Stat-Ease Online DOE Summit

2023

Application of DOE in the Development of an Alcoholic Beverage

(Sensory Attributes and Preference as Primary Measures)

Douglas Scott

Optimal Formulations, LLC

Self Introduction

- BS in Pharmacy - University of Cincinnati 1986
- PhD in Pharmaceutics – University of Maryland at Baltimore 1991
- Procter & Gamble (26 years)
  - Rx 1991 – 1995 **First introduced to Stat-Ease [Design-Ease and Design-Expert]**
  - OTC Medicines 1995 - 1998
  - Oral Care Technology Division 1998 - 2017
- AS in Brewing Science - Midwest Culinary Institute – Cincinnati State 2018 - 2020
- Consultant 2018 – Present
  - Primary Client is MadTree Brewing - Cincinnati

Challenges with Sensory Studies

- Data is inherently variable.
  - Using humans as an instrument to grade on a scale.
  - Hedonic measures (Preference/Liking) are the ultimate subjective measure.
- Need to limit study legs to avoid grading fatigue in tasting sessions.
  - May need to sacrifice replicates or lack of fit points.
- Sometimes have to accept less than desirable data.
  - There is truth in the data – needs to be sorted out without fooling oneself.
- Need to assess the Risk:Benefit ratio of using lean designs, and “messy” data sets.
  - What’s the worse thing that could happen?
    - A very good product is identified that may not be the absolute optimum.

Context of Sensory Work with The Brewery

- Predominately New Product Development
- Two Stages of Development
  - Exploratory Stage
    - Serial DOE with employee sensory panels.
    - Need good direction setting with reasonable prediction.
    - Can be more “*liberal*” with statistics and models.
    - Want good understanding of the flavor system.
      - **WHY** is an optimum an optimum?
  - Final Selection Stage:
    - 3-4 prototypes tested head-to-head.
    - Includes a broader employee pool and actual taproom patrons.

MADTREE

Brewery

Cincinnati, Ohio

- Have performed over 100 individual DOE's in 5 years:
  - Sensory optimization for Beers, Ciders, Mixed Drinks and RTD Cocktails.
  - Other non-sensory work.

Ready To Drink (RTD)
Vodka Cocktails

- RTD's are simple formulations and easy to prepare.
- Lend themselves nicely to DOE.

Design

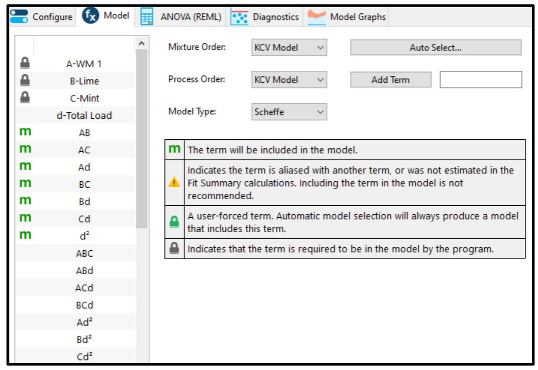
Component	Name	Units	Type	Minimum	Maximum
A	Watermelon	Fraction	Mixture	0.4	0.7
B	Lime	Fraction	Mixture	0.2	0.4
C	Mint	Fraction	Mixture	0.1	0.4

Factor	Name	Units	Type	4 Levels
D	Total Load	uL/100ml	Discrete Numeric	100, 267, 433, 600

- Second study in serial DOE.
- KCV (Kowalski, Cornell, and Vining) Split-Plot design.
  - Each tasting session had a fixed Total Flavor Load.
- All subjects intended to assess all products (employee sensory panel).

Design Layout and KCV Model Terms

Group	Run	Component 1 A:WM 1	Component 2 B:Lime	Component 3 C:Mint	Factor 4 d:Total Load uL/100ml
1	1	0.4	0.2	0.4	100
1	2	0.56	0.32	0.12	100
1	3	0.7	0.2	0.1	100
1	4	0.4	0.4	0.2	100
1	5	0.5	0.28	0.22	100
2	6	0.4	0.4	0.2	600
2	7	0.7	0.2	0.1	600
2	8	0.57	0.33	0.1	600
2	9	0.4	0.2	0.4	600
2	10	0.51	0.27	0.22	600
3	11	0.55	0.2	0.25	267
3	12	0.55	0.2	0.25	267
3	13	0.4	0.31	0.29	267
3	14	0.61	0.29	0.1	267
3	15	0.5	0.4	0.1	267
4	16	0.4	0.31	0.29	433
4	17	0.5	0.4	0.1	433
4	18	0.56	0.2	0.24	433
4	19	0.54	0.28	0.18	433
4	20	0.4	0.31	0.29	433



- Total Load was fixed in each testing session to avoid sequence effects on the subjects' ability to grade (going from High Load to Low Load).

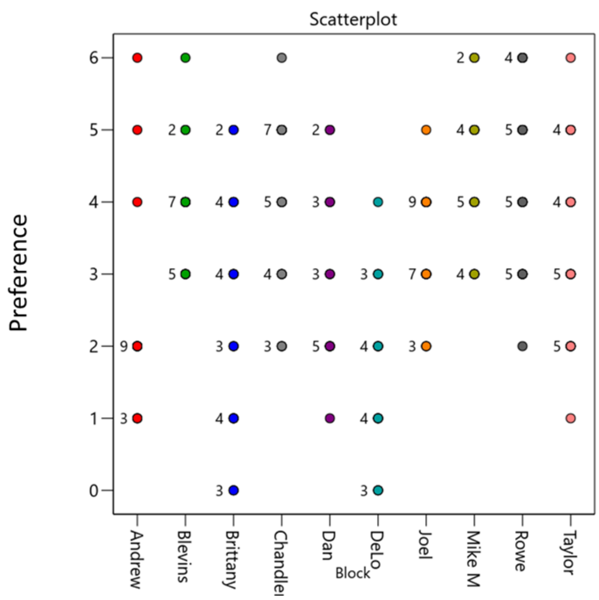
- KCV designs are ideal for these type studies:
  - ➢ Get information on linear effects and interdependencies of all factors.

Responses

- 15 responses graded on a 0 - 7 point scale.
- Simultaneous Hedonic and Attribute grading.

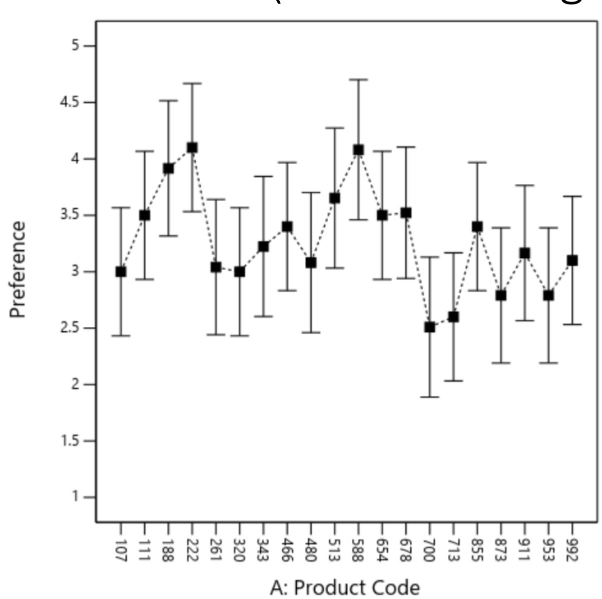
Hedonic Measure	Aroma Attributes	Flavor/Taste Attributes
Preference	Aromatic Intensity Watermelon Aroma Mint Aroma Lime Aroma Alcohol Aroma Sweet Aroma Artificial Aroma	Flavor Intensity Watermelon Taste Mint Taste Lime Taste Alcohol Taste Artificial Taste Tartness

Individual Subject Grading



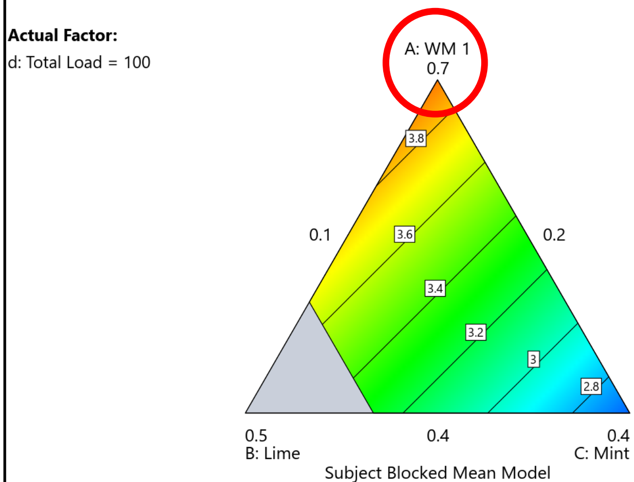
- Blocking for subject helps to account for differences in use of the grading scale.

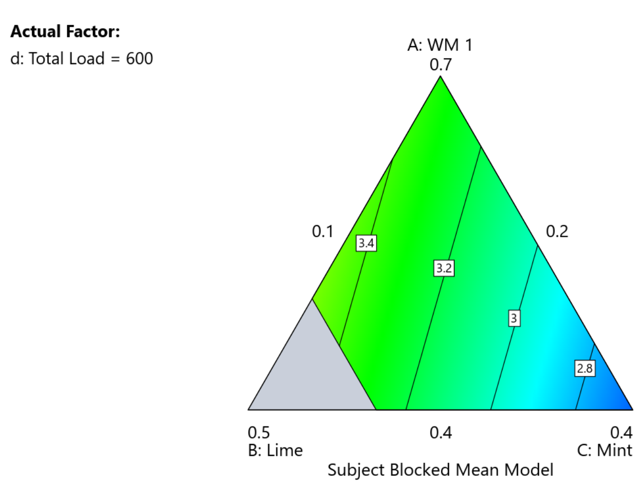
Preference (Runs as a Single Categoric Variable)



- First look to see spread amongst the individual products.
  - Mean Preference range = 2.5 – 4.1.
- Mean values are used in Mean DOE modeling.

Preference (Mean Model)

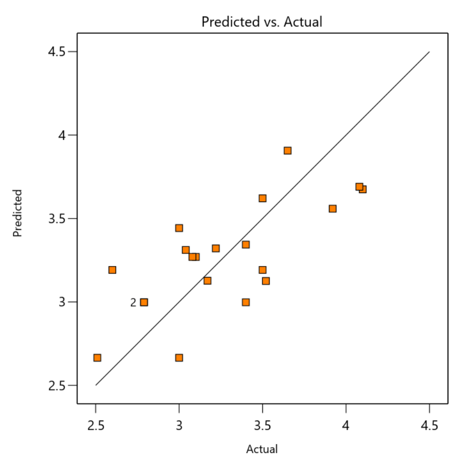




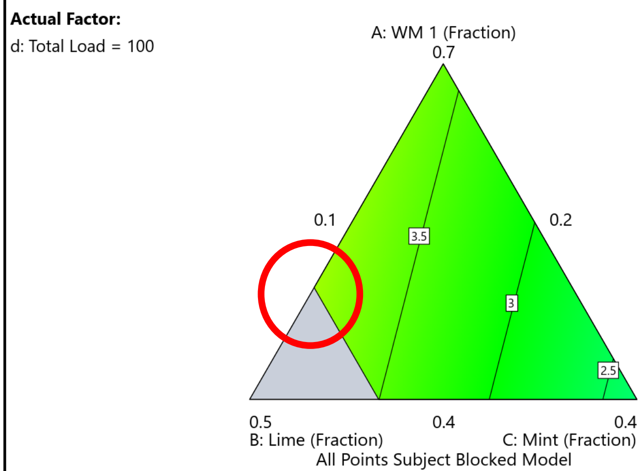
- 70% WM 1, 20% Lime, 10% Mint, and Low Total Load is most optimum.

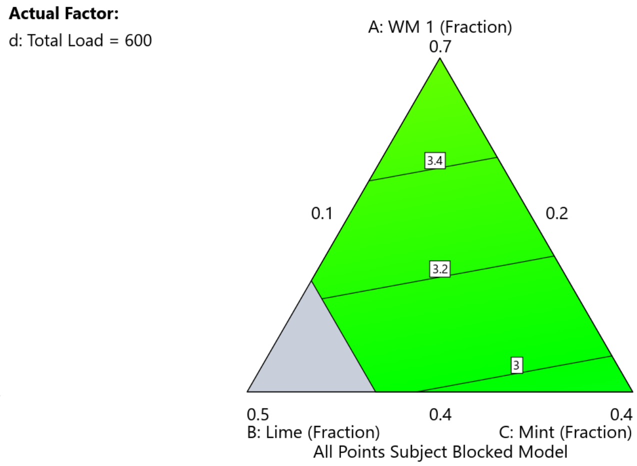
Statistics and Fit (Mean Model)

Source	**p-value**
Subplot	0.0076
Linear Mixture	0.0051
Ad	0.1713
R^2	0.5155
Adjusted R^2	0.3424



Preference (All Individual Points – Subject Blocked)

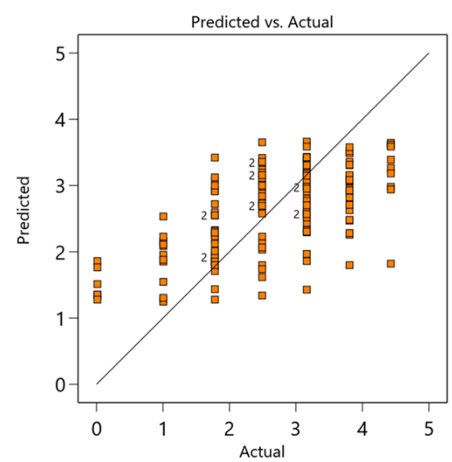




- 50% WM 1, 40% Lime, 10% Mint, and Low Total Load is most optimum.
- Why does this model differ from the Mean model?

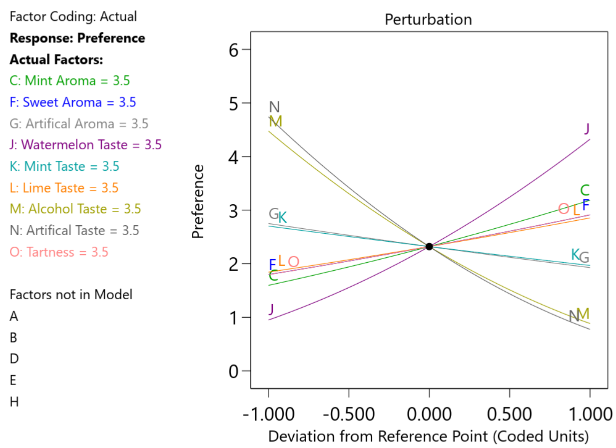
Statistics and Fit (All Individual Points – Subject Blocked/Groups Ignored)

Source	p-value
Model	0.0193
Linear Mixture	0.0088
Bd	0.1660
Cd	0.2811
Lack of Fit	0.8292
R^2	0.0704
Adjusted R^2	0.0472
Predicted R^2	-0.1121
Adequate Precision	10.0890



Note: If group is included, get same model and improved R^2's.

Drivers of Preference

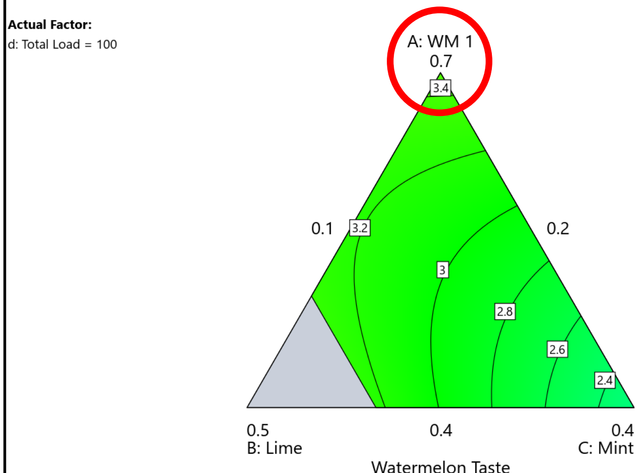


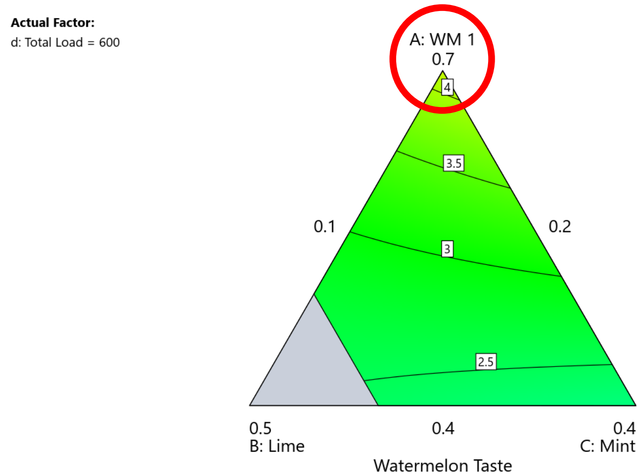
- Preference is modeled as a function of sensory attributes – multiple linear regression.
- Can optimize via maximizing positive drivers and minimizing negative drivers simultaneously.
- **Positive Drivers of Preference** are: Mint Aroma, Sweet Aroma, Watermelon Taste, Lime Taste, and Tartness.
- **Negative Drivers of Preference** are: Artificial Aroma, Mint Taste, Alcohol Taste, and Artificial Taste.

➢ 9 of 14 attributes are important – unusually high.

Watermelon Taste

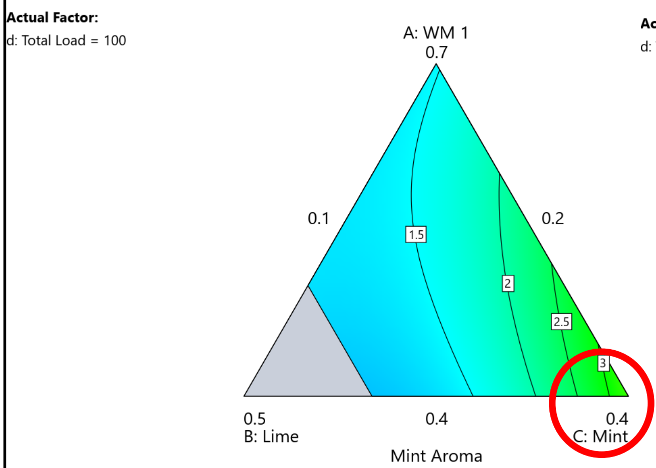
➢ Positive Driver of Preference

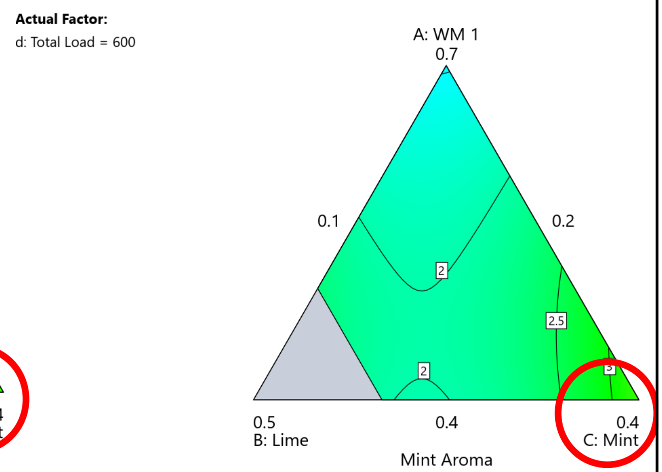




Mint Aroma

➢ Positive Driver of Preference

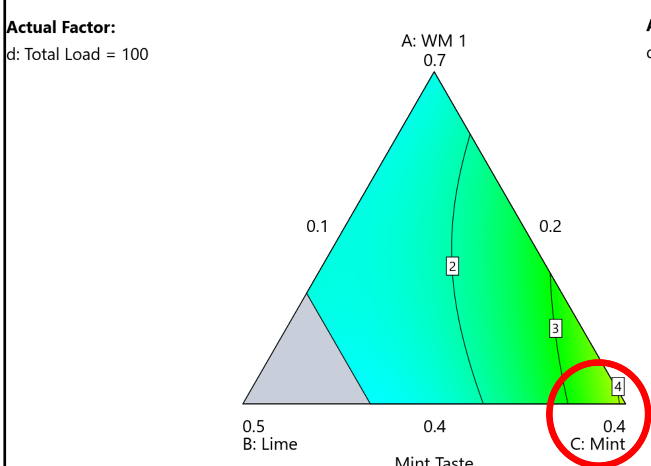


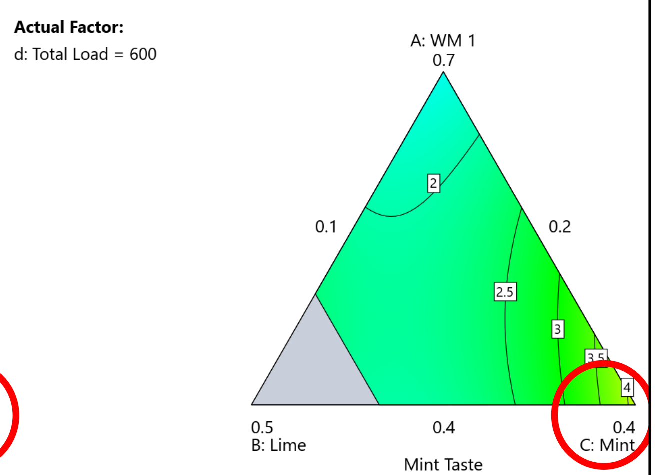


- Higher Mint fraction yields higher Mint Aroma; however, load has no effect in presence of high Mint.

Mint Taste

➢ Negative Driver of Preference

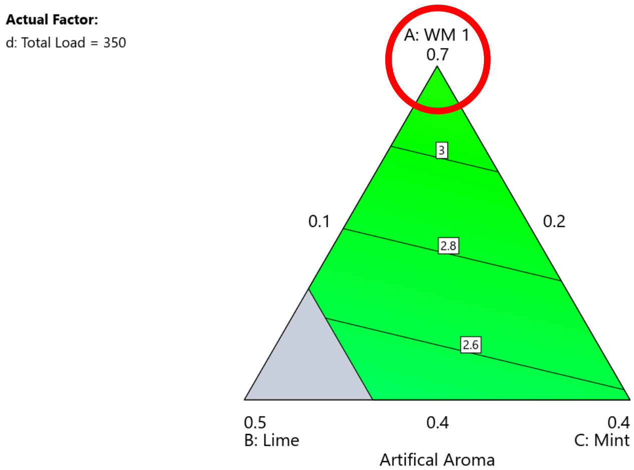




- Similar behavior as Mint Aroma.
- Is Mint overwhelming at high fractions?

Artificial Aroma

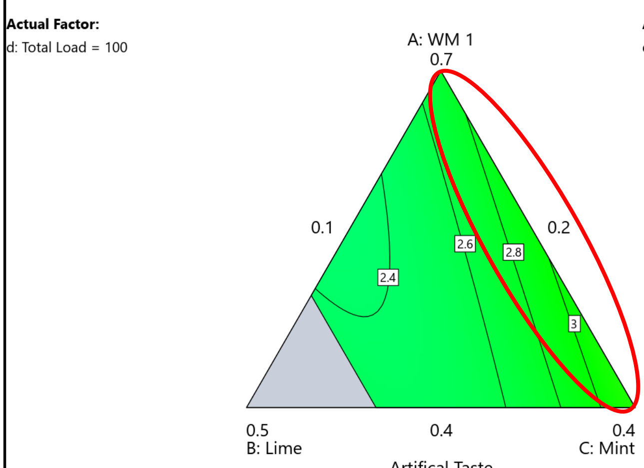
➢ Negative Driver of Preference

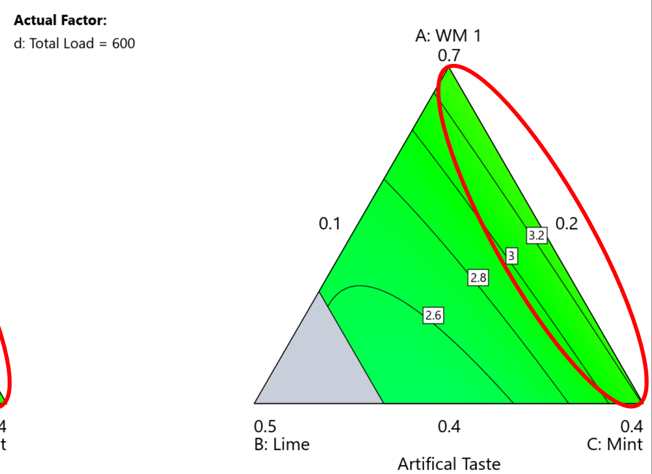


- Load Not Significant

Artificial Taste

➢ Negative Driver of Preference





- Mint and Watermelon increase Artificial Taste.
- Lime reduces Artificial Taste.

Optimums Based on Various Criteria

Preference as a Measure Optimum (*Mean Model*)

WM 1	Lime	Mint	Total Load	Preference
0.7	0.2	0.1	100	3.91

Preference as a Measure Optimum (*Individual Points Model*)

WM 1	Lime	Mint	Total Load	Preference
0.5	0.4	0.1	100	3.94

Drivers of Preference Optimums (*Maximize Positive and Minimize Negative Drivers*)

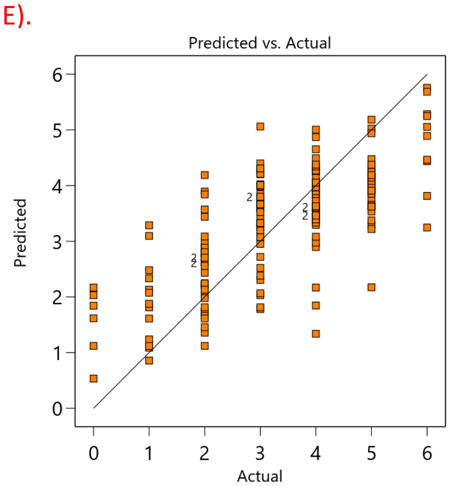
WM 1	Lime	Mint	Total Load	Preference
0.7	0.2	0.1	344	NA
0.4	0.4	0.2	361	NA

- The various approaches yield multiple optimums for further testing.
  - ➢ It's OK. That's the Goal !
- Preference scores at this stage in development are desired to be ≥ 4.
- Polarizing ingredients can drive down overall scores across the whole design space.

Preference Modelling with Subjects Included as a Categoric Variable

- Each subject can have their own equation - enables prediction of individual subject optimums.
- Allows for identification of potentially polarizing factors.
  - ➢ Look for interactions that involve Subject (Factor E).

Source	p-value
Model	< 0.0001
Linear Mixture	0.0077
Ad	0.1344
AE	0.0343
BE	0.0027
CE	0.0013
dE	0.0422
Lack of Fit	0.6783
R^2	0.5228
Adjusted R^2	0.3840
Predicted R^2	0.1327
Adequate Precision	9.5026



Optimums for Individual Subjects

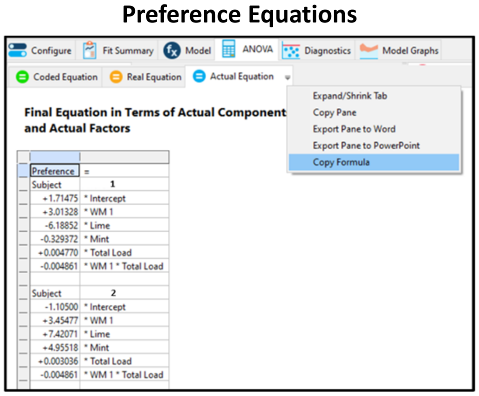
WM 1	Lime	Mint	Total Load	Absolute Mint	Subject	Preference
Fraction	Fraction	Fraction	uL/100ml	uL/100ml		
0.7	0.2	0.1	600	60	1	3.37
0.4	0.4	0.2	600	120	2	4.89
0.5	0.4	0.1	100	10	3	3.98
0.4	0.2	0.4	600	240	4	4.38
0.7	0.2	0.1	100	10	5	4.17
0.7	0.2	0.1	100	10	6	3.33
0.7	0.2	0.1	100	10	7	3.66
0.49	0.4	0.11	100	11	8	6.13
0.7	0.2	0.1	100	10	9	5.25
0.7	0.2	0.1	600	60	10	5.76

- 6 of 10 prefer High Watermelon.
- 7 of 10 prefer Lower Lime.
- 8 of 10 prefer Low Mint - only 1 subject prefers Highest Mint.
- 6 of 10 prefer Low Load.
- Mint has a unique sensory character relative to the other ingredients.
  - Need to think about Absolute Mint Level (Fraction x Load).
    There is a wide range of preferred absolute Mint Levels.

"One Formula to Rule Them All" Optimization

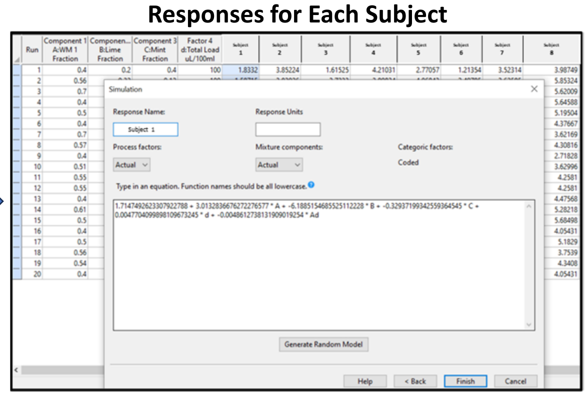
- Subject specific equations from the individual subject model are used in a new model.
  - Each subject's equation is entered as an individual "simulated" response.
  - Numeric optimization is performed maximizing Preference for all subjects simultaneously.

Subject Specific Preference Equations



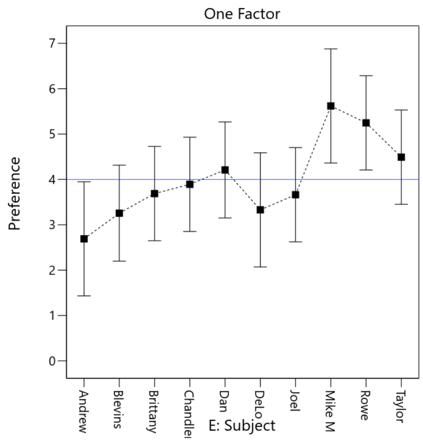
Via Excel

New File with Simulated Preference Responses for Each Subject

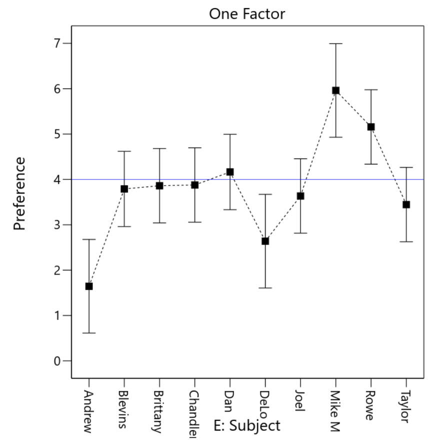


Individual Acceptance of "One Formula to Rule Them All" Optimums

Actual Components:
A: WM 1 = 0.7
B: Lime = 0.2
C: Mint = 0.1
Actual Factor:
d: Total Load = 100



Actual Components:
A: WM 1 = 0.58
B: Lime = 0.32
C: Mint = 0.1
Actual Factor:
d: Total Load = 100

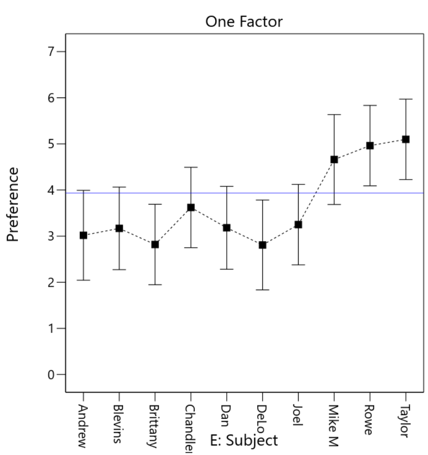


- Same or similar to the Preference as a Measure Optimums (Mean Model and Individual Points Model).

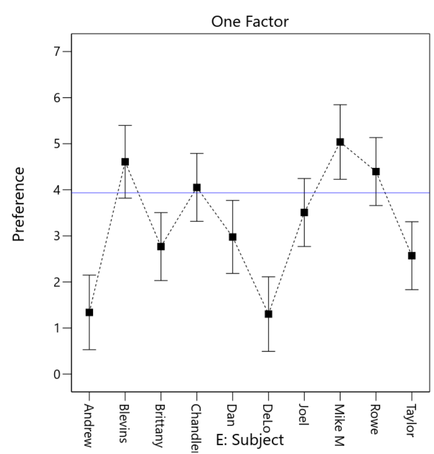
➢ No single formula pleases the greater majority subjects.

Individual Subject Acceptance of Drivers of Preference Optimums

Actual Components:
A: WM 1 = 0.7
B: Lime = 0.2
C: Mint = 0.1
Actual Factor:
d: Total Load = 340

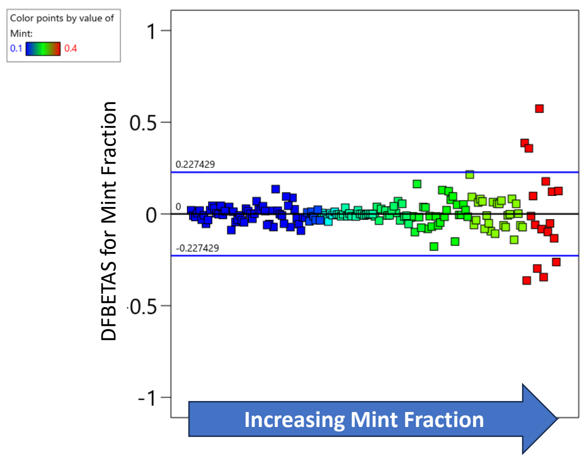


Actual Components:
A: WM 1 = 0.4
B: Lime = 0.4
C: Mint = 0.2
Actual Factor:
d: Total Load = 340



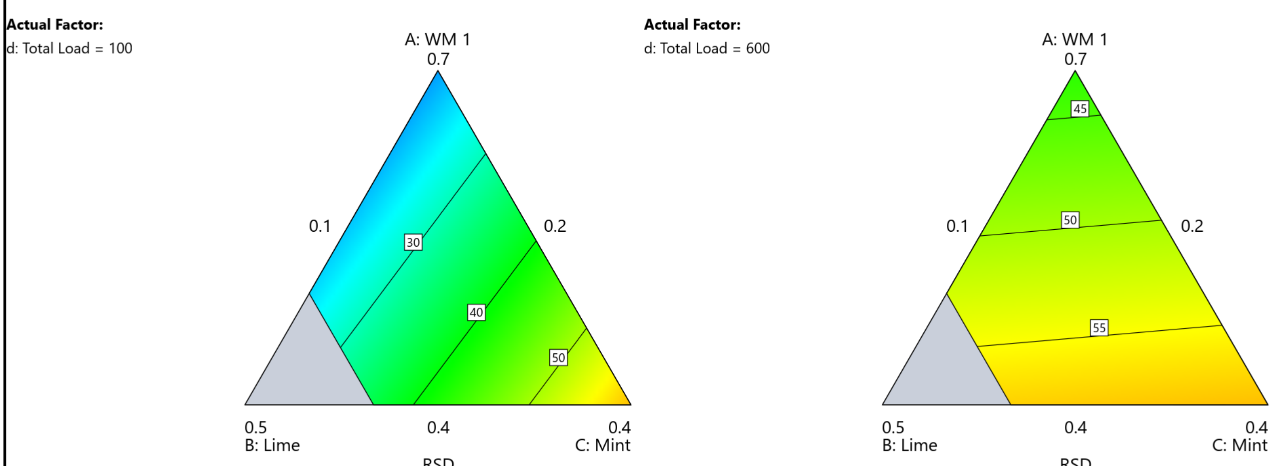
➢ No single formula pleases the greater majority of subjects.

Diagnostic Plots for Identification of Polarizing Ingredients



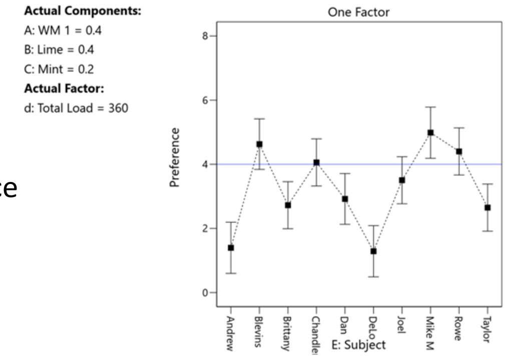
- Individual subjects score Preference very differently for compositions with high Mint.
- The effect is compounded by Total Flavor Load.
- Evidence that Mint is polarizing.

Grading Variation Between Subjects - Modeling of Preference % RSD



- Higher Mint Fraction and Higher Total Load have more variation in grading between subjects.
- Consistent with high levels of Mint being polarizing and overall not desirable for the masses.
- Lime becomes more polarizing at high loads.

Final Formula Selected for Launch



- Can't share exact formula due to confidentially.
- Most resembles one of the Drivers of Preference optimums.

- Selection Logic was as follows:
  - Product concept was intended to be Watermelon-Mint based on flavor trends and other marketing information.
  - "Mint" is on the label and should be noticeable during consumption.
  - It is 1 of 4 flavors in a Variety Pack; it's OK to have a somewhat polarizing product.
  - Preference scores are generally higher outside the context of DOE and employee panels.

➢ Is one of the more popular flavor combinations in the variety pack.

Conclusions

- Results were difficult to interpret due to polarization but were reliable.
- Sensory data is inherently variable, and particularly true for hedonic measures.
- Being *generous but judicious* with model selection is key; however, you need to have knowledge of the subject matter.
- Multiple analysis approaches were crucial in having confidence in optimum selections and identification of Mint as a polarizing ingredient.
- Modeling via subject averages has its place in sensory analysis.
  - However, it is not ideal in representing polarizing study legs.
- The client had more than sufficient data to make a business decision and understood the risks involved with selection of a potentially polarizing product.

MADTREE
Brewery
Cincinnati, Ohio

Special Thanks Go To:

Founders and Owners:

Kenny McNutt and Brady Duncan

Experimental Team:

Ryan Blevins – Head Brewer

Chandler Cottrell – Food Safety and Quality Manager

Taylor Dreves – Quality Lab Technician

Brittany Frey – Production Manager

Sensory Panel Members