presentation in a 3-AFC using a skimming strategy.

How this SPA model would apply to Duo-trio?

Water inter-stimuli rinsing: the best, WSS
HOW DISCRIMINATION METHODS BECOME MORE DISCRIMINATING

Variable Duo-Trio Procedures
1. The traditional Duo-Trio

2. The Duo-Trio with the reference tasted in the Middle, between the two test samples (DTM)
   - Orange-flavored beverage
   - Fixed design
   - Water inter-stimulus rinsing
   - Tested sequences: SSW, WWS

**Duo-Trio & DTM**

(Rousseau & O’Mahony, 2002)
**DUO-TRIO, DTM & DTFR**  
*(LEE & KIM, 2008)*

1. The traditional **Duo-Trío** with the reference tasted **First** *(DTF)*

2. The **Duo-Trío** with the reference tasted in the **Middle**, between the two test samples *(DTM)*

3. The **Duo-Trío** with the reference tasted twice, **First and last as a Reminder** *(DTFR)*

Reference is also directly compared to the comparison sample.
• Using salt model systems, in a roving design to use comparison of distances (COD) decision strategy
• No rinsing between samples
• Tested all sequences
STUDY III

COMPARISON OF $d'$ ESTIMATES PRODUCED BY THREE VERSIONS OF A DUO-TRIO TEST FOR DISCRIMINATING TOMATO JUICES WITH VARYING SALT CONCENTRATIONS: THE EFFECTS OF THE NUMBER AND POSITION OF THE REFERENCE STIMULUS

Objectives:

- To investigate the relative performance of the DTF, DTM and DTFR, focusing on the comparison of distances (COD) strategy
- To examine the sensitivity predictions from the new SPA model for the same sequences of DTF and DTM

**Duo-Trio, DTM & DTFR**  
*(Kim, Lee & Lee, 2010)*

- Tomato juice
- Roving design to use comparison of distances (COD) decision strategy
- Tested all sequences

**All sequences**

**Only WWS, SSW**

![Graphs showing d' estimates for different sequences](image_url)
### SPA Prediction for WWS, SSW

#### Applying comparison of distances (COD) decision strategy

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Protocol</th>
<th>Triadic sequences for each protocol</th>
</tr>
</thead>
<tbody>
<tr>
<td>1&lt;sup&gt;st&lt;/sup&gt;</td>
<td>SPA</td>
<td>W ← S → S ← W → S ← W ← S → S ← W</td>
</tr>
<tr>
<td>2&lt;sup&gt;nd&lt;/sup&gt;</td>
<td></td>
<td>W ← S ← W ← S ← W ← S ← W ← S ← W</td>
</tr>
<tr>
<td>3&lt;sup&gt;rd&lt;/sup&gt;</td>
<td></td>
<td>W ← S ← W ← S ← W ← S ← W ← S ← W</td>
</tr>
</tbody>
</table>

**Position of stimuli**
- W: Weaker
- S: Stronger
- Perception of stimuli accounting for contrast appearing in duo-trio test

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Protocol</th>
<th>Triadic sequences for each protocol</th>
</tr>
</thead>
<tbody>
<tr>
<td>1&lt;sup&gt;st&lt;/sup&gt;</td>
<td>SPA</td>
<td>W ← S → S ← W → S ← W ← S → S ← W</td>
</tr>
<tr>
<td>2&lt;sup&gt;nd&lt;/sup&gt;</td>
<td></td>
<td>W ← S ← W ← S ← W ← S ← W ← S ← W</td>
</tr>
<tr>
<td>3&lt;sup&gt;rd&lt;/sup&gt;</td>
<td></td>
<td>W ← S ← W ← S ← W ← S ← W ← S ← W</td>
</tr>
</tbody>
</table>

**Sensitivity rank based on COD strategy**
- 1<sup>st</sup> WWS
- 2<sup>nd</sup> SSW

**SPA 8-distribution prediction**
- DTF
- DTM

For SPA prediction, apply the comparison of distances (COD) decision strategy. There is no inter-stimuli rinsing.

**Example**: For WWS, the sequence is W ← S → S ← W → S ← W ← S → S ← W. The sensitivity rank based on the COD strategy is 1<sup>st</sup>.
## SPA PREDICTION FOR WWS, SSW

### Applying comparison of distances (COD) decision strategy

#### Water inter-stimuli rinsing

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Protocol</th>
<th>Triadic sequences for each protocol</th>
<th>Position of stimuli</th>
<th>Perception of stimuli accounting for contrast appearing in duo-trio test</th>
<th>Sensitivity rank based on COD strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DTF</td>
<td>W&lt;sub&gt;x&lt;/sub&gt; W W S S</td>
<td>W&lt;sub&gt;x&lt;/sub&gt;</td>
<td>W&lt;sub&gt;x&lt;/sub&gt; W W S S</td>
<td>1&lt;sup&gt;st&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>DTM</td>
<td>W W S S W&lt;sub&gt;x&lt;/sub&gt; W&lt;sub&gt;x&lt;/sub&gt;</td>
<td>W W S S W&lt;sub&gt;x&lt;/sub&gt; W&lt;sub&gt;x&lt;/sub&gt;</td>
<td>W W S S W&lt;sub&gt;x&lt;/sub&gt; W&lt;sub&gt;x&lt;/sub&gt;</td>
<td>2&lt;sup&gt;nd&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>DTF</td>
<td>S&lt;sub&gt;x&lt;/sub&gt; S W W S</td>
<td>S&lt;sub&gt;x&lt;/sub&gt;</td>
<td>S&lt;sub&gt;x&lt;/sub&gt; S W W S</td>
<td>2&lt;sup&gt;nd&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>DTM</td>
<td>S S W W S&lt;sub&gt;x&lt;/sub&gt; S&lt;sub&gt;x&lt;/sub&gt;</td>
<td>S S W W S&lt;sub&gt;x&lt;/sub&gt; S&lt;sub&gt;x&lt;/sub&gt;</td>
<td>S S W W S&lt;sub&gt;x&lt;/sub&gt; S&lt;sub&gt;x&lt;/sub&gt;</td>
<td>1&lt;sup&gt;st&lt;/sup&gt;</td>
</tr>
</tbody>
</table>
SPA PREDICTION FOR WWS, SSW

Applying comparison of distances (COD) decision strategy

- The difference between DTM and DTF was successfully explained by the SPA model. The differential results found between Kim et al. (2010) and Rousseau & O’Mahony (2002) could be due to the differential inter-stimuli rinsing scheme.

- But it is still possible that there might be a difference in the cognitive decision strategy used for the experiment.
  - Kim et al (2010): a roving design
  - Rousseau & O’Mahony (2002): a fixed design
Why would the Duo-Trio with the reference tasted twice, First and last as a Reminder (DTFR) perform better?

Is this test same as the “AB-X (matching to the sample)”?

- It is possible that the first reference was tasted as a mere primer in the DTFR and the task was performed as the ABX design.
- It can also be hypothesized that when the DTFR is used as the ABX, subjects may be able to use a $\beta$-decision strategy.
- This is a topic for future research.
FACTORS INFLUENCING PERFORMANCES OF DIFFERENCE TESTS

- Decision strategy
- Response bias

- Cognitive perception strategy
- Affective/hedonic state of mind

- Factors determining the nature and clarity of the perceived information

- Order of samples on a test

Factors determining proportion of correct responses ($P_c$)
In order to predict accurate sensory difference or discriminability in flavor discrimination as an index comparable across different experiments, **no matter whether you use P(c) or Signal Detection measures such as d’,**

the test procedure and experimental context should be carefully standardized in a way that:

1) **Appropriate attention can be driven to the food sample.**
2) **A decision rule could be applied in a consistent manner.**
3) **Physiological and cognitive interference can be minimized.**

The TM/SDT models are currently the most advanced model accommodating decision strategy used for the test method.

Extended TM/SDT model should also be explored to take account for the physiological and cognitive complication in the temporal flavor discrimination.
ACKNOWLEDGEMENT

Ewha Womans University

- Prof. Kwang-Ok Kim
- Food Design/Ergonomics Lab members:
  - Young-Mi Lee, Min-A Kim, Ji-Eun Chae
  - Yoon-Jung Choi

Collaborators

Prof. Michael O’Mahony

Unilever “Sensation Perception & Behaviour” team, Danielle van Hout

Dr. Michael J. Hautus
Thank you