

ASBC Annual Meeting

June 4–7 ■ Fort Myers, Florida

See what SCIENCE can brew for you

Application of Methods for Quality Control

Lindsay Barr, MS

ASBC Sensory Workshop 2017





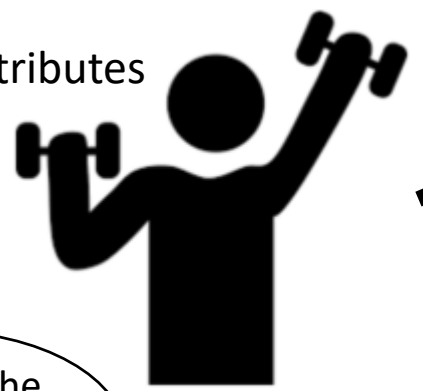




Raw
Materials

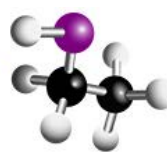
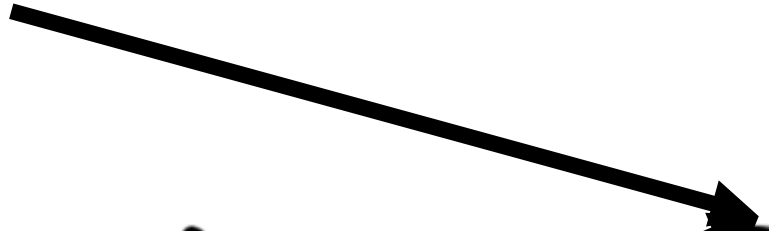


Attributes



A five on the
geraniol
scale.

Descriptive
Analysis



Raw Materials



Attributes



A five on the geraniol scale.

Descriptive Analysis



New Product Development

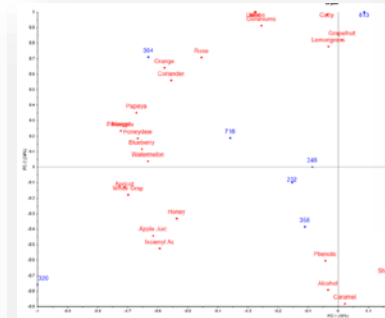
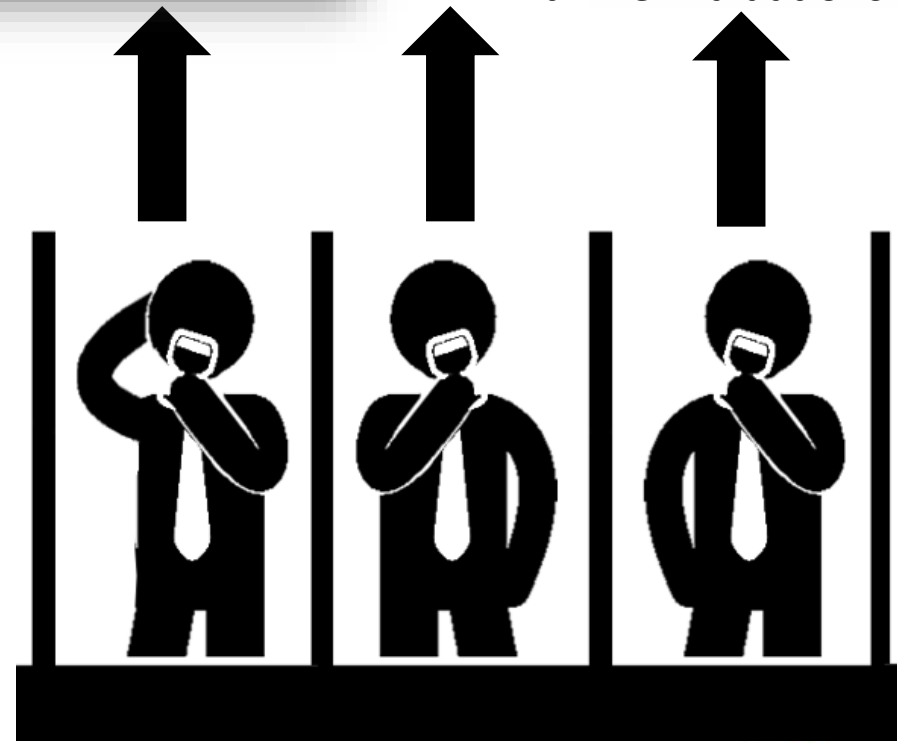


Figure 1. PCA with brands as scores and sensory hop aroma attributes as loadings

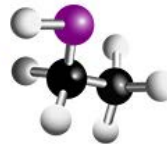
Raw Materials Evaluations



At-Line Evaluations



2017 ASBC Meeting



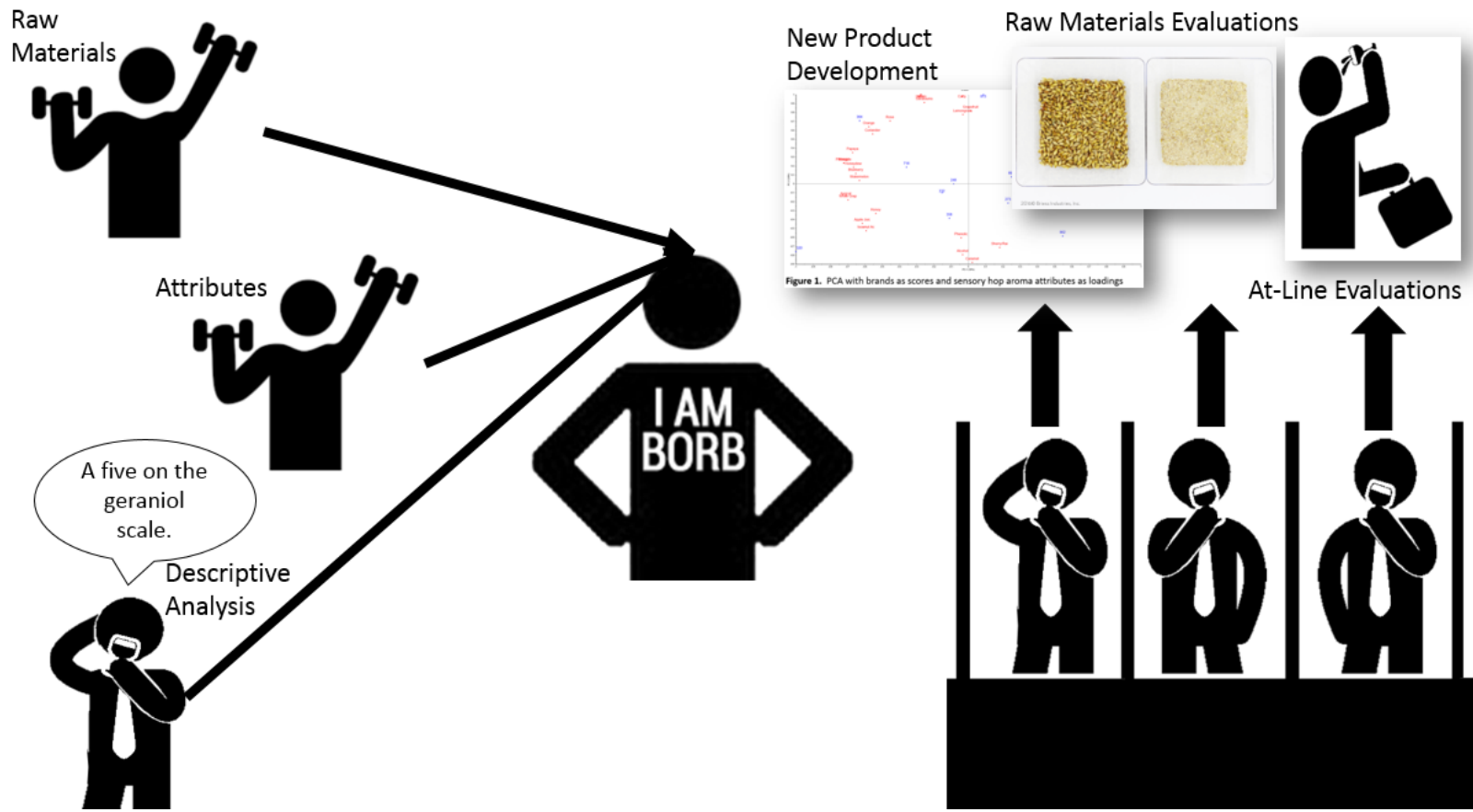
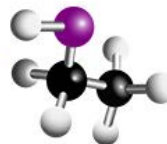


Figure 1. PCA with brands as scores and sensory hop aroma attributes as loadings

Quality Assurance







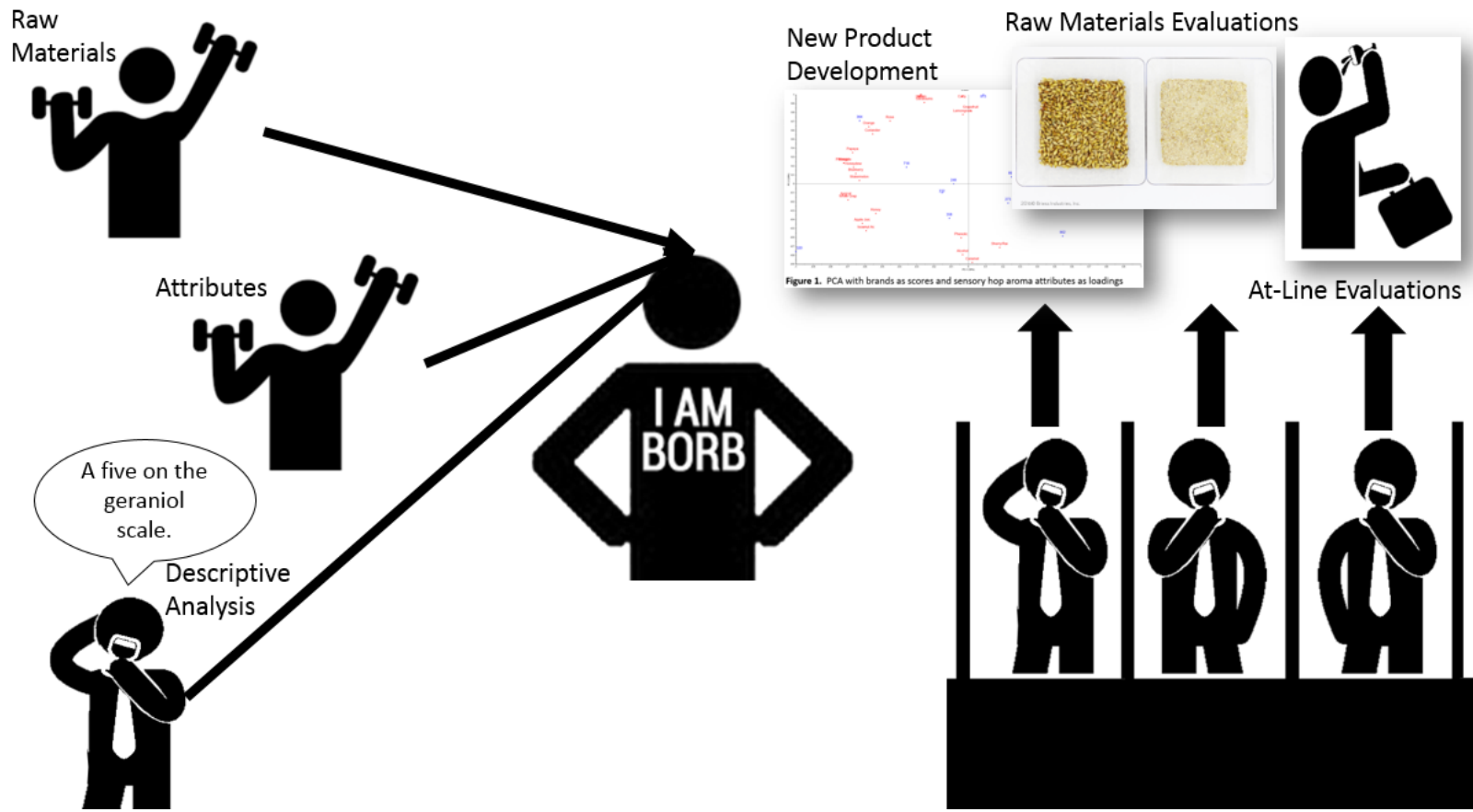


Figure 1. PCA with brands as scores and sensory hop aroma attributes as loadings

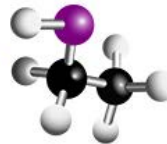
Quality Assurance

Methods

Data Analysis

Reaction Plans

Calibration & Validation



Quality Assurance Method Requirements

Established Specifications

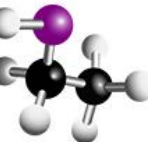
- Characteristics of the ideal or average product

Established Tolerance Limits

- Acceptable range of variation to deem a product “in spec”

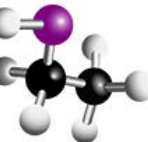
Appropriate Sampling Plan

- Reasonable number of samples, taken at appropriate process stages



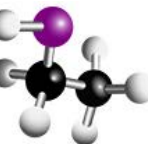
(Some) Sensory Methods

- In/Out
- Descriptive Analysis
- Quality Ratings
- Difference Testing



In/Out Method

- Purpose: Identify products that deviate from “normal” production.
- Method: Production samples are evaluated by a trained panel as being either “in”-spec or “out” of specification.
- Results: The percentage of panelists who deem the product “in” spec.

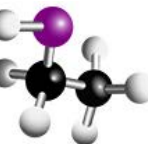


Consider this scenario...



In/Out Method

- ❌ Established Specifications
 - No rejection guidelines
- ❌ Established Tolerance Limits
 - No formal trainings
 - Biased panelists
- ❌ Appropriate Sampling Plan
 - Too many samples, high fatigue level



Consider this scenario...



Image shamelessly stolen from The Polish Academy of Sciences

In/Out Method

- ✓ Established Specifications
 - Defined flavor target
- ✓ Established Tolerance Limits
 - Trained panelists
 - Calibrated panelists
 - Unbiased panelists
- ✓ Appropriate Sampling Plan
 - Appropriate number of samples

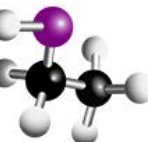


In/Out Method

- ✓ Established Specifications
 - Defined flavor target
- ✓ Established Tolerance Limits
 - Trained panelists
 - Calibrated panelists
 - Unbiased panelists
- ✓ Appropriate Sampling Plan
 - Appropriate number of samples



Yes but...

- **Why is the sample out of spec?**
- **To what degree?**



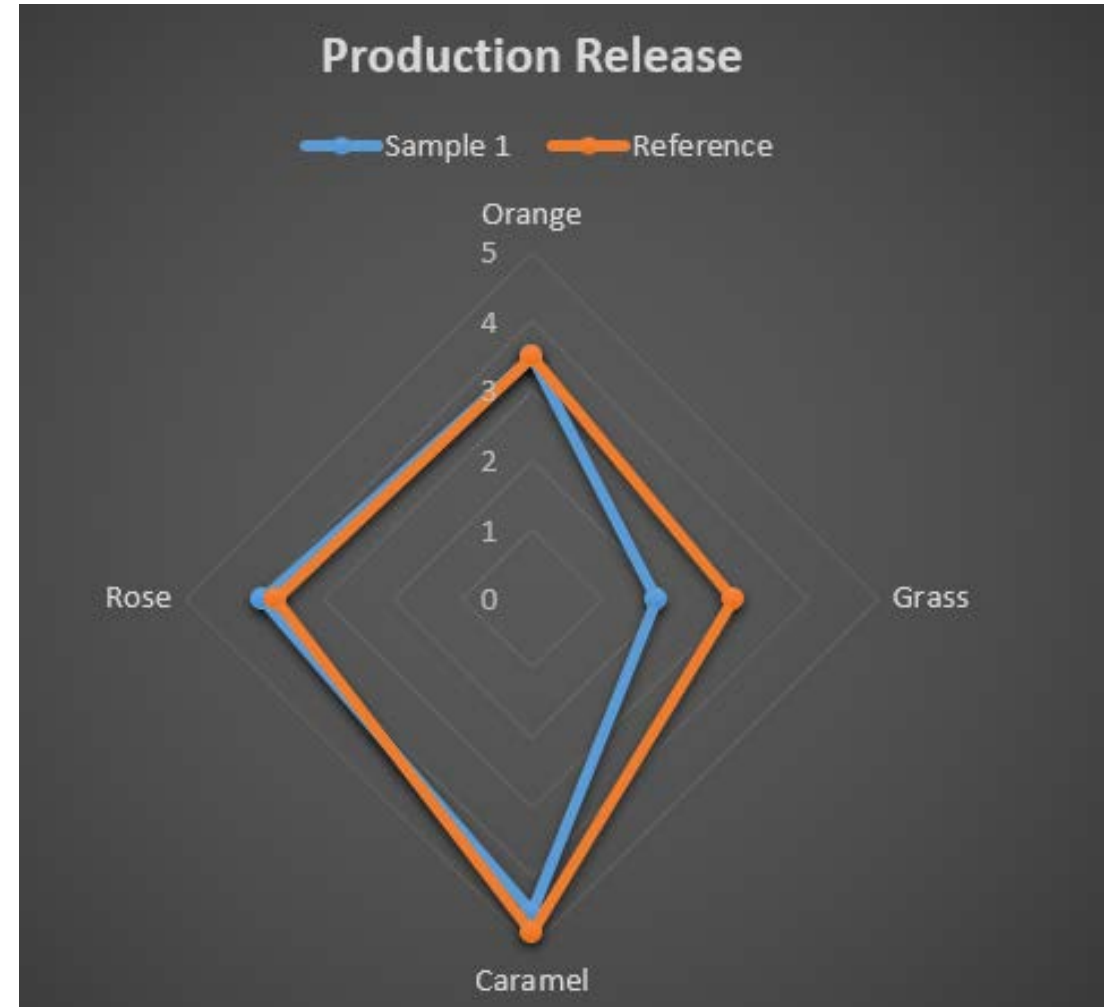
Descriptive Method

- Purpose: Identify products that vary in specific flavor attributes.
- How: Trained panelists evaluate the intensity levels for a small set of attributes for each product.
- Results: Individual intensity ratings for each attribute.

Panelist				
	4	2	3.5	5
	3	1.2	6	2.9

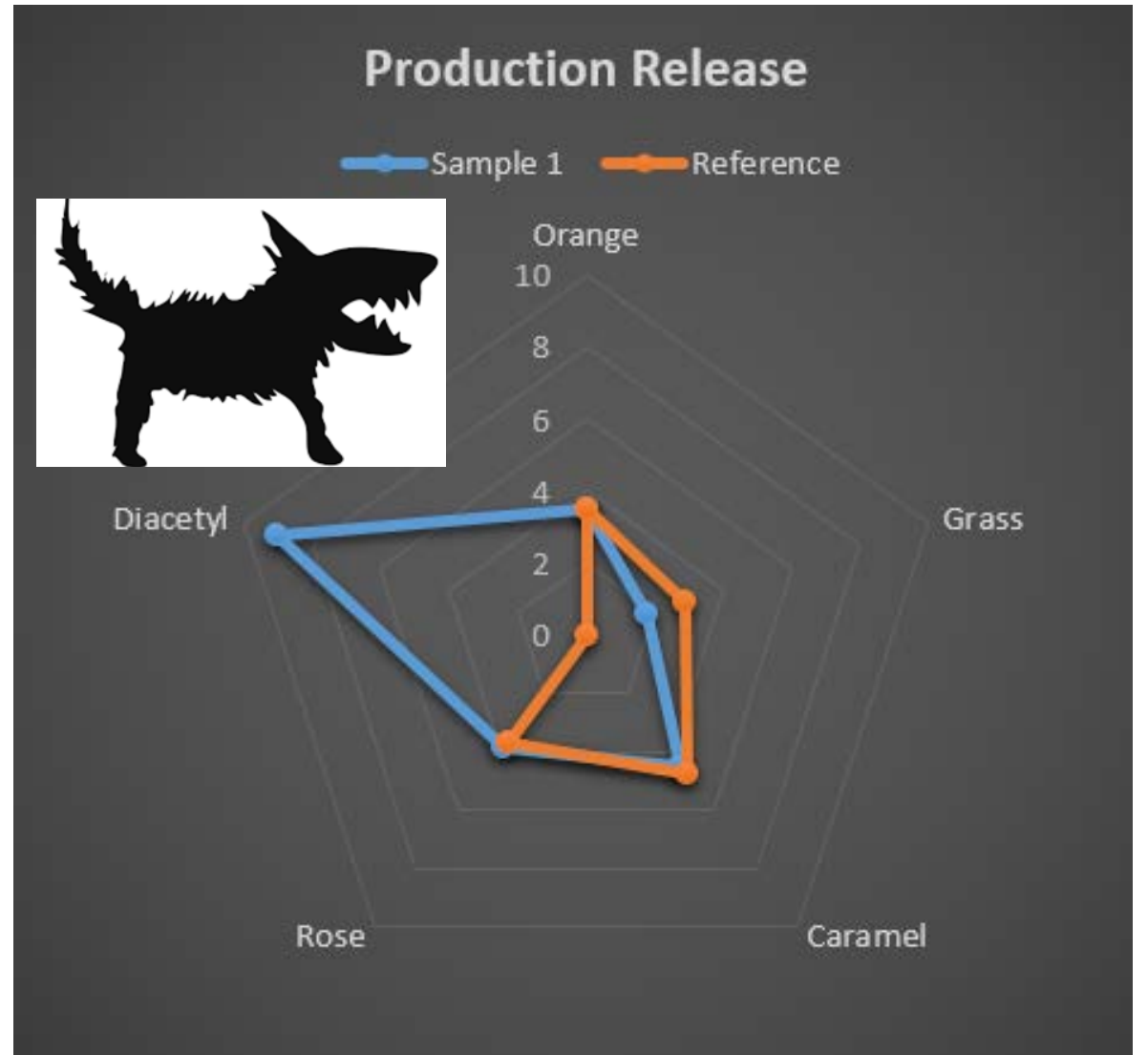


**Grass level is low...is that meaningful?
Should I react?**





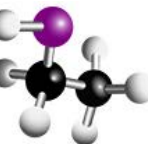
**HOLY?! Diacetyl
is high! Sound
the alarm!**



Descriptive Method

- ✓ Established Specifications
 - Defined flavor targets for multiple attributes
- ✓ Established Tolerance Limits
 - Set acceptable attribute variation ranges
 - Trained panelists
 - Calibrated and unbiased (?) panelists
- ✓ Appropriate Sampling Plan
 - Appropriate number of samples

This is pretty time consuming and the data has room for a lot of room for noise...



Quality Rating

- Purpose: Determine the overall level of quality for each product.
- How: Panelists are asked to use their own perception of quality or some set of established quality criteria to scale products on a scale indicating the overall level of quality.
- Results: Average quality score that is to indicate if the product is acceptable for release.



Very Poor

Poor

Fair

Good

Excellent



 Established Specifications

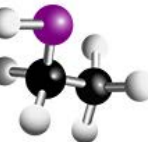
- The concept of quality is vague
- Assessing holistic concepts requires complex judgement processing
- The scale is subjective, thus largely opinion-based

 Established Tolerance Limits

- Most evaluators use the scale in a binary sense, like the in/out method

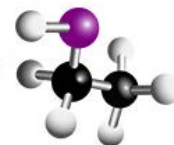
 Appropriate Sampling Plan

- ...sure





What about something more sensitive? Some test that can tell you if there is a difference between samples?!

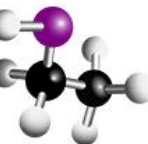


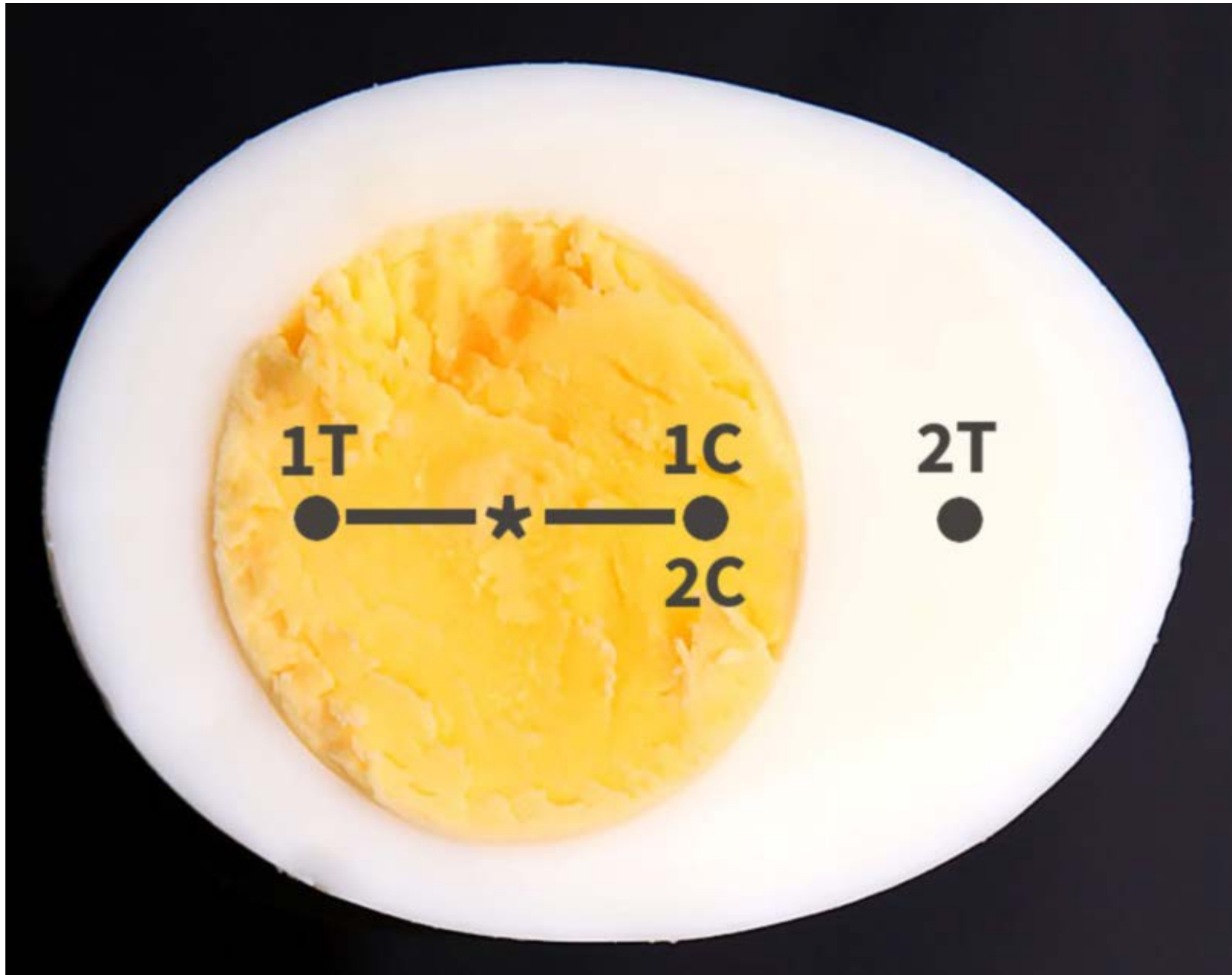
Difference Testing

- Purpose: Assess if the product is different than a “gold standard” control sample.
- How: Panelists identify and select which of three samples is different.
- Results: The number of panelists that recognized the odd sample. This number is used to assess if there is a statistically significant flavor difference between the samples.



“Which is different?”





Difference Testing

Established Specifications

- No established specific flavor specification
- Requires a static control sample
- Results do not indicate the source of the difference

Established Tolerance Limits

- Does not leave room for much product variation
- Does not recognize the range of acceptable variability

Appropriate Sampling Plan

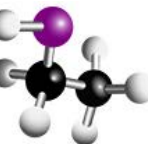
- Requires a large number of participants to achieve appropriate statistical power



Limitations of the Mighty Triangle Test

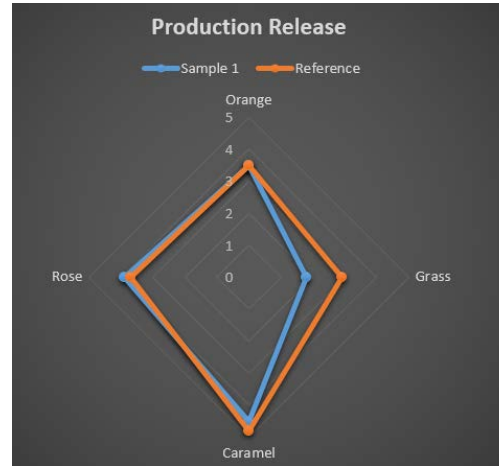


Lindsay Barr
May 31, 2017





Am I in spec?
 Yes
 No

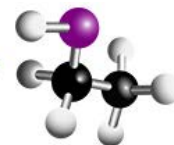


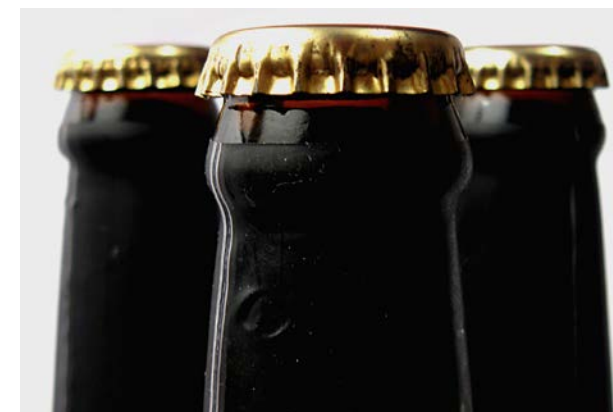
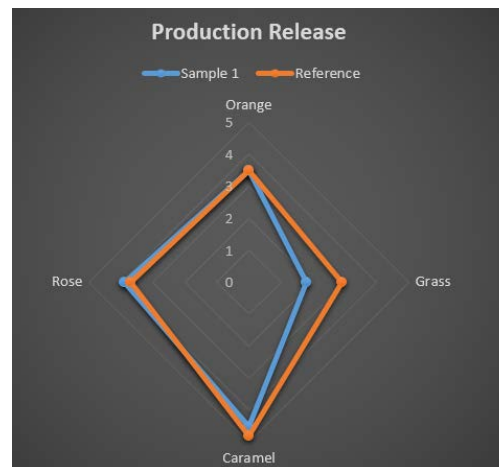
Trueness To Target Test

Reaction Plans

True to Target Test

- In/Out → TTT/Not TTT
 - Diminishes pressure
 - Focuses attention of the objective evaluation
- Descriptive Analysis → Comments and/or CATA
 - Open text or CATA to understand the nature of derivations
 - Enables the panel leader to take appropriate actions based on comments
- Difference Testing → Control Charting
 - Indicates where each samples lies in the context of every “normal” sample produced...more on this later ;-)





Trueness To Target Test

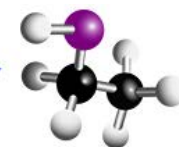
Target Generation

Sample Evaluation

Data Analysis

Reaction Plans

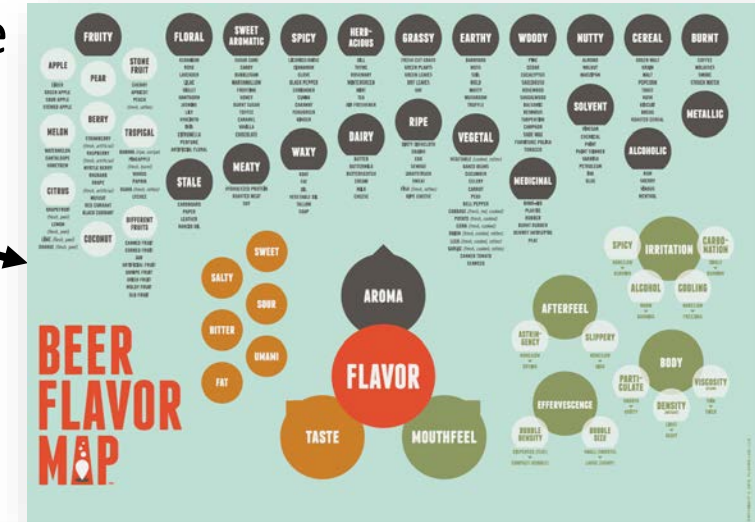
2017 ASBC Meeting



Step 1: Target Generation

✓ Established Specifications → Target Generation

- Evaluators individually describe the beer's flavor profile
 - Use common language and structure
 - Aggregate to find common attributes
- Build and/or adjust your target



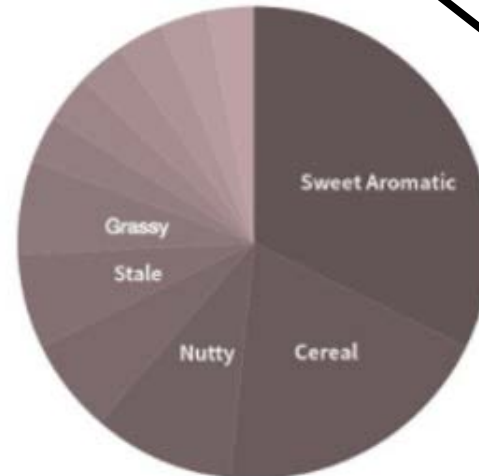
The flavor description is the current baseline target from which all subsequent batches will be compared against.

VISUAL: A translucent ruby brown color with thin light brown foam.

AROMA: Toffee, vanilla and molasses with hints of toast. Some malty and nutty aromas.

TASTE: Mildly sweet with some bitterness.

MOUTHFEEL: Light body, some tingling carbonation, and slightly mouthwatering.



SWEET AROMATIC



CEREAL



Step 2: Sample Evaluation

- How closely does each batch adhere to the set target?
 - Panelists evaluate a representative sample of every batch against the set brand profile.
 - Each modality is evaluated separately and deemed either “true to target” or not “true to target”.



AROMA TRUE TO TARGET

Based on the aroma description below, is this sample reasonably true to target?

AROMA: Toffee, vanilla and molasses with hints of toast. Some malty and nutty aromas.

True to target?

YES

NO

Comments (required if no):

Diacetyl

Step 3: Data Analysis

✓ Established Tolerance

- Targets allow for normal variation
- Acceptable range of flavor

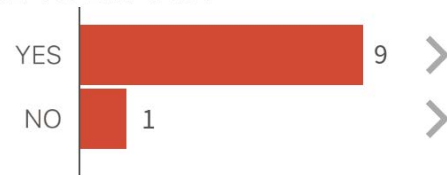
How is the acceptable range of variation determined by the process? Why is 20% not TTT but 10% is?! What do you call an crocodile in a vest?

process itself

Percent Defect

Percent defect refers to the amount of tasters that rated this sample as not true to brand. 10% is inside your control limits.

TRUE TO BRAND?



NOT TRUE TO BRAND

Modality	% Not TTT	TTT or Not TTT?
Visual	10%	TTT
Aroma	0%	TTT
Taste	10%	TTT
Mouthfeel	20%	Not TTT
Overall	0%	TTT

JR Jim Rossette



Lighter and hazier with more pronounced citrus



P-Charting!

- Completed panel data: represented by dots, each dot represents the average number of panelists who selected “Not TTB” for a specific modality.
- Center Line (CL): Average (mean) number of panelists that typically select “Not TTB.”

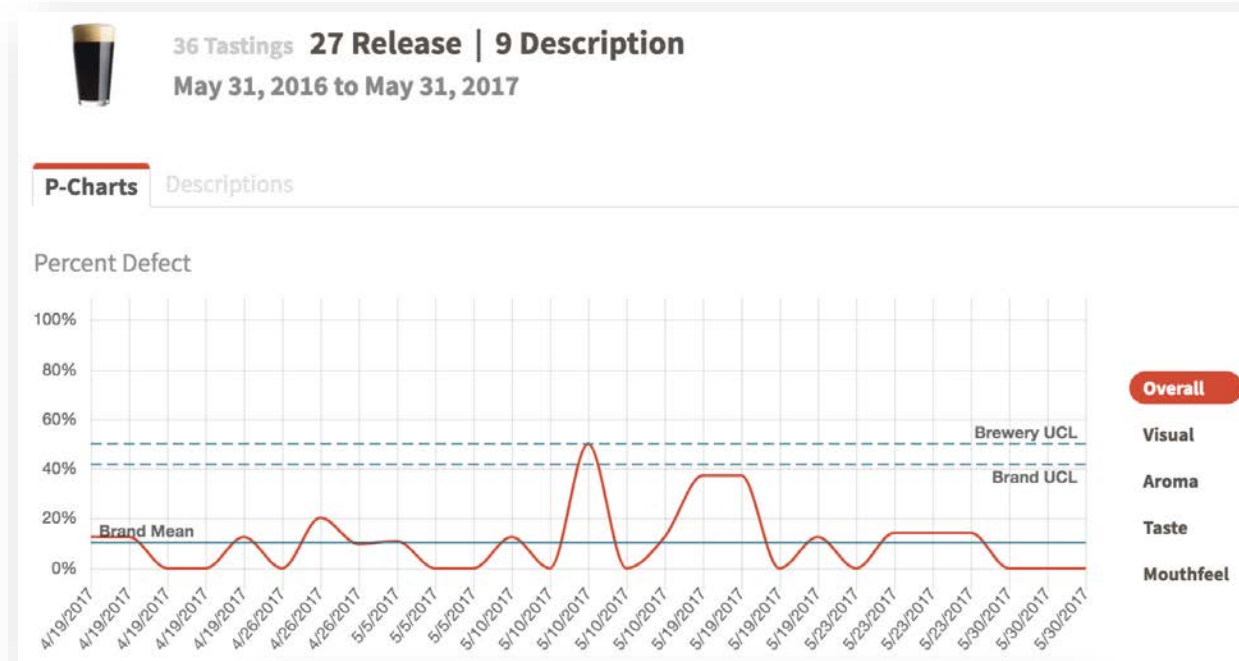
$$\bar{x} = \frac{x_1 + \dots + x_n}{n}$$

An Investigator!

- Upper Control Limit (UCL): Representing the highest level of “Not TTB” evaluations for a product Lower Control Limit (LCL): the lowest level of “Not TTB” evaluations from line, and it is always zero. It doesn't make sense as you could never have a negative number of assessments.

$$UCL = \bar{x} + 3 * \sigma$$

✓ Established Tolerance Limits



Batch-to-Batch Variation in Brewing: Let P-Charts Do the Work

January 15, 2017

Step 4: Data Reaction

Track

- Panelist Comments
- Level of sensory fail
- Brand and package code
- Location
- Stage of the Process
- Root Cause
- Corrective Action
- Disposition and Final Outcome

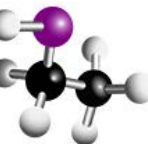
Date	BN	Package	Panelist	Visual	Aroma	Taste	MF/B	Overall	Comments
4/20/2017	170419091	can	HUG1						sl dull
			RAD1		Not TTT			Not TTT	H2S
			DAR1		Not TTT			Not TTT	H2S
			BAR1		Not TTT			Not TTT	H2S
			MIT1		Not TTT			Not TTT	H2S
			CON1		Not TTT			Not TTT	H2S
			CHR1		Not TTT			Not TTT	H2S

QA Anomaly Report Time and Date * 4/20/17 2 PM 00

Anomaly

Brand

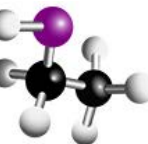
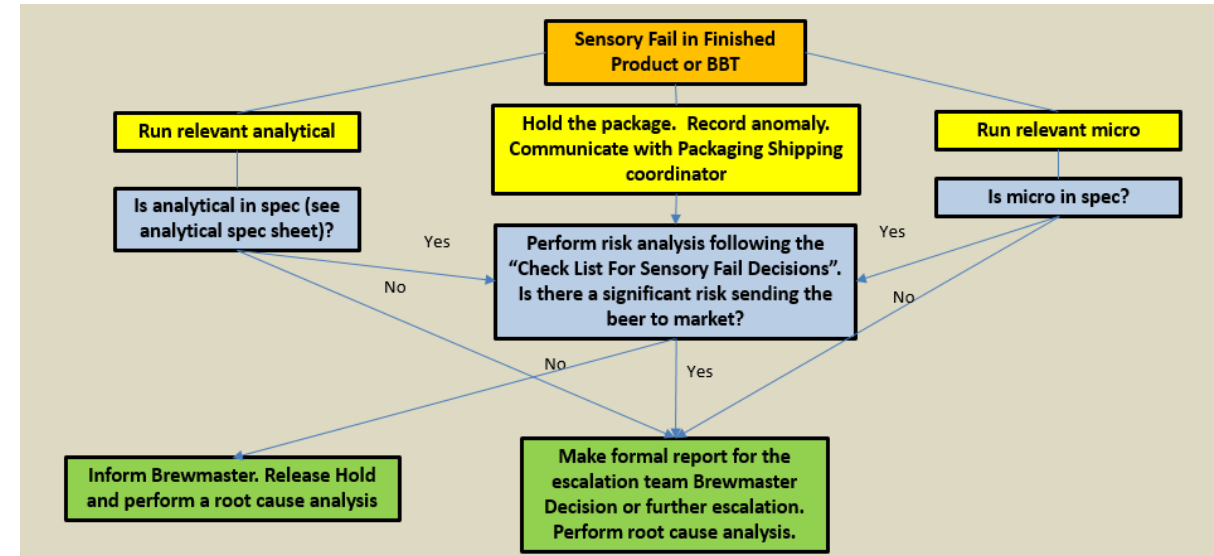
NBB Site *

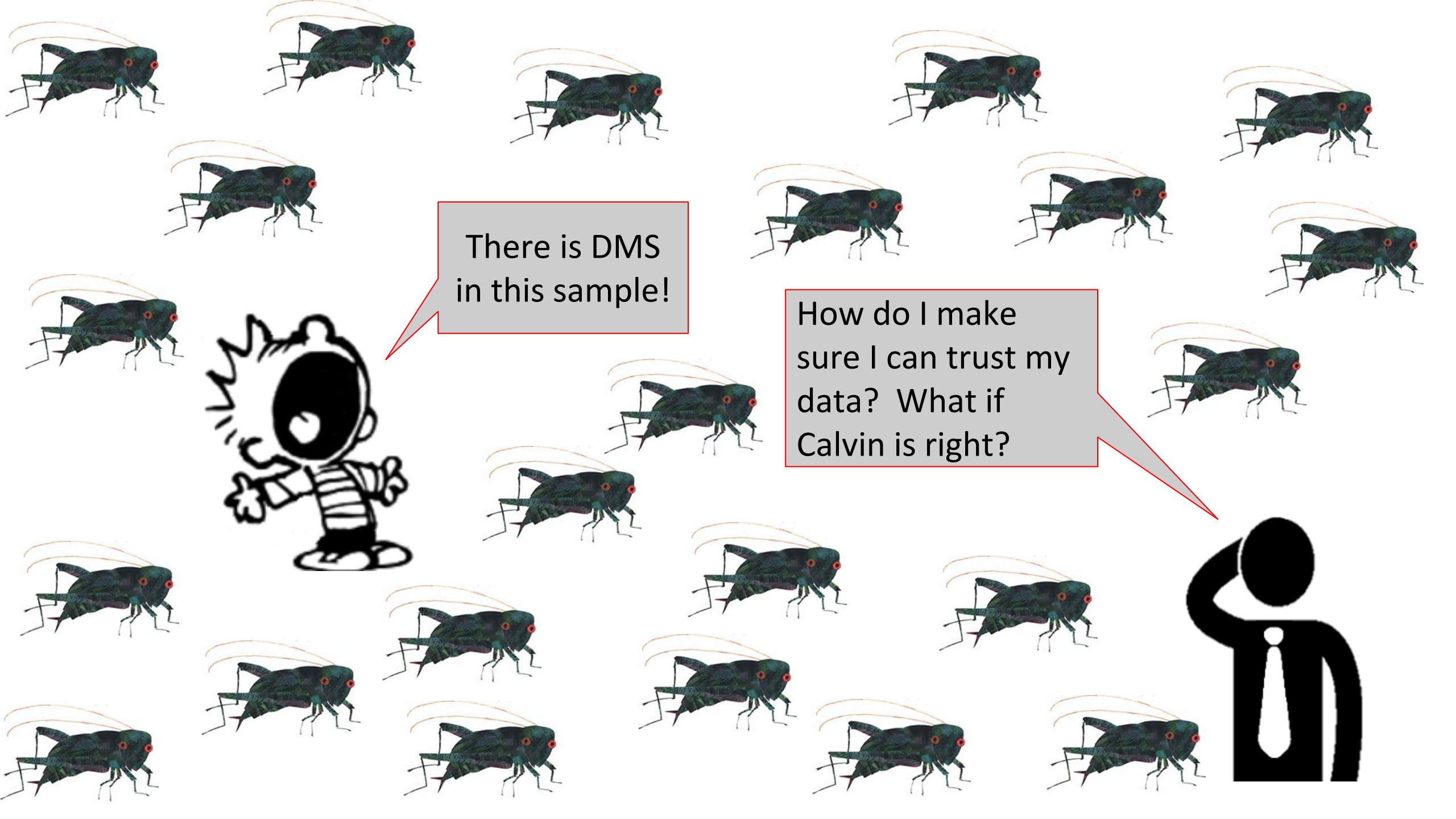


Decision Making Tools

Brewmaster Ready Check List

- Relevant Analytical Information
 - Are any analytical parameters out of spec? If yes, what is the risk?
- Relevant Micro Results
 - Is micro clean? If no, what is the risk?
- What is the potential shelf life impact?
- Are there any salvaging solutions? If so, what?
- Is there a root cause and has it been addressed?
- Have there been past instances where we were in a similar situation? What was done and what was the impact? Consult the anomaly tracker.
- What is the overall quality risk if the beer were to release? (Shelf life decrease, turbidity, aroma, sour, etc..)
- Could we expect consumer complaints associated with this beer release?





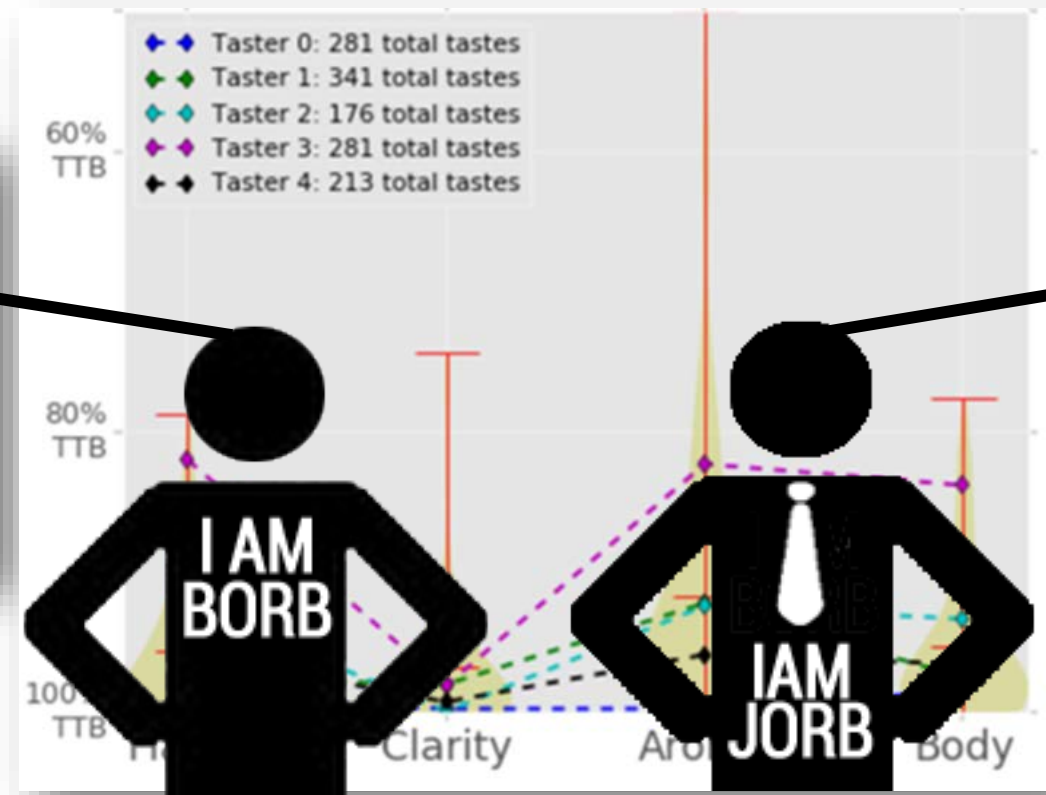
There is DMS
in this sample!

How do I make
sure I can trust my
data? What if
Calvin is right?

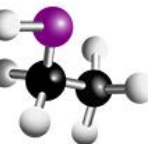


Panelist Calibration and Validation

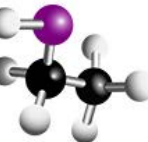
BOU1	Biased Taster
BORB	Biased Taster
IRB1	Biased Taster
MAD1	Biased Taster
BAR1	Biased Taster
GUE2	Biased Taster
VAN1	Biased Taster



SED1	Cautious Taster
PEO1	Cautious Taster
BAI1	Cautious Taster
NAB1	Cautious Taster
JORB	Cautious Taster
STE1	Cautious Taster
RAD1	Cautious Taster
FIS1	Cautious Taster
MCC1	Cautious Taster
COG1	Cautious Taster
DIC2	Cautious Taster
VIL1	Cautious Taster
KAR1	Cautious Taster



**How do I train my
panelists to identify out
of spec beer repeatedly
and without bias?**



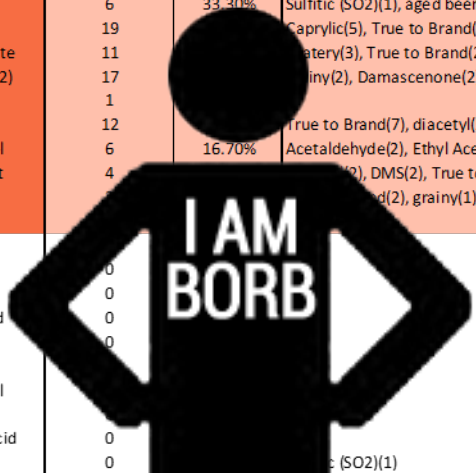
Borb

6x Attributes	Times Seen	% Correct	Often Confused With:
Metallic	13	100.00%	Sulfitic (SO2)(1)
Papery	9	100.00%	
Acetaldehyde	17	94.10%	3-Hexanol(2), Ethyl Acetate(1), Myrcene(1)
ethyl butyrate	11	90.90%	butyric(8), Linalool(1)
Phenolic (4VG)	11	90.90%	aged beer(1)
sulfidic (H2S)	11	90.90%	Mercaptan(1), Lightstruck(1)
Chlorophenol	9	88.90%	4-Ethyl Phenol(2)
geosmin	9	88.90%	Leathery(2), 3-Hexanol(1)
Acetic Acid	7	100.00%	diacetyl(1)
Ethyl Hexanoate	7	100.00%	
Geraniol	7	100.00%	
aged beer	7	85.70%	grainy(2), Methional(2), 3-Hexanol(1)
Lightstruck	7	85.70%	sulfidic (H2S)(1)
Linalool	7	85.70%	ethyl butyrate(1)
isoamyl acetate	13	76.90%	Ethyl Acetate(2), True to Brand(1)
Watery	13	76.90%	True to Brand(4), Ethyl Acetate(3), diacetyl(2)
4-Ethyl Phenol	8	75.00%	Chlorophenol(2), True to Brand(1)
Catty	8	75.00%	3-Hexanol(1), Sulfitic (SO2)(1)
Sweet	15	73.30%	Tru
Isovaleric	10	70.00%	but
Leathery	10	70.00%	geo
Sour	12	58.30%	Tru
True to Brand	18	55.60%	Cap
butyric	14	50.00%	eth
Mercaptan	14	50.00%	Tru
Damascenone	10	50.00%	Sulf
Clarity	4	75.00%	diac
Different Beer	4	75.00%	Wa
Methional	6	50.00%	Sulf
DMS	13	46.20%	Tru
grainy	4	50.00%	Sulfitic (SO2)(2), aged beer(2), styrene(1)
malty-biscuity	5	40.00%	True to Brand(1), Sulfitic (SO2)(1), Caprylic(1)
Myrcene	5	40.00%	3-Hexanol(1), Acetaldehyde(1), Sour(1)
Onion	6	33.30%	Sulfitic (SO2)(1), aged beer(1), diacetyl(1)
diacetyl	19		Caprylic(5), True to Brand(4), Watery(2)
Ethyl Acetate	11		Watery(3), True to Brand(2), 3-Hexanol(2)
Sulfitic (SO2)	17		grainy(2), Damascenone(2), Methional(2)
Bitter	1		
Caprylic	12		True to Brand(7), diacetyl(5), Watery(1)
3-Hexanol	6	16.70%	Acetaldehyde(2), Ethyl Acetate(2), aged beer(1)
Astringent	4		(2), DMS(2), True to Brand(1)
styrene			(2), grainy(1)
Smokey			
honey	0		
Indole	0		
Kerosene	0		
malic acid	0		
mousey	0		
musty	0		
Rancid Oil	0		
skatole	0		
succinic acid	0		
Worty	0		(SO2)(1)

Attribute Training?

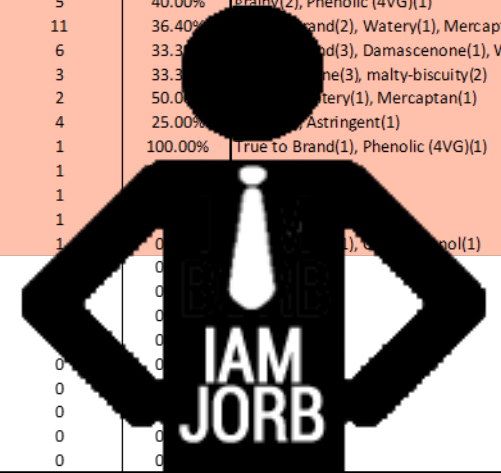
Both Borb and Jorb have high attribute on but one us in panel the other is

Attribute trainings do not
 a valid panelist make!!
 What else is needed?



Jorb

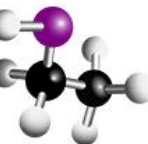
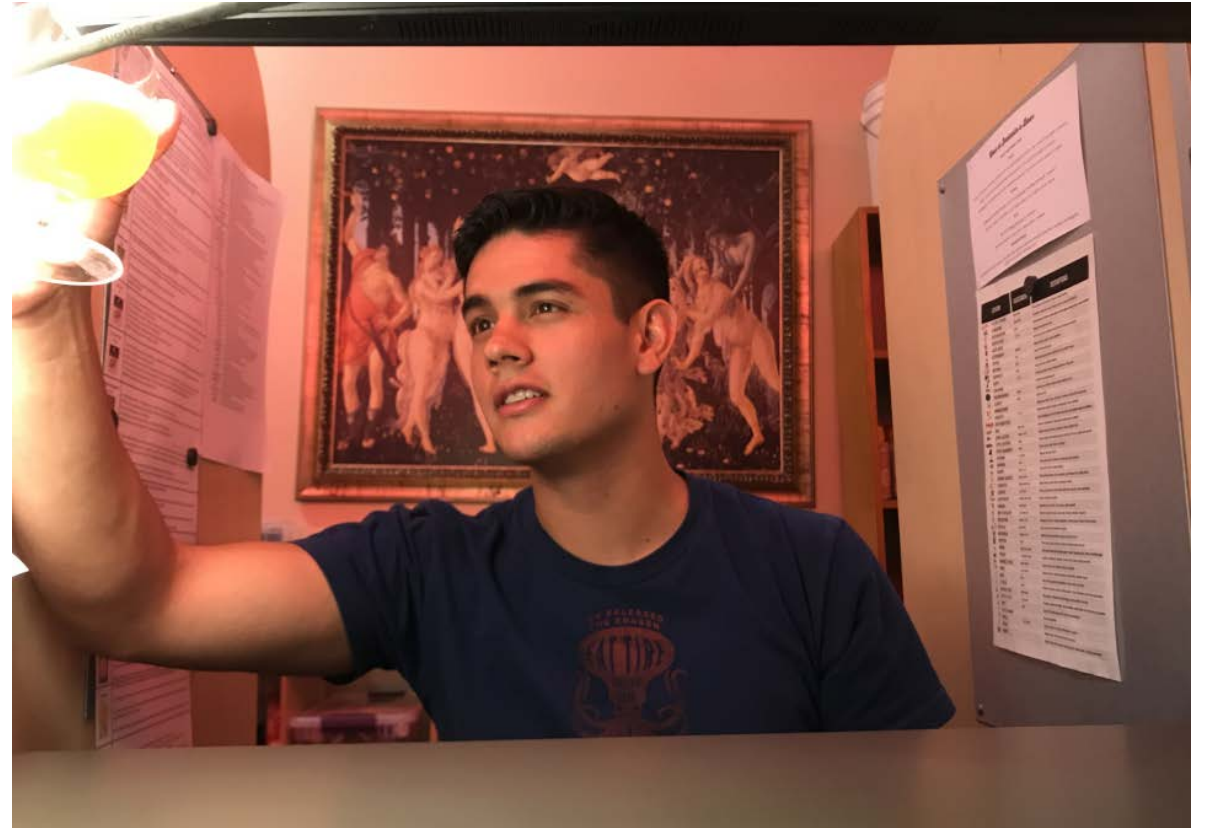
Acetaldehyde	15	100.00%	
Metallic	12	100.00%	
Sulfitic (SO2)	10	100.00%	DMS(1)
Caprylic	8	100.00%	True to Brand(1), Watery(1)
Chlorophenol	9	88.90%	Watery(1), Different Beer(1), styrene(1)
isoamyl acetate	9	88.90%	DMS(1), Ethyl Acetate(1)
diacetyl	14	85.70%	Watery(1), Clarity(1), Sour(1)
Papery	7	100.00%	True to Brand(1)
butyric	14	78.60%	Isovaleric(4), Mercaptan(2), ethyl butyrate(1)
Phenolic (4VG)	9	77.80%	malty-biscuity(1), Bitter(1), Smokey(1)
Ethyl Acetate	6	100.00%	isoamyl acetate(1)
ethyl butyrate	6	100.00%	Isovaleric(1), butyric(1)
Methional	6	100.00%	
sulfidic (H2S)	6	100.00%	
Sour	7	71.40%	True to Brand(2), Sweet(1), diacetyl(1)
Acetic Acid	6	83.30%	True to Brand(1)
Watery	14	64.30%	Sweet(2), True to Brand(2), Astringent(1)
4-Ethyl Phenol	5	100.00%	
Sweet	10	60.00%	Watery(2), True to Brand(1), Phenolic (4VG)(1)
Catty	5	80.00%	True to Brand(1)
Leathery	5	80.00%	
True to Brand	12	58.30%	Astringent(3), Watery(2), Sour(2)
geosmin	7	57.10%	True to Brand(2), Watery(1)
Isovaleric	7	57.10%	butyric(4), ethyl butyrate(1)
Mercaptan	7	57.10%	butyric(2), DMS(1), Onion(1)
3-Hexanol	4	100.00%	
aged beer	4	100.00%	
Geraniol	4	100.00%	Linalool(4), Watery(1)
Lightstruck	4	100.00%	
Linalool	8	50.00%	Geraniol(4)
Clarity	4	75.00%	diacetyl(1)
Different Beer	4	75.00%	Chlorophenol(1)
Myrcene	4	75.00%	Sour(1)
malty-biscuity	5	40.00%	grainy(2), Phenolic (4VG)(1)
DMS	11	36.40%	True to Brand(2), Watery(1), Mercaptan(1)
Astringent	6	33.30%	DMS(3), Damascenone(1), Watery(1)
grainy	3	33.30%	styrene(3), malty-biscuity(2)
Onion	2	50.00%	Watery(1), Mercaptan(1)
Damascenone	4	25.00%	Astringent(1)
Bitter	1	100.00%	True to Brand(1), Phenolic (4VG)(1)
Ethyl Hexanoate	1		
Worty	1		
Smokey	1		
styrene	1	0.00%	(1), Caprylic(1)
honey	0		
Indole	0		
Kerosene	0		
malic acid	0		
mousey	0		
musty	0		
Rancid Oil	0		
skatole	0		
succinic acid	0		



TTT: Brand Familiarization

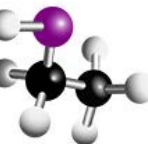
- Trueness to brand trainings
- Calibrating before booths
- Negative controls in booths
- Positive controls in booths

“The general who wins the battle makes many calculations in his temple before the battle is fought. The general who loses makes but few calculations beforehand.” --Sun Tzu



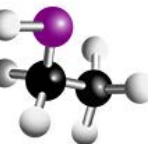
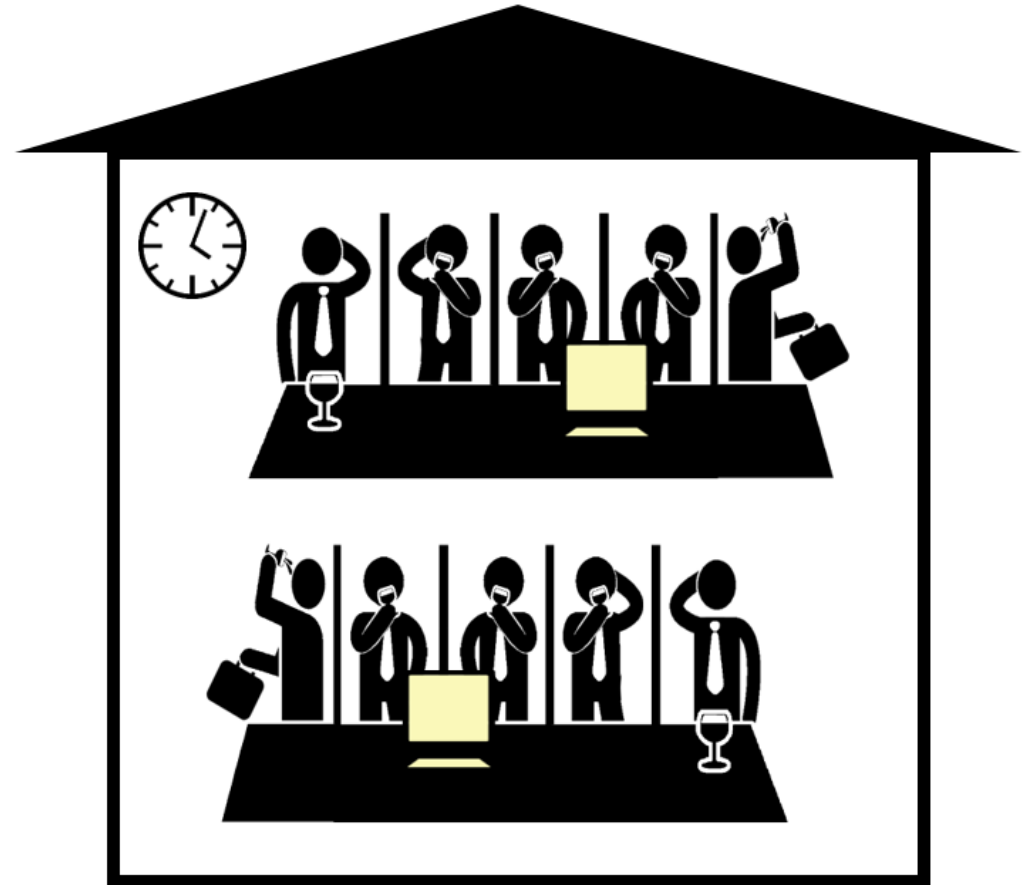
Panelist Selection

- A good panelist is...
 - Sensitive
 - Consistent
 - Aligned with the panel
 - Motivated
 - Articulate
- Make sure you keep them around
 - Ask for no more than 2.5% of their time
 - Communicate frequently
 - What motivates a panelist? Ask them!



Where and When

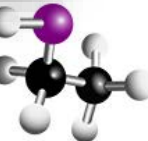
- Consistency is key
 - Same Place
 - Same Time
 - Same Frequency
- Free of distractions
 - Aromas
 - Noise
 - Panelists 😊
- Beware of fatigue
 - Keep sample maximum to 8/panel
 - Beware of sample order and fatigue level



Documentation

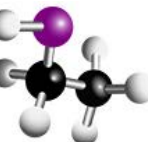
- **Cover your a**!**
 - Within the brewery
 - Within the program
- **Consistency**
 - Panelist Bias
 - Leadership within the brewery

Appendix	
BASIC	..#
PANELIST ONBOARDING	
NEW PANELIST TRAINING	
TRACKING ATTENDANCE AND PROGRESS OF NEW PANELISTS	
PANEL	#
DATA ORGANIZATION AND REPORTING	
MONTHLY DATA RECAP REPORTING	
FINISHED PRODUCT BATCH RELEASE SHEET	
TASTE RELEASES	#
SETTING UP AT-LINE TASTINGS	
COMMUNICATING AT-LINE RESULTS	
MAINTENANCE OF TRACKING SYSTEMS FOR AT-LINE RELEASES	
MAINTENANCE	#
PANELIST REPORTING AND TRIMESERLY TRACKING	
LAB CLEANING CHECKLIST	
NEW BRANDS, SHELF LIFE, TESTS	#
NEW BRANDS RELEASE CHECKLIST	
SET UP AND EXECUTION OF THE DoD	
SCHEDULING BRANDS FOR SHELF LIFE ANALYSIS	
SHELF LIFE REPORTING	



Growing the Program

- Focus on QC first, this is your greatest risk. Training always fuels the program, keep focusing there.
- Continue to introduce new products
- Add evaluation points (MV, raw materials)
- Introduce new attributes
- Develop new methods like Descriptive Analysis and Difference Testing
- Keep 'em motivated!



Hey, thanks!

