Variety Evaluation for Sourdough Baking and Sensory Quality

Lisa Kissing Kucek[1], June Russell[2], Elizabeth Dyck[3], Liz Clark[4], David Benscher[1], Mark E. Sorrells[1], Julie Dawson[5]

[1]Cornell University; [2]Greenmarket, Grow NYC; [3]Organic Growers Research and Information-Sharing Network; [4]Gimme! Coffee; [5]University of Wisconsin-Madison

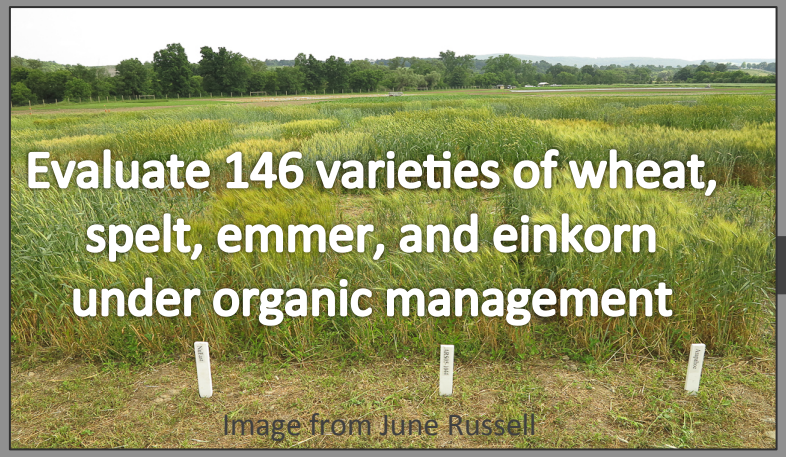
We thank Stefan Senders of Wide Awake bakery for hosting and Jeffrey Hamelman of King Arthur Flour for facilitating the baking trial. We are grateful for the many bakers and tasters who dedicated time and effort in completing this evaluation.

Preliminary Data - 24 January 2015

This work is part of “Value-added Grains for Local and Regional Food Systems”, supported by the Organic Agriculture Research and Extension Initiative

Evaluation Process:

Wheat varieties were screened for use in local organic food systems



Overview of Results

Table 1. Selected wheat variety performance for bread quality (green: higher scoring; red: lower scoring; *indicates statistical significance)

Type	Variety	Market	Variety	Yield	Test Weight	Protein	Baking	Bread Height	Bread Taste	Crumb Texture	Surface Texture	Bread Ability to Dissolve	Bread Graininess	Bread Dryness	Whole Grain Taste	Whole Grain Size
	Name	Class	Age	Rank	Rank	%	Score	cm	Intensity	10=hearty	10=rough	seconds	10=grainy	10=moist	Intensity	10=large
Winter Wheat	Appalachian White	Hard White	Modern	8 of 33	5 of 33	10.4	3.9*	6.5	5.2	6.7	5.6	20.3	5.1	4.5	3.3*	2.5*
	Frederick	Soft White	Modern	11 of 33	22 of 33	9.7	5.5*	5.1*	5.5	7.9*	6.7*	20.7	5.6*	3.8*	4.7	2.8
	Fulcaster	Soft Red	Heritage	27 of 33	15 of 33	10.8	6.2	5.9	5.1	6.9	5.0	19.5	5.3	4.0	4.1	3.0*
	Warthog	Hard Red	Modern	2 of 33	6 of 33	11.1	6.5	8.0*	4.8*	6.6	5.6	20.3	5.4	4.0	5.4*	2.7
Spring Wheat	Red Fife	Hard Red	Heritage	19 of 22	14 of 22	15.3	6.8	6.3	5.7*	6.9	4.8	21.9	4.7	4.8	4.0	2.6
	Tom	Hard Red	Modern	1 of 22	2 of 22	16.7	7.6*	7.4*	5.4	6.5	3.9*	23.5	4.7	4.6	4.2	2.9
	Glenn	Hard Red	Modern	8 of 22	1 of 22	16.0	7.7*	7.5*	5.3	5.4*	3.7*	27.8*	3.9*	5.6*	3.7	2.5*

Results the baking and sensory evaluation were based on only one method and baking session.

Sourdough Baking Trial Results

8 bakers evaluated 7 varieties in replicate

- **Glenn and Tom:** top categories for baking quality, height, and weight
- **Warthog:** intermediate for baking and weight, top category for height
- **Red Fife and Fulcaster:** intermediate in most categories
- **Appalachian White:** second lowest for baking, poor weight
- **Frederick:** lowest for baking, height, and weight

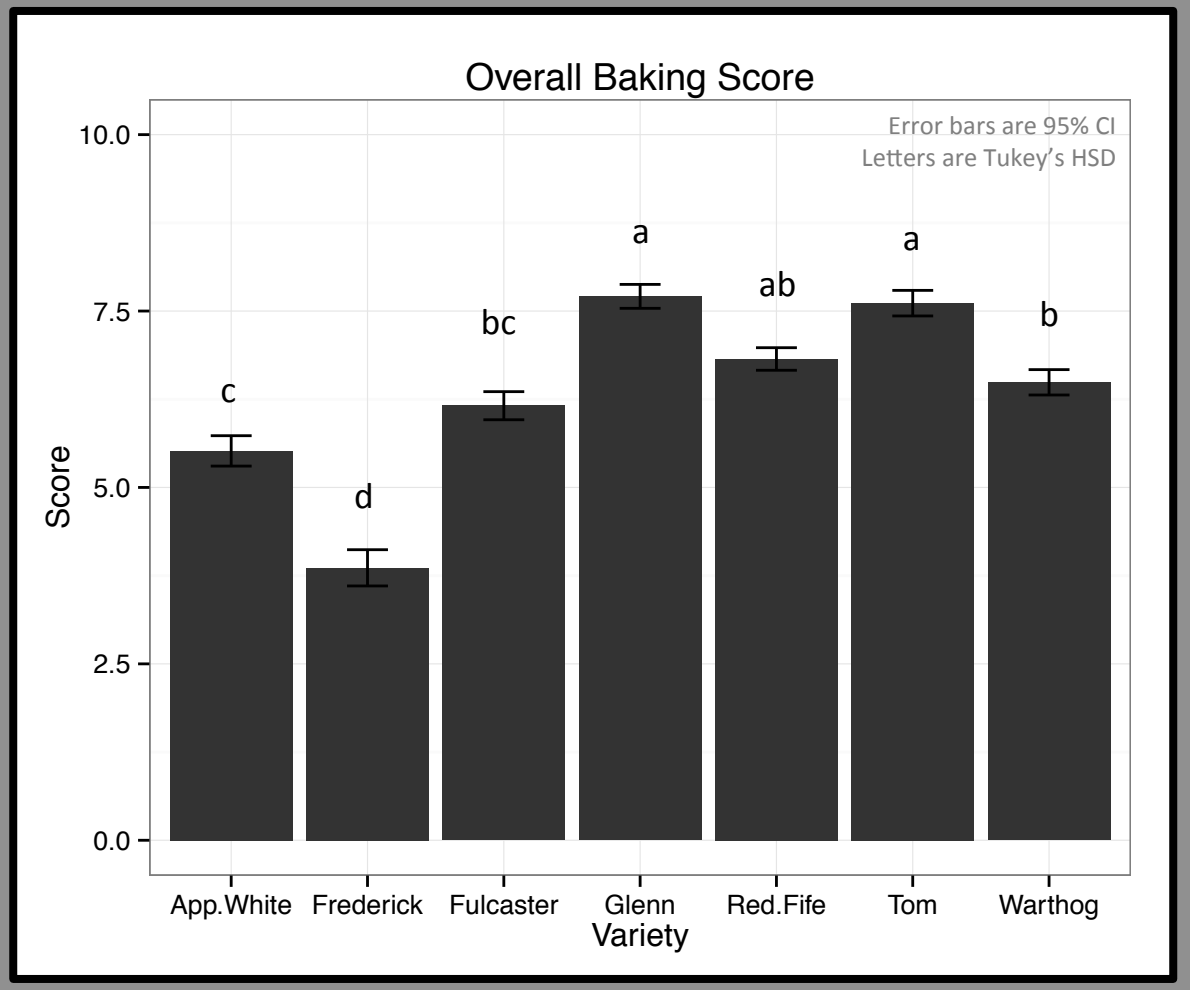
Type III ANOVA with Sattherwaite approximation

$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu_6 = \mu_7$; $\alpha \leq 0.05$

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$$

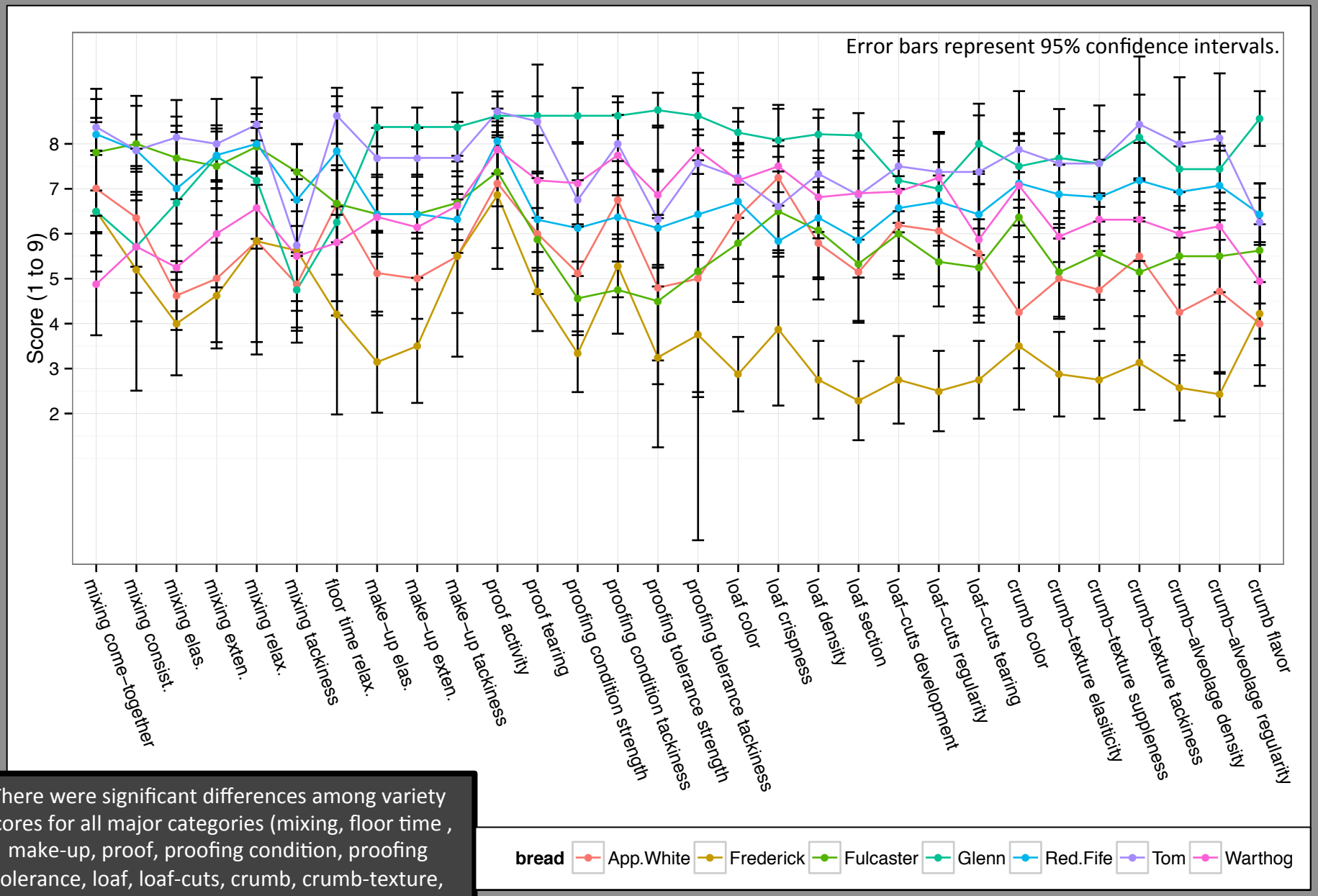
y_{ij}: response for variety i and baker j
μ: overall mean response
α_i: fixed effect of variety i
B_j: random effect of baker j
ε_{ijkl}: experimental error associated with response I,j

To validate model assumptions, errors and random effects were checked for normal distribution, homogeneous variance, and independence.



There were significant differences in scores among varieties at $p<0.0001$. n=1567

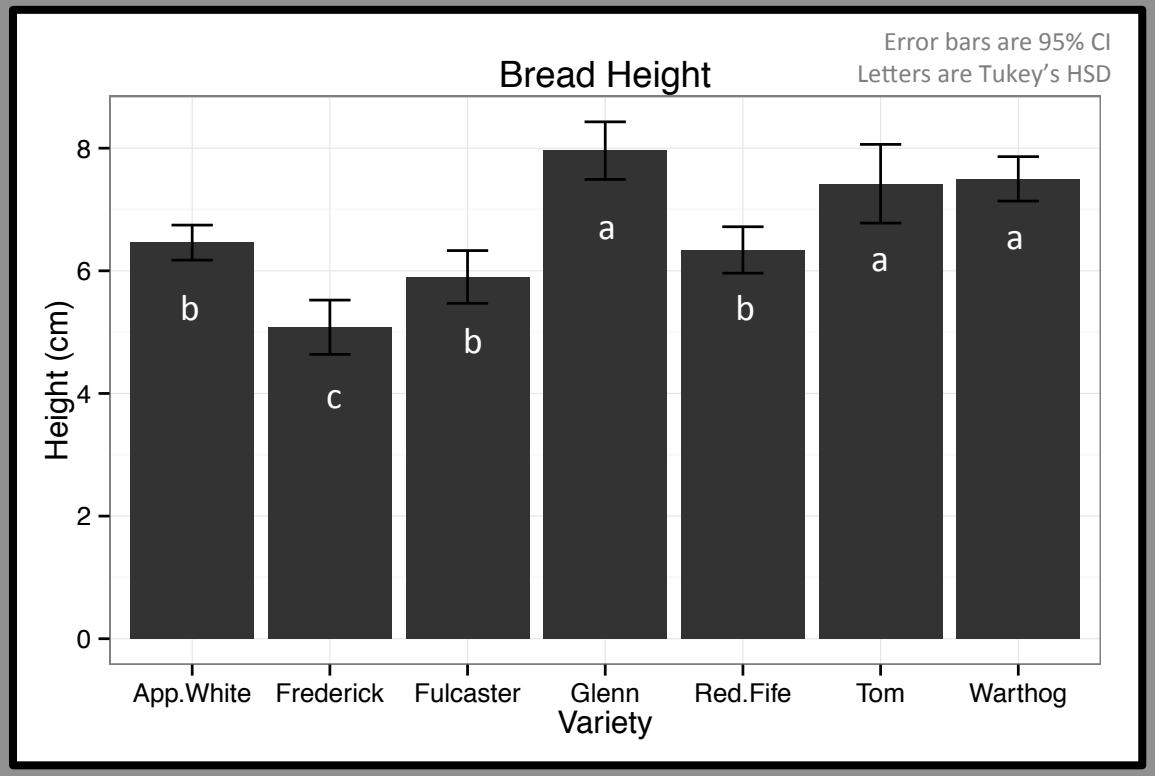
Sourdough Baking Trial Results

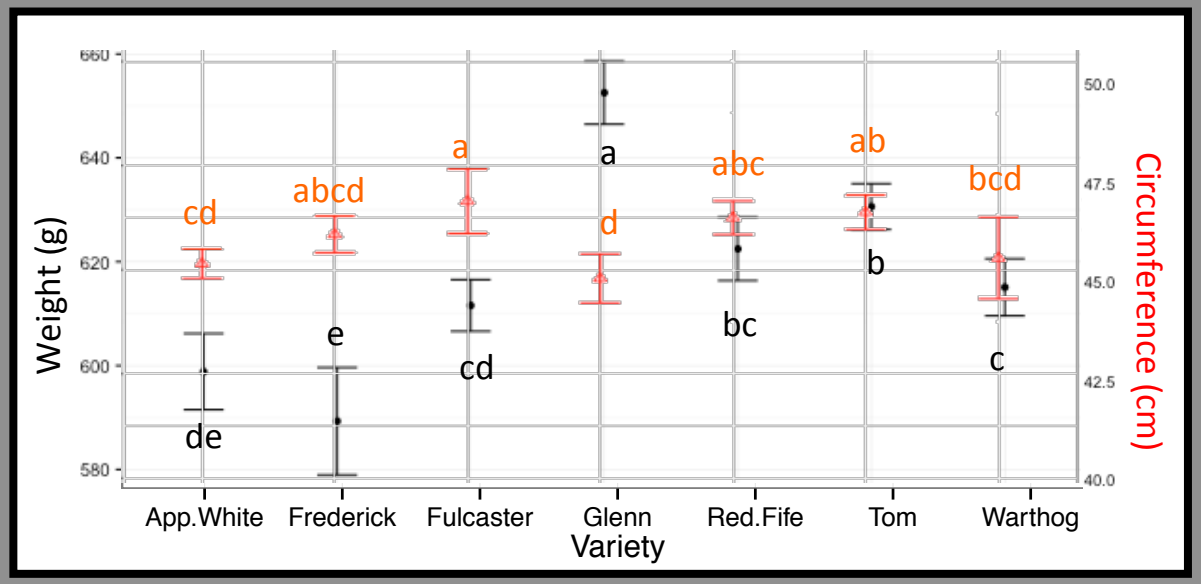


There were significant differences among variety scores for all major categories (mixing, floor time , make-up, proof, proofing condition, proofing tolerance, loaf, loaf-cuts, crumb, crumb-texture, and crumb-alveolage ($p<0.0001$)

Figure to the right:
There were significant differences in height among varieties at p<0.0001. n=35

Figure below:
There were significant differences in weight and circumference among varieties at p<0.0001. n=81





Not shown:
Volume (p=0.109) and density (p=0.33) of loaves were not significantly different among varieties. n=21

Sourdough Sensory Evaluation

30 tasters evaluated 7 varieties over 2 replicates

- **Glenn:** smoothest surface texture category and most delicate crumb texture, longest time to dissolve, lowest graininess, highest moisture
- **Red Fife:** taste intensity higher than Warthog, earthier flavors
- **Tom:** smoothest surface texture category, largest air bubble size
- **Fulcaster:** second highest air bubble size
- **Appalachian White:** smallest air bubble size
- **Warthog:** taste intensity lower than Red Fife
- **Frederick:** roughest surface texture, most hearty crumb texture, highest graininess, driest bread

Type III ANOVA with Sattherwaite approximation

$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu_6 = \mu_7$; $\alpha \leq 0.05$

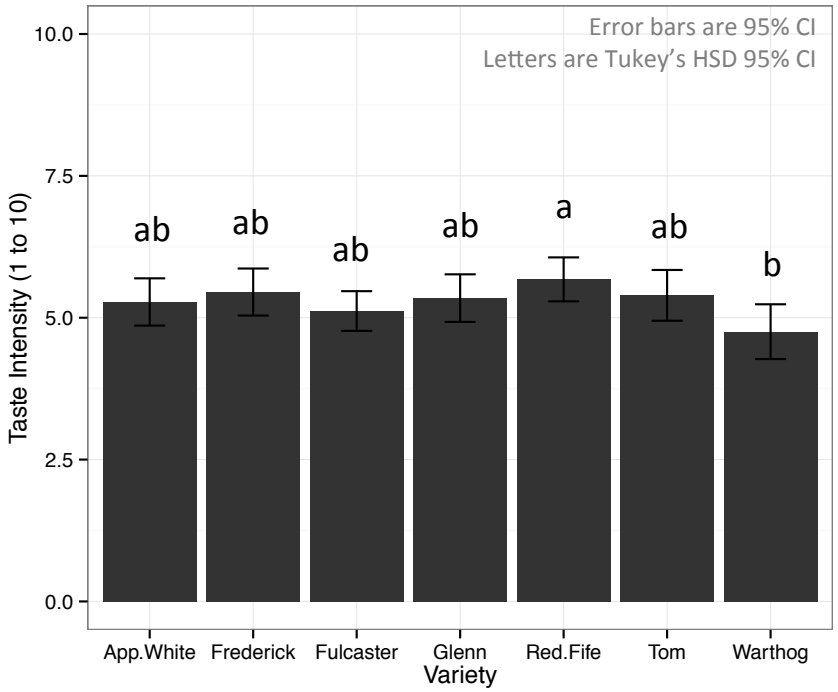
$$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \varepsilon_{ijk}$$

y_{ij}: response for variety i, rep j, order k, and subject l
μ: overall mean response
α_i: fixed effect of variety i
B_j: fixed effect of rep j
γ_k: random effect of subject k
ε_{ijk}: experimental error associated with response I,j,k

To validate model assumptions, errors and random effects were checked for normal distribution, homogeneous variance, and independence.

Taste Intensity

(1 = no flavor, 10 = intense) n=380

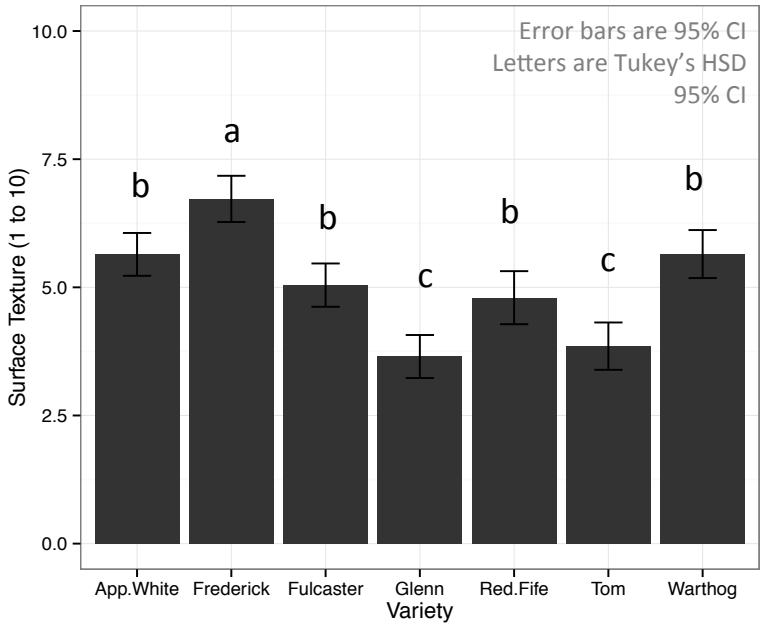


There were significant differences in taste intensity among varieties at p=0.021. Subject accounted for 17.68% of variation.

Sourdough Sensory Evaluation

Surface Texture

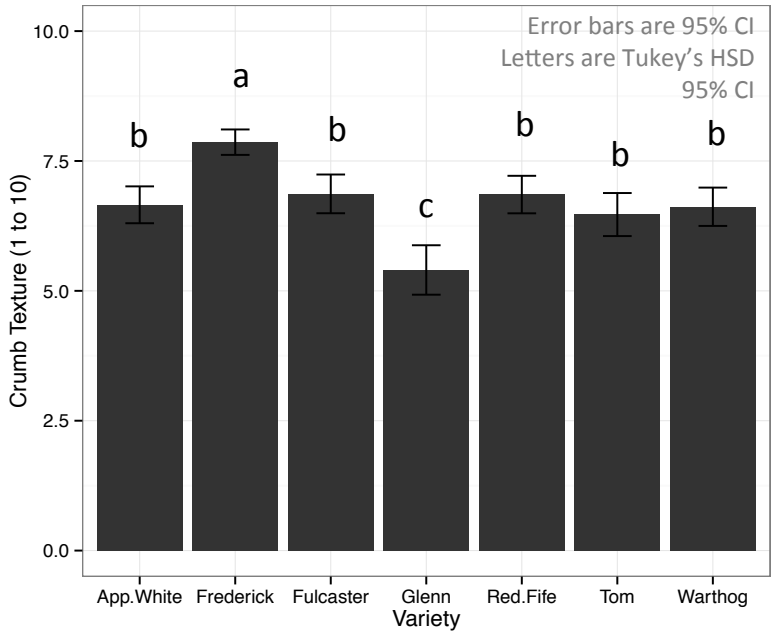
(1 = even and smooth, 10 = heavily textured)
N=400



There were significant differences among varieties at p<0.0001. Subject accounted for 11% of variation.

Crumb Texture

(1 = delicate, 10 = most hearty)
n=402



There were significant differences among varieties at p<0.0001. Subject accounted for 21% of variation.

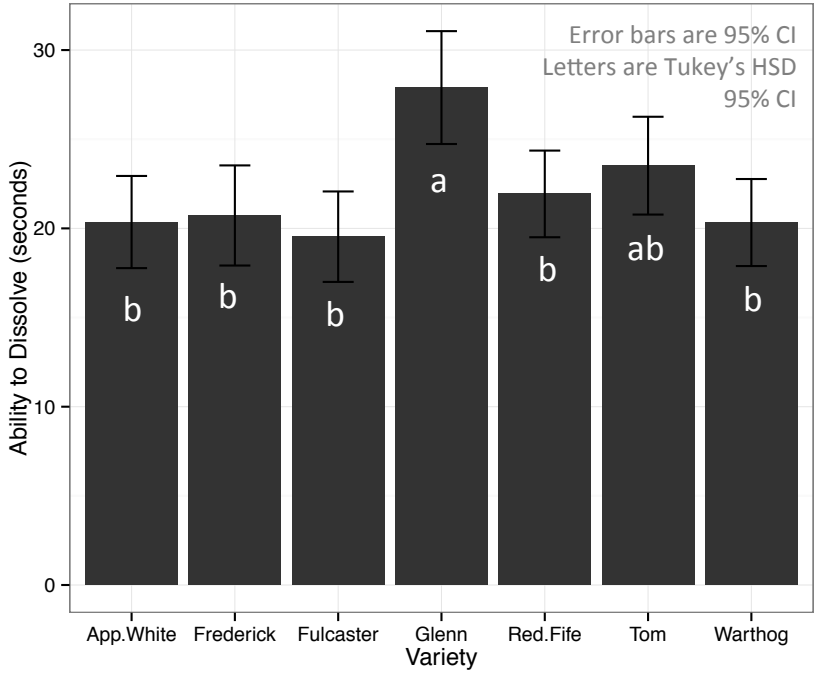
Although there were no significant differences between varieties for any aromatics categories, replicate number influenced aromatics. Replicate 2 produced significantly higher values for whole sample aromatics (p=0.0134), crust aromatics (p=0.0242), and crumb aromatics (p=0.0341). Subject accounted for 22.18%, 39.66%, and 34.01% of variation, respectively.

Sourdough Sensory Evaluation

Texture: ability to dissolve

(seconds to dissolve in mouth)

n=400

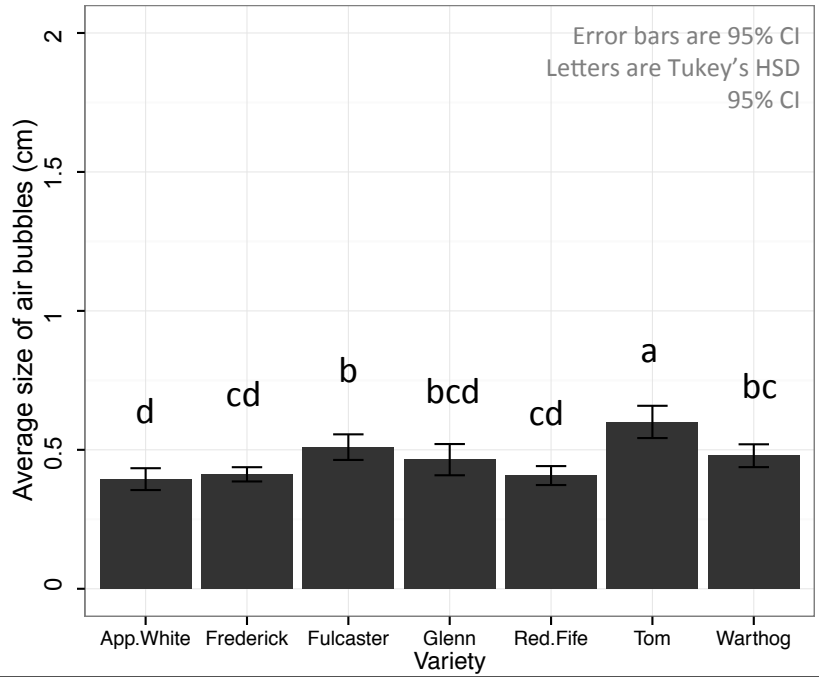


There were significant differences in time to dissolve among varieties at p<0.0001. Subject accounted for 56.47% of variation.

Air Bubble Size

(centimeters)

n=411



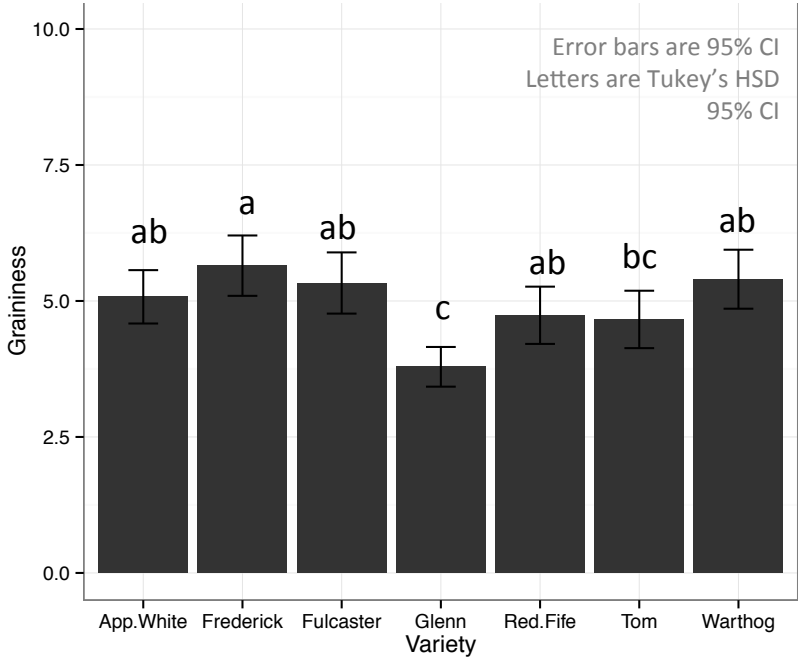
There were significant differences in reported average air bubble size among varieties at p<0.0001. Subject accounted for 16.53% of variation.

Sourdough Sensory Evaluation

Graininess: amount of small particles

(1=no graininess, 10 = overwhelming graininess)

n=397

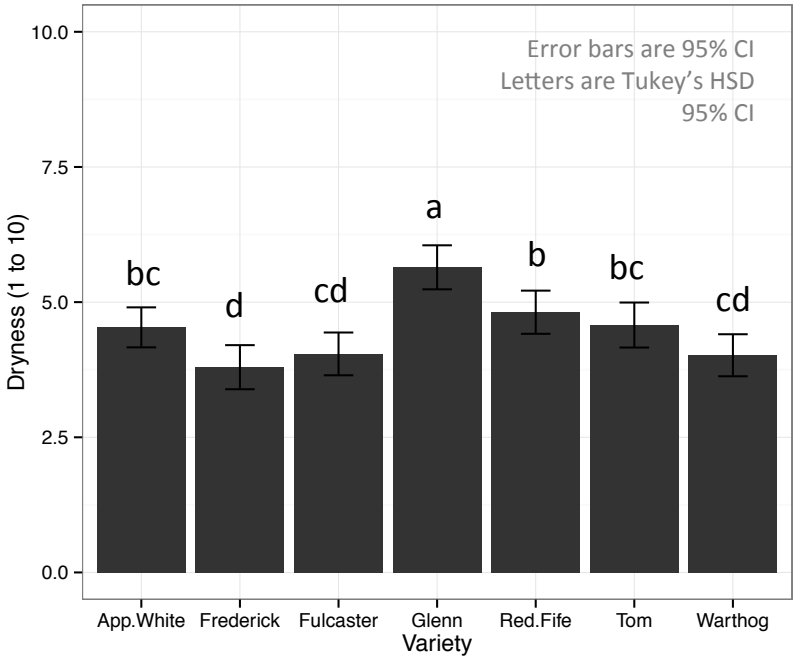


There were significant differences in graininess among varieties at p<0.0001. Subject accounted for 42.46% of variation.

Dryness: saliva taken from tongue

(1 = very dry, 10 = moist)

n=414



There were significant differences in dryness among varieties at p<0.0001. Subject accounted for 32.81% of variation.

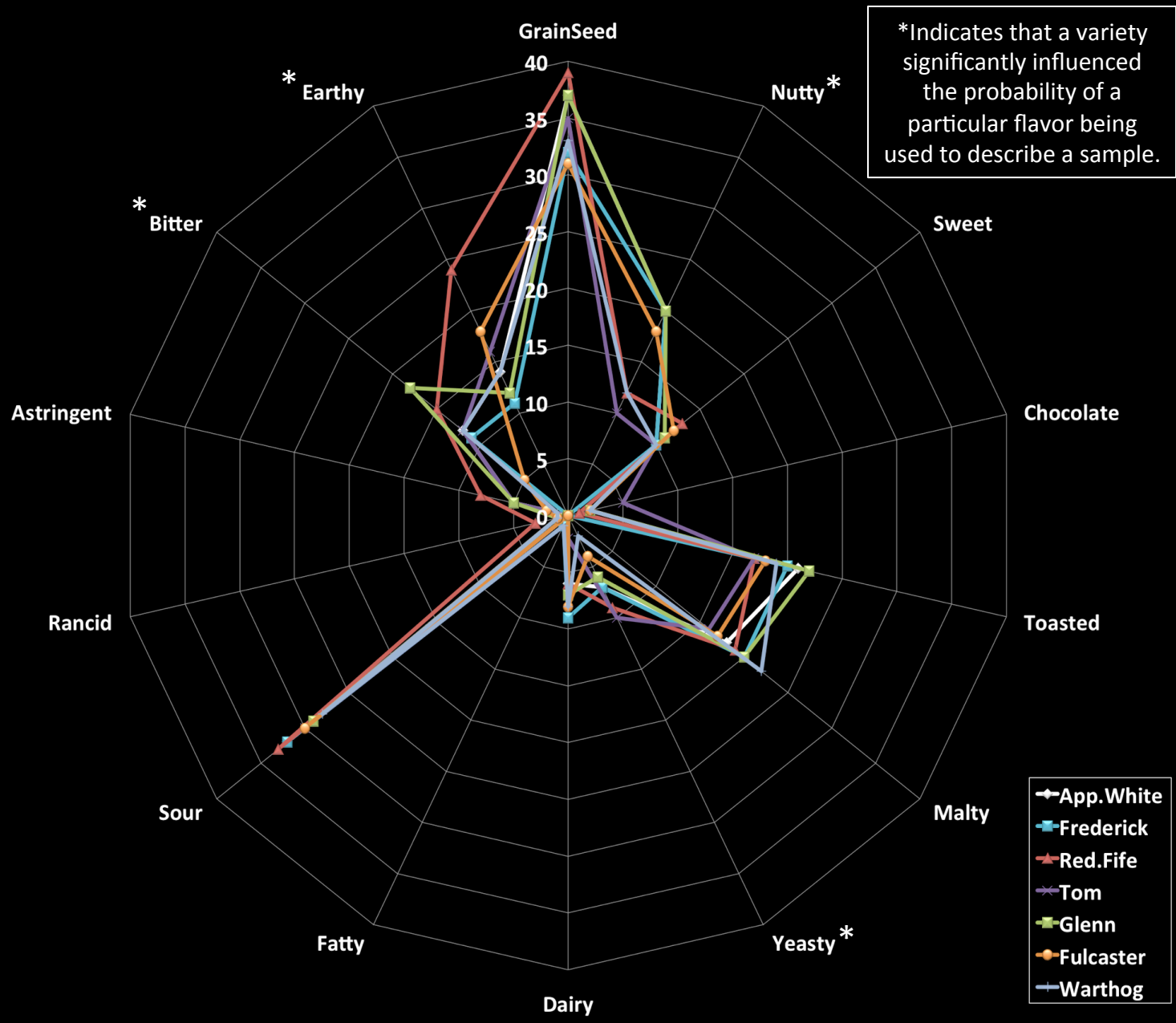
Sourdough Sensory Evaluation

Tom (p=0.024), Red Fife and Warthog (p= 0.073) lowered the odds for nutty flavors.

Warthog lowered the odds for yeasty flavors (p=0.060).

Fulcaster lowered the odds for bitter flavors (p=0.042).

Red fife increased the odds for earthy flavors (p=0.035).



*Indicates that a variety significantly influenced the probability of a particular flavor being used to describe a sample.

Wald χ^2 test binomial distribution

H_0: $\beta_1=0$; $\alpha \leq 0.10$

$Y_{ijk} = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3}$

Y_{ij}: log odds of a flavor used for sample
B_0: intercept log odds App.White rep 1
β_1: partial slope associated with variety
x_{i1}: fixed variable of variety i
B_2: partial slope associated with rep
x_{i2}: fixed variable of rep i
β_3: partial slope associated with taster
x_{i3}: random variable of taster I

To validate model assumptions, $n*\pi>5$ and $n*(1-\pi)>5$. n: number of observations; π: sample probability mean. See final slide for more details.

Cooked Whole Grain Sensory Evaluation

30 tasters evaluated 7 varieties over one replicate

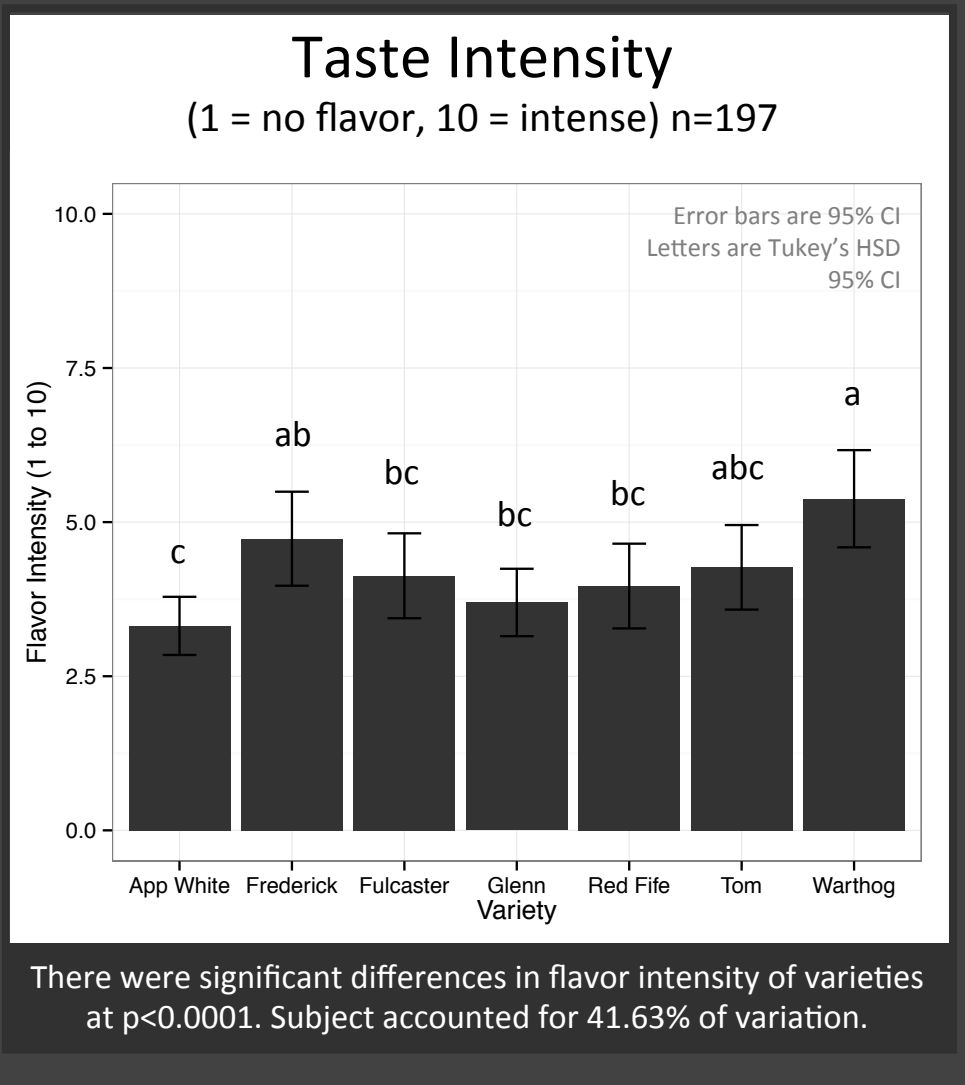
- **Warthog:** most intense flavor, sweeter and less grainy/seedy flavors
- **Red Fife and Fulcaster:** nuttier flavors
- **Frederick:** yellow color and more dairy flavors
- **Glenn:** less nutty and less sweet flavors
- **Tom:** intermediate in all categories
- **Appalachian White:** least intense flavor

Type III ANOVA with Sattherwaite approximation
H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu_6 = \mu_7$; $\alpha \leq 0.05$

$$Y_{ijk} = \mu + \alpha_i + \beta_j^* + \gamma_k + \varepsilon_{ijk}$$

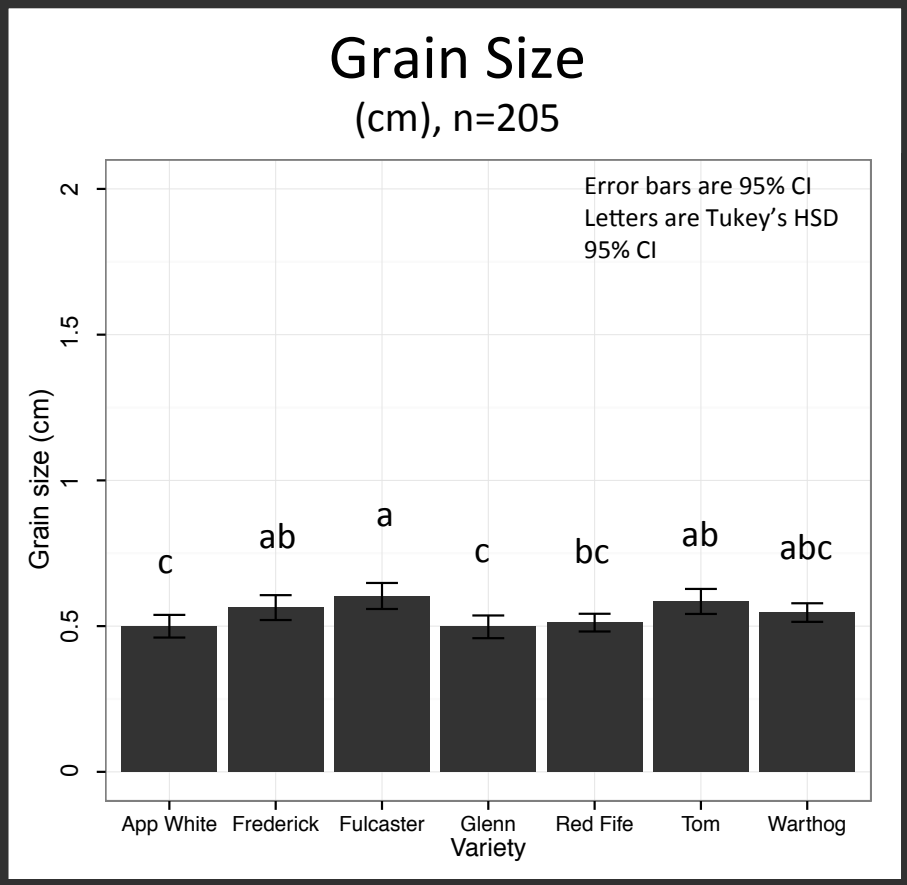
y_{ij}: response for sample i, order k, subject l
μ: overall mean response
α_i: fixed effect of variety i
β_j: fixed effect of order j*
γ_k: random effect of subject k
ε_{ijk}: experimental error associated with response I,j,k

To validate model assumptions, errors and random effects were checked for normal distribution, homogeneous variance, and independence.
*Only included in the dryness model. See last slide for details.



There were significant differences in flavor intensity of varieties at $p<0.0001$. Subject accounted for 41.63% of variation.

Cooked Whole Grain Sensory Evaluation



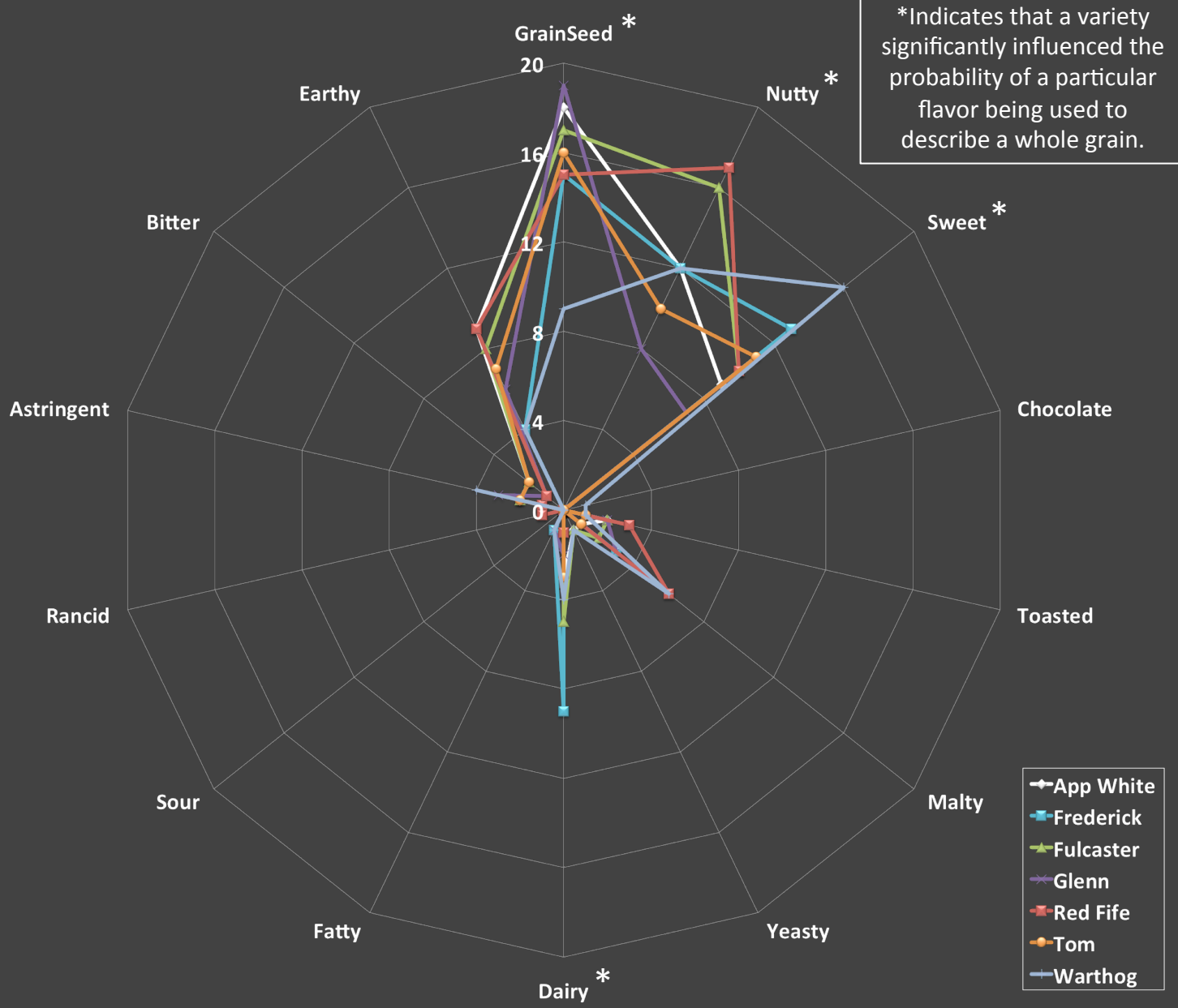
There were significant differences in grain size of varieties at $p<0.0001$. Subject accounted for 60.02% of variation.

Whole grain dryness was not significantly different by variety ($p=0.946$). The first sample tasted (order=1) was reported to have significantly higher moisture ($p=0.0434$).

Cooked Whole Grain Sensory Evaluation

Warthog lowered odds for grainy/seedy flavors (p=0.091).
Glenn lowered (p=0.035), while Red Fife (p=0.036) and Fulcaster (p=0.074) increased odds for nutty flavors.
Glenn significantly lower odds for sweet flavors (p=0.067).
Frederick increased odds for dairy flavors (p=0.002).

*Indicates that a variety significantly influenced the probability of a particular flavor being used to describe a whole grain.



Wald χ^2 test binomial distribution
H_0: $\beta_1=0$; $\alpha \leq 0.10$
$Y_{ijk} = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3}$

Y_{ij}: log odds of a flavor used for sample
B_0: intercept log odds
β_1: partial slope associated with variety
x_{i1}: fixed variable of variety i
B_2: partial slope associated with order
x_{i2}: fixed variable of order i
β_3: partial slope associated with taster
x_{i3}: random variable of taster I

To validate model assumptions, $n*\pi>5$ and $n*(1-\pi)>5$. n: number of observations; π: sample probability mean.

Details of Statistical Models

Analyses completed in R and JMP. Order, an ordinal variable from 1 to 7, was a candidate to be included as a covariate in all models. However, order was not linearly related to the responses evaluated in the models, and consequently, violated the assumptions of an ANCOVA model. Despite randomization, some samples are overrepresented in certain orders (e.g. Red Fife in order 2). When the model was run for samples that were balanced, there was not a significant effect for order. Order was only included as a fixed effect in the analysis of whole grain dryness, to interpret deviations found between the first sample tasted and all other orders.

Baking evaluation R code:

```
model=lmer(Y~Variety+(1|baker))
summary(bakemodel)
anova(bakemodel,Type=3)
```

Bread sensory R code:

```
model=lmer(Y~Variety+Rep+(1|Subject))
summary(model)
anova(model, Type=3)
```

Bread flavor descriptors R code:

```
model=glmer(Y~Variety+Rep+ (1|Subject),
control=glmerControl(optimizer="bobyqa",optCtrl=list(maxfun=100000)),family="binomial",data=sens)
summary(model)
anova(model)
```

Whole Grain taste and size R code:

```
Ymodel=lmer(Y~Variety+(1|Subject))
```

Whole Grain dryness R code:

```
model=lmer(Y~Variety+Order+(1|Subject))
```

Whole Grain flavor descriptors R code:

```
model=glmer(Y~Variety+(1|Subject), control=glmerControl(optimizer="bobyqa",optCtrl=list(maxfun=100000)),family="binomial",data=sens)
summary(model)
anova(model)
```