



# ***Engaging the Power of Sensory Science to Meet Quality Assurance Objectives***

**Jean Ann Hankins**

**Sensory Operations Manager, Silliker Inc. – Bentonville Facility**

**Quality Assurance Association - May 11<sup>th</sup>, 2010**

# Sensory Evaluation

- **A scientific discipline that applies principles of experimental design and statistical analysis to the use of human senses—sight, smell, taste, touch and hearing—for the purpose of evaluating consumer products**

- ◆ Evokes
- ◆ Measures
- ◆ Analyzes
- ◆ Interprets



# Business Role of Sensory Science

- **Sensory evaluation provides valid and reliable product information which can assist in expediting sound business decisions**
- **Gives you answers to questions about your product:**
  - ◆ Are these products different?
  - ◆ How much is my product liked?
  - ◆ Which is the preferred product?
  - ◆ How do these products compare to the competition in terms of texture, aroma, flavor, color and appearance?
- **Experimental design must be set based on the desired business objective of the project**
- **The QA group must define the project's objective for testing before any testing begins or is approved**

# Technical Role of Sensory Evaluation

- To conduct reliable, ethical sensory experiments using accepted sensory methods to ensure that the data being collected is the appropriate information to relate to the project's objective
- To control variables that will assure a safe and clean testing atmosphere that is fair and impartial to both the respondent and the product

# Sensory Evaluation as a Consulting Tool

- **Sensory data can be used by a number of departments in order to make business decisions**
  - ◆ Research and Development
    - ▶ Formulation development and changes
    - ▶ Bench top screenings
    - ▶ New product approvals
  - ◆ Manufacturing
    - ▶ Product reviews (cuttings)
    - ▶ **Quality Control/Assurance**
  - ◆ Marketing
    - ▶ What is the right product to fill the desired niche?
    - ▶ Positioning of product compared to competition
- **It's like working with a net!**

# Sensory for Quality Assurance

- **Big movement toward aligning sensory profiles with category leaders**
- **To assist QA in determining whether or not a product is maintaining it's intended sensory properties over time and across production runs/multiple plants locations/etc.**
- **It all comes down to “objective”**
  - Is your objective to “match” or align with another product?
  - Is your objective to roll out a product that is as “good” as a competitors?
  - Do you want to ensure that you do not limit your product's potential to be better than a competitor's or “gold standard?”
  - Is your objective to match certain sensory characters of a product but maybe not all of them?

# The QA Sensory Toolbox

- **Two basic drawers in your sensory toolbox**
  - **More than one type or category of information is often needed to answer all the sensory objectives of a project**
- **Must have multiple tools in each drawer**
  - **Different tools (or tests) are needed to collect different types of data**

**Analogy: It takes more than a hammer to build a house and it takes more than a triangle test to answer all of your sensory objectives!**

# The Drawers in Your Toolbox

- Drawer #1 is used for collecting information about people's feelings, opinions, desires and intentions ... or

## AFFECTIVE INFORMATION

- Drawer #2 is used for collecting measurements about what is perceived and how intensely ... or

## OBJECTIVE INFORMATION

# Type I—Affective Measures

- **Answers the questions:**
  - ◆ What do you think?
  - ◆ How do you feel?
- **Provides data relative to people’s opinions, feelings, beliefs and attitudes towards a product**
- **In most instances of affective testing, it is necessary to have a large number of respondents to provide enough “power” for decision making**

## Type I—Affective Measures

- **Draws upon people's imaginations, emotions and personal preferences**
  
- **What kind of data result from these types of sensory projects?**
  - ◆ Preference/ranking = Consumer Preference
  - ◆ Hedonic measurements/acceptance = like or dislike
  - ◆ Diagnostic information = too much, too little or Just About Right (JAR)

## Preference / Ranking

- **Quantitative methods used to measure either preference or acceptance of products**
  - ◆ Two products compared to each other = preference
  - ◆ Multiple products compared to each other = ranking
- **Between 50 and 100 assessors are required in order for valid conclusions to be drawn**
- **Not necessary for panelists to be users of the products, but must be “non-avoiders” or “non-rejecters” of the product**
- **May or may not include a “no preference” option**
  - ◆ Pros and cons of this—do you want to force a choice?
- **No indication of liking or disliking so prior knowledge of whether the product is liked is necessary**
- **Ballot can allow for open ended probing of why a panelist preferred the sample chosen**

# Preference

Panelist ID# \_\_\_\_\_

Name \_\_\_\_\_

## Potato Chips

Please taste the samples as they appear from left to right. After you have tasted both samples, please indicate which sample you prefer by circling the corresponding 3-digit code.

**Which one do you prefer? (Circle your preference)**

**384**

**464**

**Please tell us why you preferred the sample you circled.**

---

---

# Ranking

## Potato Chips

The sample numbers below appear in the order in which they were presented to you. In the blank space provided, please rank the samples in the order you preferred them, by placing a 1, 2 or 3 on the line next to the corresponding sample number.

Place a "1" next to the sample you most preferred, a "2" next to your second choice and a "3" next to your least preferred sample.

781 \_\_\_\_\_

962 \_\_\_\_\_

453 \_\_\_\_\_

## Possible Outcome of Preference Data

- With n=51 people
- 36 of 51 people preferred sample 384
- Comment Trend = 36 of the 51 people tested chose 384 due to “more crispy texture and fresher oil flavor”
- 36 of 51 responses would be a significant preference for sample 384 at alpha = .05
- Charts are available for fast and easy analysis of this type of data

## Acceptance/Hedonics (Liking/Disliking)

- **Objective is to determine how much people like one or more products**
  - ◆ Useful for new product concepts either standing alone or in comparison to a market leader
- **For data to be as meaningful as possible, panelists must be users of the product type being tested**
- **Demographic parameters such as usage group, usage frequency, age, sex, income, geographic location, nationality and others may be important for sampling of some products**
- **These parameters which define the target population, should always be carefully considered before choosing the demographics of the sample for a given test**
- **Most common scale used is a nine-point hedonic scale**
  - ◆ High end is “like extremely”
  - ◆ Low end is “dislike extremely”
- **Scale can be represented either by words or, when testing warrants children’s responses, facial expressions can be used.**

# Hedonics (Liking/Disliking)

1) Please observe and taste this sample. All things considered, which statement best describes your **OVERALL IMPRESSION** of this product?

Dislike	Dislike	Dislike	Dislike	Neither Like	Like	Like	Like	Like
Extremely	Very Much	Moderately	Slightly	Nor Dislike	Slightly	Moderately	Very Much	Extremely
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

2) Which statement best describes your impression of the **APPEARANCE** of this product?

Dislike	Dislike	Dislike	Dislike	Neither Like	Like	Like	Like	Like
Extremely	Very Much	Moderately	Slightly	Nor Dislike	Slightly	Moderately	Very Much	Extremely
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

3) Which statement best describes your impression of the **FLAVOR** of this product?

Dislike	Dislike	Dislike	Dislike	Neither Like	Like	Like	Like	Like
Extremely	Very Much	Moderately	Slightly	Nor Dislike	Slightly	Moderately	Very Much	Extremely
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

# Diagnostics (Too Much, Too Little, Just About Right)

- **Measures the intensity of one attribute of a product at a time**
  - ◆ Appearance = size and color
  - ◆ Aroma = chocolaty, fruity, fishy
  - ◆ Flavor includes=
    - ▶ Basic Tastes - sweetness, saltiness, sourness and bitterness
    - ▶ Aromatics- chocolaty, fruity
    - ▶ Chemical Feeling Factors- heat/burn, cooling
  - ◆ Texture = smooth, crunchy, chewy
  
- **Bipolar scale**
- **This scale is typically five categories wide**
- **Main disadvantage is that perception of what is “just right” varies across individuals**
  - ◆ Reduce this variability by increasing panel size

# Diagnostics (Too Much, Too Little, Just About Right)

1) Considering only the COLOR of this sample, which statement best describes your impression of this product?

Much Too Light

Too Light

Just About Right

Too Dark

Much Too Dark

2) Considering only the SALTINESS of this sample, which statement best describes your of this product?

Much Too Little

Too Little

Just About Right

Too Much

Much Too Much

3) Considering only the CRISPNESS of this sample, which statement best describes your overall impression of this product?

Not At All Crisp Enough

Not Crisp Enough

Just About Right

Too Crisp

Much Too Crisp

## Bacon and Cheese Hamburger Patty

	Company X	Company Y
<b>Appearance</b>	5.7 a	5.6 a
<b>Overall Impression</b>	5.8 a	6.3 a
<b>Flavor</b>	6.2 b	7.0 a
<b>Texture</b>	5.9 b	6.7 a
<b>Overall FI JAR</b>	2.9 a (29%TL, 52%JAR, 19%TM)	3.0 a (11%TL, 75%JAR, 15%TM)
<b>Beef Flavor JAR</b>	2.9 a (25%TL, 59%JAR, 16%TM)	2.9 a (13%TL, 80%JAR, 7%TM)
<b>Cheese FI JAR</b>	2.3 a (64%TL, 29%JAR, 7%TM)	2.4 a (49%TL, 48%JAR, 3%TM)
<b>Bacon FI JAR</b>	2.3 a (57%TL, 33%JAR, 9%TM)	2.3 a (59%TL, 33%JAR, 8%TM)
<b>Saltiness JAR</b>	3.0 a (13%TL, 69%JAR, 17%TM)	3.2 a (7%TL, 71%JAR, 23%TM)
<b>Texture JAR</b>	3.2 a (12%TL, 56%JAR, 32%TM)	3.1 a (9%TL, 73%JAR, 17%TM)
<b>Oiliness JAR</b>	3.2 b (13%TL, 59%JAR, 28%TM)	3.6 a (0%TL, 53%JAR, 47%TM)

Means with the different letters for each attribute are significantly different ( $p < 0.05$ ) using the Least Significance Difference method.

The original JAR scale was collapsed into a 3-point JAR scale: Too Little (TL), Just About Right (JAR) and Too Much (TM).

## Next Drawer: Type II—Objective Measures

- Provides objective data relative to a product's sensory profile
- Requires a smaller number of panelists/respondents
- Panelists must be trained in the methodology

## Type II—Objective Measures

- **Methodologies that fall under this category include the following:**

- ◆ Difference testing
  - ▶ Duo Trio
  - ▶ Paired Comparison
  - ▶ Triangle Test
  - ▶ Degree of Difference/Difference from Control

- ◆ Descriptive Analysis

## Duo Trio

- **Determines if a difference exists between two samples**
- **Assessors are presented with three samples**
  - ◆ Two are blind coded
  - ◆ One is labeled as the reference sample
- **Panelists must determine which of the blind coded samples are most similar to or different from the reference sample**
- **This technique is most useful for samples that are not homogeneous**

## Duo Trio Ballot

- **Instructions:** Taste samples from left to right. The sample furthest to the left is a reference. Determine which of the two samples matches the reference and indicate by placing an “X” in the corresponding box.
- If no difference is apparent between the two unknown samples, you must guess.

Reference

Code 261

Code 597

# Paired Comparison

- **Objective is to determine if a difference exists between two samples with regard to a specified attribute**
  - ◆ Sweetness/flavor
  - ◆ Hardness/texture
  - ◆ Intensity of aroma
- **Samples should vary in intensity ONLY of the attribute in question in order for this to be an appropriate test**
- **Useful for determining whether a product change leads to a perceived difference in a particular product attribute**
  - ◆ Formulation
  - ◆ New ingredient / supplier
  - ◆ Seasonal variation

## Paired Comparison

**Instructions:** Taste each pair of samples from left to right and enter your answer by the question below. If no difference is apparent, enter your best guess, however uncertain. “No difference” verdicts are permitted, but only as a last resort.

### TEST PAIRS

*Which sample is more salty?*

679      154

811      215

931      422

# Triangle Test

- **Objective is to determine if a difference exists between two samples**
- **Difference between this test and a duo trio is that panelists are told that two samples are the same and one is different**
  - ◆ They are asked to determine which is the odd sample
  - ◆ May be asked to describe the nature of the difference

# Triangle Test

- **Instructions:** Two of the samples are the same and one is different. Taste the samples in the order they appear from left to right and identify the odd or different sample.
- Identify the odd or different sample by circling the corresponding code.

973

775

149

*Optional:*

Please describe what you found different about the sample you circled

---

## Degree of Difference (Difference from Control)

- **Difference from Control or Degree of Difference determines**
  - ◆ That a difference exists between two samples
  - ◆ The magnitude of the difference that exists (how different are they?)
- **Useful when relative size of difference from the control is important for decision making**
- **Useful for Quality Assurance/Quality Control programs**

# Degree of Difference (Difference From Control) Ballot

SIMS2000 - Sensory Information Management System Page: 1 (This Page Incomplete)

Peanut Butter

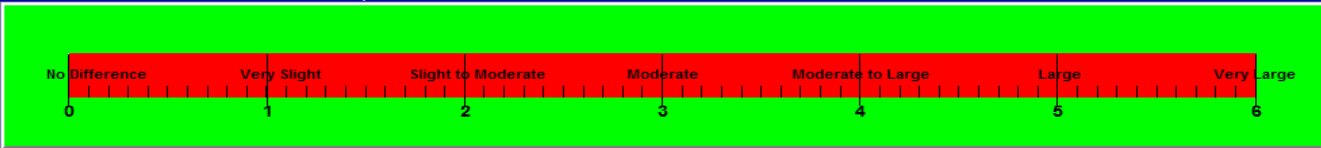
Instructions:

- 1) First observe and taste the sample marked CONTROL.
- 2) Then observe and taste the sample marked with the 3-DIGIT CODE.
- 3) Considering ALL sensory attributes, assess the overall degree of difference between the two samples.

Peanut Butter

Please indicate the degree of difference by clicking on the appropriate hash mark below.

How different are these two samples?



The scale is a horizontal bar with a red-to-yellow gradient. It is marked with integers from 0 to 6. Below the bar, the following labels are positioned: 'No Difference' at 0, 'Very Slight' at 1, 'Slight to Moderate' at 2, 'Moderate' at 3, 'Moderate to Large' at 4, 'Large' at 5, and 'Very Large' at 6. There are small tick marks between the integers.

## Explanation and Possible Outcome of DOD Data

The Student's  $t$  Test is used to indicate whether a corrected DOD mean value is significantly different from 0 (i.e. "no difference" on scale). A  $t$ -score is calculated based on the corrected DOD and the standard error of the panel. This value is compared to a table value, and if greater, the sample is "Rejected" meaning it is significantly different. If the value is less than or equal to the table value, the sample is "Accepted" meaning any differences noted are not significant.

Different?

24	# Panelist	24	# Panelist
0.029	Blind Bias	0.167	Blind Bias
0.204	average	0.442	average
0.326	sd	0.565	sd
0.107	var	0.319	var
0.067	se	0.115	se
2.627	$t$	2.385	$t$
2.500	$t$ table	2.500	$t$ table
Reject		Accept	

# Descriptive Analysis

## ■ Descriptive Analysis

- ◆ Needed when detailed sensory information (i.e. appearance, aroma, flavor and texture) must be obtained
- ◆ The output is a “sensory profile” outlining
  - ▶ Attributes of the product
  - ▶ Intensity for each attribute
- ◆ Can be used for, but is not limited to:
  - ▶ Defining sensory characteristics of a target for product development
  - ▶ Defining a standard
    - Documenting product characteristics before consumer testing
    - Tracking product sensory changes over time
- ◆ Characteristics identified in the product allow to create a product profile or thumbprint
- ◆ Complete product profile includes appearance, aroma, flavor and texture attributes

# QDA<sup>®</sup> vs. Spectrum<sup>®</sup>

Two most commonly used methods are QDA<sup>®</sup> and Spectrum<sup>®</sup>

## QDA<sup>®</sup>

- 10 – 12 panelists
  - Product specific training
  - Free lexicon
  - Unstructured line scale
  - Few scale references
  - No scale calibration
  - Data analysis by ANOVA
- Individual panelists  
Panel as a whole

## SPECTRUM<sup>®</sup>

- 6 – 12 panelists
  - comprehensive product training
  - Fixed Lexicon
  - Structured line scales
  - Many scale references
  - Scale calibration
  - Data analysis by ANOVA
- Individual panelists  
Panel as a whole

# Descriptive Analysis Ballot

## CANNED PASTA PRODUCT

Characteristic

208 (Test)

319 (Control)

### Basic Tastes

Sweet

Salt

Sour

Bitter

### Aromatics

Tomato Powder

Cooked Meat

Starchy

Metallic

### Aftertaste

Salt

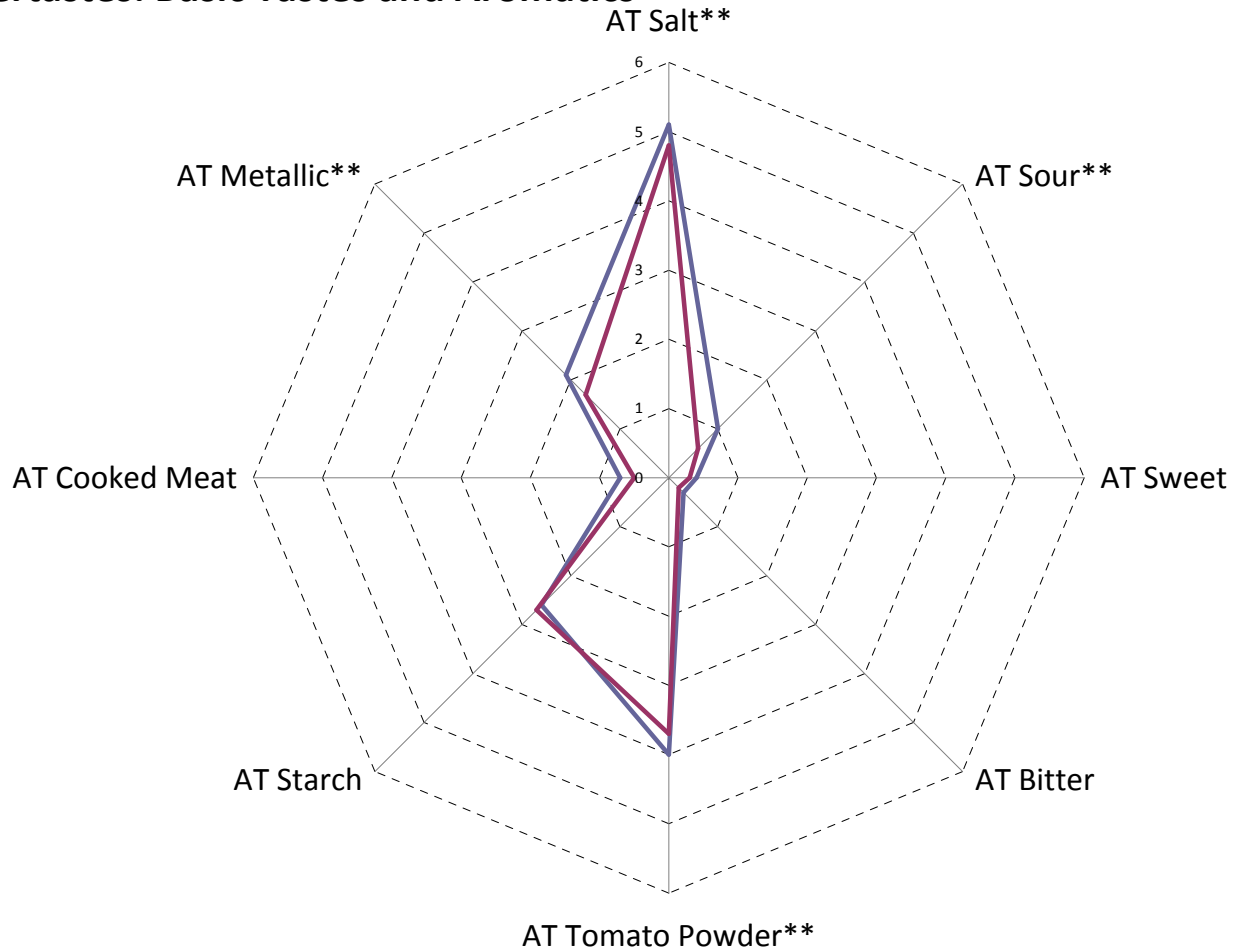
Sour

Tomato Powder

Metallic

# Possible Outcome and Output for Descriptive Data

## Aftertastes: Basic Tastes and Aromatics



# Limiting Factors for Sensory Testing

- **Stringent timelines**
- **Lack of financial resources dedicated to collection of sensory information**
- **Sample constraints**
  - ◆ Availability of samples—pilot plant or bench top
  - ◆ Amount of sample—is there enough to accommodate the chosen method?
  - ◆ Sample integrity—can product reach testing destination in a timely fashion in acceptable condition?
- **Importance of decision**
  - ◆ Are you prepared to act on your findings
  - ◆ What is the impact of the decision to the business

# Choosing the Proper Sensory Strategy for Your Product

- **What question or questions are you trying to answer about the product?**
  - ◆ How much is the product liked?
  - ◆ Which product is preferred?
  - ◆ How does my product compare to another?
  - ◆ Can a trained panelist tell the difference between my product and a national brand equivalent?
  - ◆ Where does my product fit into the marketplace?
- **What is your budget?**
  - ◆ Large scale consumer testing is costly
  - ◆ Methodology modifications can be carefully implemented depending on the weight/importance of the decision being made using the data collected
- **What is your timeline?**
  - ◆ More complex projects have longer lead times
- **Answers to the above questions will determine the appropriate sensory methodology for your project**

# Thank You!

- Any questions?

## Objective: Among QA professionals, which potato chip is most preferred and why?

### Testing parameters

- Users or “non-rejecters” of plain potato chips
- Demographic considerations – Multiple Regions
- N=~50
- Hedonics to determine “degree of liking/disliking”
- Diagnostics to assist in understanding why one is like/disliked more than the other
- Samples have been screened to ensure we get the information necessary to answer our objective

# Potato Chip Ballot

- **Hedonic Questions**

  - Overall**

  - Appearance**

  - Flavor**

  - Texture**

- **Diagnostics**

  - Potato Flavor**

  - Saltiness**

  - Crisp/Crunch**