Using Orthonasal Aroma Evaluation to Predict Consumer Liking

Christopher T. Simons, Ph.D.
Givaudan Flavors Science & Technology
Cincinnati, OH

Society of Sensory Professionals
November 2008
Limitations using traditional consumer testing protocols

To accurately predict liking, large multifactor designs require panelists to evaluate a considerable number of samples.

- The number of samples a consumer can evaluate in a single sitting is limited.
  - Depending on the product, the maximum number of samples ranges from 5-10.
Limitations using traditional consumer testing protocols

To accurately predict liking, large multifactor designs require panelists to evaluate a considerable number of samples.

- Several factors limit the number of products that a consumer can evaluate in a single sitting.
  - Satiety
  - Carry-over
  - Adaptation
Limitations using traditional consumer testing protocols

To accurately predict liking, large multifactor designs require panelists to evaluate a considerable number of samples.

- Sample preparation requires support from applications groups.
- Logistical support (labeling cups, serving utensils, maintaining product temperature, etc.) requires significant planning.
Limitations using traditional consumer testing protocols

We hypothesized that using orthonasal evaluations to collect hedonic information of flavor samples might overcome some of the limitations associated with traditional consumer testing.

- Issues surrounding satiety are minimized when smelling.
- Relatively easy to cleanse the “nasal palette” and therefore sample carry-over and adaptation can be reduced.
- Sample preparation and logistical support is easier.
  - Cartridges easy to fill.
  - Applications support not needed until validation step.
Virtual Aroma Synthesizer (VAS) and MiniVAS™

- Givaudan’s proprietary flavor creation tools
- It combines various aromas into a single aroma profile.
- Computerized controls
- Proprietary programs translate the smell to taste and provide a beginning flavor formulation.
MiniVAS Operating Principle

1. VAS software
2. Adjust ingredient intensity
3. Mixing tube
4. Air flow
5. Aroma of base product + aroma chemical
6. Key A
7. Key B
8. Key C
9. Base
To develop methodology, two key questions had to be answered:

- How many samples can be evaluated by a consumer before they become fatigued or bored?
  - Must be substantially more than 10 samples otherwise methodology not worth the investment.

- How do evaluations by nose compare to evaluations by mouth?
  - Do orthonasal and retronasal perceptions differ? If they do, is it a problem for the methodology?
  - What you like by smell do you like by taste?
Can we evaluate more samples by sniff than taste?

- How many samples can a person evaluate before becoming fatigued or bored?
- Hypothesis: If a person becomes fatigued or bored, their hedonic ratings for a replicated sample will display non-random variation.
### Special Causes Analysis

- Special cause variation is caused by known factors (such as fatigue or lack of motivation) that result in a non-random distribution of output.
- A statistical technique used routinely in quality control.
- Special Causes (non-random variation)
  - 1 point more than $3\sigma$ from center line
  - 9 points in a row on same side of center line
  - 6 points in a row, all increasing or decreasing
  - 14 points in a row alternating up and down
  - 2 out of 3 points > $2\sigma$ from center line (same side)
  - 4 out of 5 points > $1\sigma$ from center line (same side)
  - 15 points in row within $1\sigma$ from center line (either side)
  - 8 points in row >$1\sigma$ from center line (either side)
Can we evaluate more samples by sniff than taste?

Study design

- Consumers:
  - Adults: ages 18-66
  - Kids: ages 8-12

- Sample presentation:
  - 25 strawberry flavor types
  - 111 samples presented without break or feedback
  - 1 sample repeated every 2-3 samples (tracking sample T)
  - Tracking sample always preceded by anchor sample (A)

1 2 A T 3 4 5 A T 6 7 A T 8 9 A T 10 11 12 AT ... 111
Can we evaluate more samples by sniff than taste?

Results: Adults (age range 18-66 years, mean 39±12 yoa)

Average # of samples evaluated prior to signs of fatigue or boredom

<table>
<thead>
<tr>
<th>Gender</th>
<th>Samples Evaluated</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>80±4</td>
</tr>
<tr>
<td>♀ (n=61)</td>
<td>83±4</td>
</tr>
<tr>
<td>♂ (n=44)</td>
<td>77±4</td>
</tr>
</tbody>
</table>

Cumulative number of panelists showing signs of fatigue or boredom relative to sample number

Special Causes Analysis

- Mean (X): 5.643
- Upper Control Limit (UCL): 6.154
- Lower Control Limit (LCL): 5.132

- Standard Deviation (S): 1.7332
- Upper Control Limit (UCL): 2.0959
- Lower Control Limit (LCL): 1.3705
Can we evaluate more samples by sniff than taste?

Results: Kids (age range 8-12 years, mean 10±2 yoa)

Average # of samples evaluated prior to signs of fatigue or boredom

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>70±6</td>
<td></td>
</tr>
<tr>
<td>♀ (n=26)</td>
<td>81±6</td>
<td></td>
</tr>
<tr>
<td>♂ (n=27)</td>
<td>61±7</td>
<td></td>
</tr>
</tbody>
</table>

Cumulative number of panelists showing signs of fatigue or boredom relative to sample number

Special Causes Analysis
Can we evaluate more samples by sniff than taste?

- Typical sensory limitations allow for the evaluation of few samples when tasting.
  - Consumers can evaluate maximum 10 samples/session by taste

- Results from the current test suggest panelists can evaluate significantly more samples orthonasally than retronasally.
  - Provides advantage of being able to evaluate a greater proportion of the sensory space.
  - Enables the execution of complex multifactor predictive models.
Orthonasal-Retronasal comparisons
Does route of administration effect olfactory perception?
Prior studies have indicated that perceptions evoked via orthonasal evaluation of flavors are not identical to those evoked via retronasal evaluation.

- Threshold and Intensity differences exist.
  - Orthonasal thresholds tend to be lower than retronasal thresholds.
  - Perceived intensity is greater when odorants are delivered orthonasally.

Does the perceptual quality of a stimulus differ according to the route of administration?

- Dearth of information exists addressing this question.
  - Differential brain responses were obtained using fMRI when the same odorant (chocolate) was delivered ortho- versus retronasally.
  - Stimuli can be differentiated more easily with orthonasal evaluation compared to retronasal evaluation.
**Question:** Are descriptive profiles evoked by flavor stimuli the same when evaluated orthonasally and retronasally?
Orthonasal-retronasal comparisons

Question: Are descriptive profiles evoked by flavor stimuli the same when evaluated orthonasally and retronasally?
Orthonasal and retronasal flavor profiles are not identical when flavor rebalancing does not occur.

- Flavorists have the critical training and tools to rebalance flavors according to application type and make-up.
Smell-Taste Hedonic testing
Is liking similar when flavors are evaluated by sniff and taste?

- Chicken Flavors –
  - 108 consumers (age 18-55) evaluated 5 chicken flavors by smell (Cardsniff) and by taste (diced chicken).
  - 2 of the top 3 flavors by smell were the same as by taste (Fatty and Roasted). Lemon pepper was liked more by smell than by taste.
Cheese Flavors –
- 98 consumers (age 18-55) evaluated 5 cheese flavors by smell (Cardsniff) and by taste (cheese sauce on pasta).
- Significant differences of taste and sniff evaluations of cheese are identical.
- The same cheese flavor is rated the same as seen by smell and by taste.
- Identical correlations between sniff and taste data.

![Bar graph showing average hedonic scores for taste and sniff evaluations of different cheese flavors.](image)
Is liking similar when flavors are evaluated by sniff and taste?

**Chicken optimization study**

- DOE used to generate 67 chicken flavor samples evaluated by smell using the miniVAS.
- 8 flavors were added to chicken cubes for taste validation
  - 5 from DOE that spanned the liking continuum
    - Verify that samples significantly different by smell are significantly different by taste
  - 3 optimized flavors
    - Verify that optimized flavors tend to be liked more than non-optimized flavors
- Flavorist iterated compounding until taste profiles matched the profiles elicited when smelling on the miniVAS.
Is liking similar when flavors are evaluated by sniff and taste?

Chicken optimization study

- Results:
  - Flavors liked by smell tended to be liked by taste whereas flavors disliked by smell tended to be disliked by taste.
  - The optimized flavors tended to be liked more than other flavors.
  - Smell:Taste correlation: 0.71

Taste:Taste reliability: 0.61
Conclusions
Using orthonasal flavor evaluation in consumer liking studies…

- Overcomes many of the limitations associated with traditional consumer testing.
  - Satiety, fatigue and logistical complexities

- Significantly increases the number of samples that can be evaluated in a single sitting.
  - Enhances the rate and quality of guidance to the flavor creation process
  - Able to explore a greater proportion of the sensory space
  - Increases confidence that selected flavor is truly “consumer preferred”.
Acknowledgments

- **Givaudan Flavors**
  - Bob Antenucci
  - Christel Adam
  - Gaelle LeCourt
  - Guillaume Blancher
  - Hansruedi Gygax
  - Jason Parcon
  - Joe Kaiser
  - Lindsay Bellegia
  - Matt Burland
  - Rachel Mansencal
  - Sarah Kirkmeyer

- **J. Reckner Associates**
  - Cindy Ward

- **InsightsNow**
  - Caryn Crawford
  - Natosha Simpson

- **Dr. Pepper**
  - Emily Kimmins

- **Chiquita Brands Intl.**
  - Diego Luzuriaga